The Journal of
Economic Perspectives

A journal of the
American Economic Association

Fall 2023

The Journal of
Economic Perspectives

Contents

Symposia

Statement of Purpose

The *Journal of Economic Perspectives* aims to bridge the gap between the general interest business and financial press and standard academic journals of economics. The journal aims to publish articles that will serve several goals: to synthesize and integrate lessons learned from active lines of economic research; to provide economic analysis of public policy issues; to encourage cross-fertilization of ideas among the fields of economics; to offer readers an accessible source for state-of-the-art economic thinking; to suggest directions for future research; to provide insights and readings for classroom use; and to address issues relating to the economics profession. Articles appearing in the journal are normally solicited by the editors and associate editors. Proposals for topics and authors should be directed to the journal office, at the address inside the front cover.

Policy on Data Availability

It is the policy of the *Journal of Economic Perspectives* to publish papers only if the data used in the analysis are clearly and precisely documented and are readily available to any researcher for purposes of replication. Details of the computations sufficient to permit replication must be provided. The Editor should be notified at the time of submission if the data used in a paper are proprietary or if, for some other reason, the above requirements cannot be met.

Policy on Disclosure

Authors of articles appearing in the *Journal of Economic Perspectives* are expected to disclose any potential conflicts of interest that may arise from their consulting activities, financial interests, or other nonacademic activities.

Why Did the Best Prepared Country in the World Fare So Poorly during COVID?

Jennifer B. Nuzzo and Jorge R. Ledesma

Though the timing and details of the COVID-19 pandemic caught governments around the world by surprise, the possibility of a new pathogen emerging to cause a deadly pandemic had been long understood by experts in public health preparedness. Even in the months prior to the outbreak of the coronavirus in February 2020, several expert groups warned governments about the potential imminent threat of pandemics and urged them to bolster their preparedness for such events. The Global Preparedness Monitoring Board (2019), a high level independent group of national and global health leaders, published a report called *World at Risk*, in which it concluded that it was "well past time to act" to reduce global vulnerabilities to pandemics and other health emergencies (Global Preparedness Monitoring Board 2019). The Global Health Security (GHS) Index—a framework that measures the readiness of 195 countries for pandemics and other significant biological threat emergencies—determined that "no country was fully prepared" for a potential pandemic (Cameron, Nuzzo, and Bell 2019). Months later, these conclusions became self-evident with the emergence and rapid spread of the SARS-CoV-2 virus that caused the COVID-19 pandemic.

But while the lack of national readiness for a pandemic was anticipated, it was hard to anticipate the extent to which the earlier measures of pandemic preparation seemed to have little connection to the later outcomes. For example, the Global

■ *Jennifer B. Nuzzo is Professor of Epidemiology and Director of the Pandemic Center, Brown University School of Public Health, Providence, Rhode Island. Jorge R. Ledesma is an Epidemiology PhD candidate, Brown University School of Public Health, Providence, Rhode Island. Their email addresses are jennifer_nuzzo@brown.edu and jorge_ledesma@brown.edu.*

Figure 1

Global Distribution of Global Health Security (GHS) Index Scores in 2021
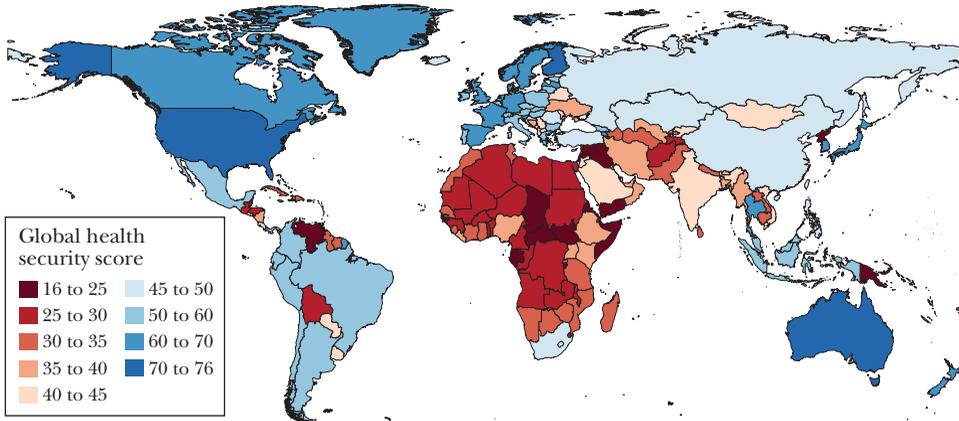
Source: Data are from the 2021 Global Health Security (GHS) Index report.
Note: The Global Health Security (GHS) Index is a comprehensive assessment and benchmarking of 195 countries' readiness for pandemics and other significant biological threat emergencies. The index ranges from 0 to 100 with lower scores indicating weaker health system capacities and higher scores suggesting stronger health system capacities for pandemic readiness. In 2021, the United States had the highest GHS Index score at 75.9.

Health Security Index ranked the United States as the best-prepared country in the world overall, as illustrated by the map in Figure 1. This assessment was based on a variety of measurements grouped in six broad categories: (1) prevention of the emergence or release of pathogens; (2) early detection and reporting for epidemics of potential international concern; (3) rapid response to and mitigation of the spread of an epidemic; (4) sufficient and robust health sector to treat the sick and protect health workers; (5) commitments to improving national capacity, financing, and adherence to international norms; and (6) overall risk environment and country vulnerability to biological threats. But when the pandemic hit, the United States struggled to respond despite its prepandemic advantages (Nuzzo, Bell, and Cameron 2020).

In total, COVID-19 has caused upwards of 6.5 million reported global deaths as of May 2023. The United States, which makes up less than 5 percent of the world's population, accounts for more than 15 percent of the officially reported COVID deaths.

However, the United States has so far failed to undertake a rigorous and comprehensive audit of its pandemic mistakes and challenges. Various academic analyses have attempted to dissect the US COVID-19 experience. Most of these analyses compare disease outcomes in US states or counties with differing COVID-19 policies. Such ecological analyses can help us generate hypotheses as to which actions may explain why disease rates may have differed throughout the country, or across countries, but they do not offer evidence of causality. Instead, causality is obscured by the fact that many communities implemented multiple policies and actions. Moreover,

communities that implemented more aggressive public health interventions likely differed in a number of ways from those communities that chose not to. Unfortunately, much of the existing literature fails to tease apart these important differences.

Collectively, the existing literature suggests that the United States mounted a response that failed to make full use of the preparedness capacities it had, was hampered by politics, made poor use of data, and neglected to overcome intrinsic social vulnerabilities that helped the virus spread and caused high mortality. Below summarizes what we think we know about the US experience during COVID-19 and where the data point toward identifying the source of the country's profound challenges.

The Tolls of COVID-19 in the United States

The US health tolls of the pandemic have been staggering. Some skeptics have tried to put the pandemic in context by comparing it to other routine infectious diseases, such as influenza. But the SARS-CoV-2 virus that causes COVID-19 has differentiated itself in terms of the numbers of infections and deaths it has caused. The Centers for Disease Control and Prevention (2022) estimates that between 2010–2020, seasonal influenza killed between 12,000 and 52,000 Americans per year. From the start of the pandemic in February 2020 through the end of 2022, COVID-19 killed more than 1.1 million Americans.

This cumulative mortality of the SARS-CoV-2 virus has been large enough that in 2020 and 2021, US life expectancy dropped for two consecutive years, reaching a low not seen since the mid-1990s (Arias et al. 2022). Indeed, the two-year decline was the biggest the United States had seen since 1921–1923. Most of these declines in life expectancy that occurred in the first two years of the pandemic were directly caused by the pandemic, as a result of increased deaths due to COVID-19. The Centers for Disease Control estimates that COVID-19 deaths contributed to 74 percent and 50 percent of the decline in life expectancy in 2020 and 2021, respectively (Arias et al. 2022).

Pandemic mortality has not been experienced equally across the United States. Throughout the pandemic, there were stark racial/ethnic disparities in the case, hospitalization, and death data, which translated to disproportionate declines in life expectancy. Overall, life expectancy decreases during 2020 and 2021 were largest among non-Hispanic Native Americans, followed by Hispanic-origin and Black people, as shown in Figure 2 (Arias et al. 2022). In the second year of the pandemic, however, these trends shifted. All racial and ethnic groups experienced declining life expectancy in 2020 and 2021, and also experienced the biggest decreases in life expectancy in the first year of the pandemic. However, the improvements in declining life expectancy were smallest for whites, compared to other groups, after safe and effective COVID-19 vaccines became available.

Again, COVID-19 was the leading contributor to declines in life expectancy in all racial/ethnic groups and among men and women during the pandemic. But

Figure 2

Temporal Trends in Life Expectancy at Birth by Racial/Ethnic Groups in the United States, 2006–2021

Panel A: Estimates of Life Expectancy at Birth, by Race, 2006–2021

Panel B: Change in Life Expectancy at Birth from the Previous Year
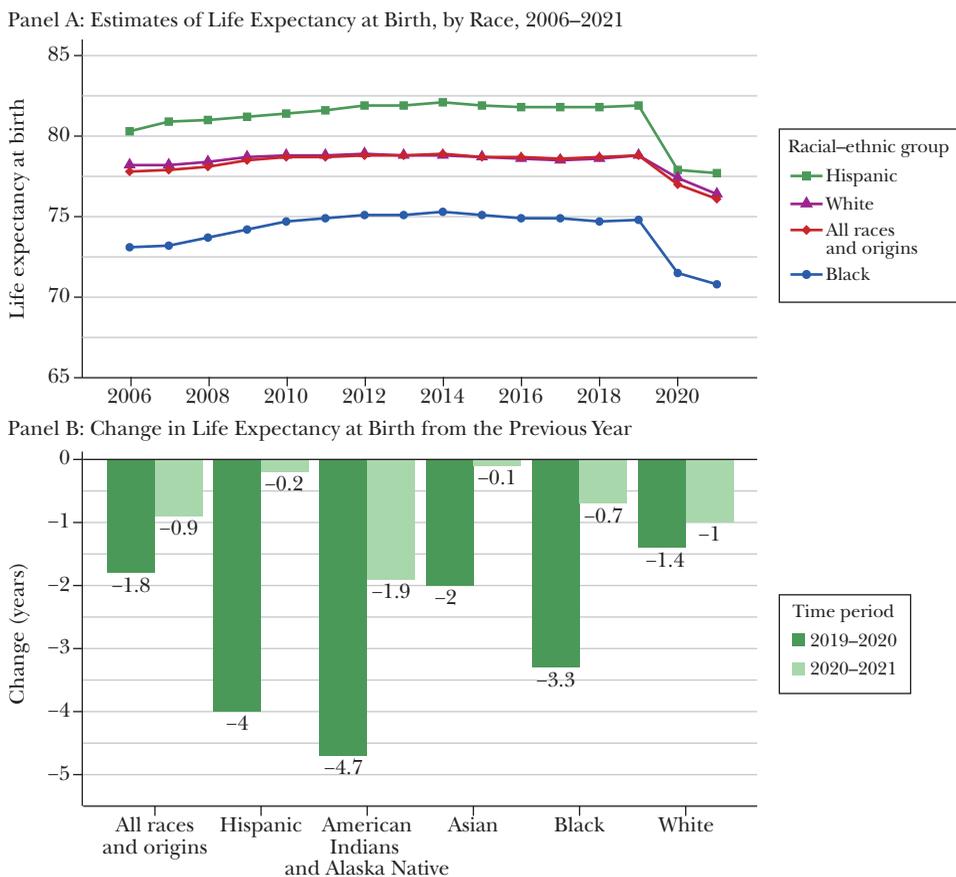
Source: Data from the National Center for Health Statistics, United States Centers for Disease Control and Prevention.
Note: Panel A visualizes annual life expectancy at birth estimates by race from 2006 to 2021. Data for American Indians and Alaska Natives (AIAN) and Asian not available from 2006 to 2018. Panel B illustrates the change in life expectancy at birth by comparing the year of interest to the year prior. All Hispanics are included in the "Hispanic" category, so other groups should be interpreted as "Non-Hispanic Whites," "Non-Hispanic Blacks," and so on.

aside from its direct harms, the pandemic may have contributed to a decline in life expectancy due to other causes.

For example, in the first two years of the pandemic, the second-leading contributor to declining life expectancy were nonintentional injuries, a category that includes overdose deaths. During this period, drug-involved overdose deaths rose sharply (National Institute on Drug Abuse 2023). An increase in substance use and disruptions in mental health and substance-use treatment during the pandemic

may be reasons for the observed increase in overdose deaths. Substance use appears to have increased during the pandemic, particularly during the first year (Panchal et al. 2023). Similarly, health claims data showed a decrease in utilization of behavioral health treatment services, possibly due to health service disruptions caused by the pandemic (Mellis, Potenza, and Hulsey 2021), but a majority of a local health departments reported in a national survey that they reduced the level of substance-use related services offered during the pandemic (Hall et al. 2022).

Another health risk is the "post-acute sequelae of COVID," or "long COVID," the name given to the occurrence of persistent symptoms among patients who have recovered from COVID-19 (as reported by Nabavi 2020). In some cases, these symptoms are debilitating and interfere with patients' ability to work and quality of life. Estimates of the prevalence of long COVID vary widely—from 7.5 percent to more than 40 percent (Ford et al. 2023). Such wide estimates speak to the need for more rigorous and standardized methodologies for studying the condition. Regardless, the evidence available suggests the total number of patients affected could reach into the millions. Rigorous studies and better clinical case definitions are needed to enumerate more precisely the effects of long COVID and to identify treatments (Munblit et al. 2022). More research is also needed to understand the prevalence of prolonged symptoms following other acute infections, which there is some early evidence may affect patients (Choutka et al. 2022)

No Country Escaped COVID, but the United States was Exceptional

In comparison with other countries, the United States reported the greatest number of COVID-19 deaths throughout the majority of the pandemic and continues to do so today. Due to differences in countries' surveillance approaches, age structures, and underlying comorbidities, understanding why the US death tolls during the pandemic were comparatively higher than other countries requires more careful analysis than simply comparing reported case and death numbers.

One misperception that arose during the pandemic was that the best-prepared countries as a group were hit hardest by the virus. This assertion was largely based on a comparison of countries' official COVID-19 surveillance data case and death numbers, which appeared to indicate that countries that scored higher in preparedness frameworks like the Global Health Security Index and the Joint External Evaluation, a monitoring and evaluation tool used by the World Health Organization, reported the highest number of COVID-19 cases and deaths (Haider et al. 2020; Aitken et al. 2020; Kim et al. 2021). These analyses and others have led some to conclude that pandemic preparedness efforts were ineffective at mitigating the health consequences from the COVID-19 pandemic (Omberg and Tabarrok 2022).

We now have evidence that the perceived inverse relationship between countries' pandemic preparedness and COVID-19 tolls was largely driven by inadequacies in global surveillance. In the early days of the pandemic, it was difficult to get an aggregate picture of which countries were being affected by COVID-19 and to what

extent. The World Health Organization initially published official country case reports in PDF files that provided limited information regarding the relative spread of the virus. Eventually, the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) COVID-19 Dashboard launched to aggregate case data, including from unofficial sources, that were reporting in real time to generate a timelier tally and map of cases (Dong, Du, and Gardner 2020; Dong et al. 2022). Other groups, like private companies and journalistic outlets, followed, creating more options for viewing timely aggregated pandemic data. As governments and organizations like the World Health Organization improved their own data-tracking options, the quality of national case data generally improved.

However, tracking cases and deaths from the COVID pandemic across countries has remained challenging. Initially, countries differed in how quickly they were able to establish testing for COVID-19. While all countries eventually gained the ability to test for the presence of the SARS-CoV-2 virus, they differed in how much testing actually happened. For example, by July 2020, it was estimated that most countries in Asia and Africa had tested less than 1 percent of their populations for COVID-19, as compared with 10–20 percent population testing coverage reported by the United States, Russia, and countries in Europe (Amukele and Barbhuiya 2020). In October 2021, the World Health Organization office for the African region estimated that only one out of every seven SARS-CoV-2 infections in the region were being detected (World Health Organization African Region 2021).

COVID-19 deaths also proved to be hard to track globally for similar reasons. An early indication that some countries were undercounting deaths came from a postmortem analysis conducted in Zambia, which estimated that, because of undertesting, as few as 10 percent of all COVID-19 deaths that occurred in the country may have been identified (Gill et al. 2022). The study also noted that testing in Zambia was rarely performed before death and almost never performed for deaths that occurred outside of healthcare facilities.

The bottom line here is that trying to track COVID-19 cases across countries may only be measuring how much or how well countries are counting. To understand the tolls of the COVID pandemic, we are better off using metrics that are less dependent on tools like the extent of COVID testing.

One measure that has turned out to be particularly useful for tracking the pandemic is "excess deaths," which can quantify the total mortality impact of the pandemic. Rather than relying on diagnostic tests to directly enumerate COVID-19 cases and deaths, we can compare deaths from all causes observed during the pandemic to what we would expect given historical patterns to derive excess mortality. Figure 3 provides an illustration of how excess mortality is computed for select countries. This measure provides a comprehensive understanding of the levels of deaths that occurred throughout the pandemic, as it captures both direct COVID-19 deaths and deaths that are an indirect effect of the pandemic, such as those resulting from health service disruptions (in this journal, see Alsan, Chandra, and Simon 2021 for a more detailed explanation of excess deaths).

All-Cause Mortality Rates per 1,000 Population for Selection Countries, 2013–2021

Panel A: Observed All-cause Mortality Rates per 1,000 Population, 2013–2021
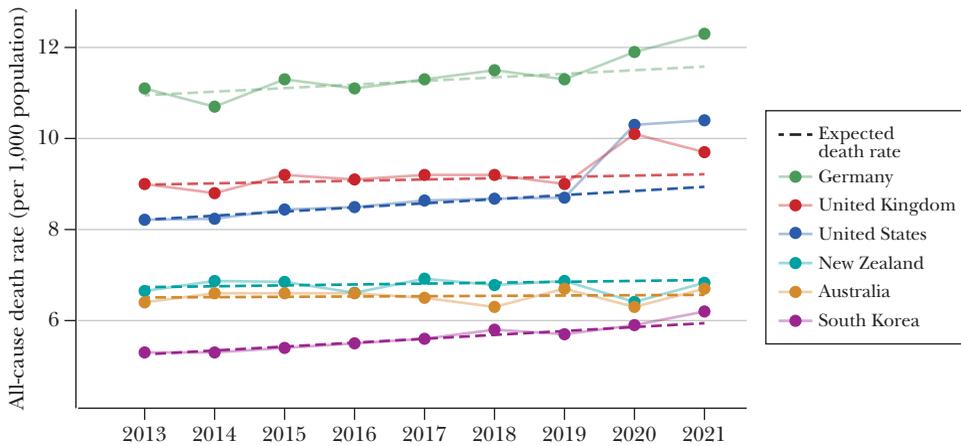
Panel B: Excess Death Rates per 1,000 Population, 2020 and 2021
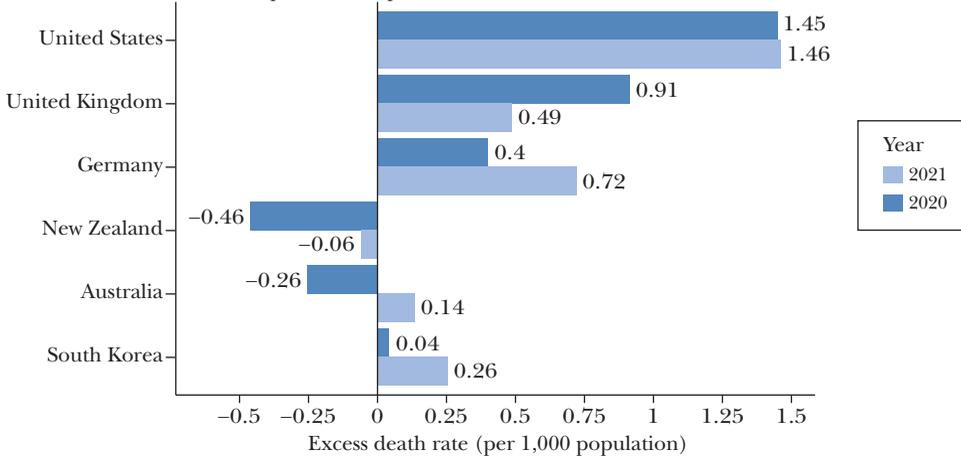
Source: All-cause death rate data are from the World Bank.
Note: Panel A demonstrates observed all-cause mortality rates per 1,000 population for six countries from 2013 to 2021. The points are observed all-cause death rates, while the dashed line is the expected death rate under normal conditions. Panel B showcases the corresponding country-specific excess death rates per 1,000 population, which were computed by subtracting the observed death rate from the expected death rate in the years 2020 and 2021. Expected death rates computed by running country-specific linear regressions where observed deaths points from 2013 to 2019 were used as input.

Measures of excess deaths suggest that many countries with lower reported COVID-19 cases may have suffered more greatly than their official statistics indicated. The relationship is even more pronounced when we account for differences in countries' demographics—a necessary step, given that age is the single biggest risk factor for dying from COVID-19.

Figure 4 provides a detailed illustration of how the relationship between the Global Health Security Index and mortality changes as we take into account the above biases in reported COVID-19 data. Panel A confirms previous studies that have demonstrated that the most prepared countries report the largest crude rates of COVID-19 deaths. However, by utilizing excess mortality as the outcome, and thus sidestepping the problems of how well COVID deaths were counted, we see a flat line (a null relationship) in Panel B.

When we further adjust for age—the single largest risk factor of COVID-19 mortality—through indirect age-standardization methods (Heuveline and Tzen 2021), we now see the expected negative relationship where the most prepared countries experienced the lowest mortality rates. The age-standardization process is a critical step before making direct country-level comparisons, because countries with a larger share of elderly individuals are expected to have higher rates of COVID-19 deaths, with the latest data confirming that over 80 percent of all global COVID-19 deaths occurred in people aged 60 years and greater (Harris 2023). Other work that has argued for no relationship between the Global Health Security Index and excess deaths does not include age-standardization, and thus lacks this crucial step for comparing health effects of the pandemic across countries.

Our own work has further confirmed the negative relationship shown in Figure 4 and the benefits of pandemic preparedness even after accounting for cross-country differences in income (Ledesma et al. 2023). That is, when we compare national Global Health Security Index scores to age-standardized excess deaths, we see across all income groups evidence that those with higher scores tended to experience lower excess mortality during the pandemic than those with lower scores. Not only do we see a difference in excess mortality between higher- and lower-prepared countries, but we also see an inverse linear relationship between age-standardized excess mortality and preparedness.

Though these findings contradict the earlier analyses that showed the better prepared countries tended to report the most COVID-19 deaths, an inverse relationship between preparedness measures and pandemic-related mortality should not be surprising. Preparedness frameworks like the Global Health Security Index count the capacities that countries use to enumerate and report infections and deaths. It is not unreasonable to think that those countries with more capacities can more easily find and report deaths.

But when we account for biases due to differences in countries' age structures and surveillance capacities, the United States stands out as a clear outlier, as shown in Figure 5. Although the United States was among the highest prepared countries as measured by the Global Health Security Index, it experienced the third-highest age-adjusted mortality. The observed excess mortality in the United States is almost more than eight times higher than what we would have expect given the observed mortality in the other highest prepared countries.

To understand how well countries will do in a pandemic, it is important to understand not only what resources they have to respond to a pandemic, but also

Relationships between the Global Health Security (GHS) Index and COVID-19 Mortality Rate Outcomes per 100,000 Population, 2020–2021

Panel A: The GHS Index and Cumulative Reported COVID-19 Death Rates per 100,000 Population, by Country

Panel B: The GHS Index and COVID-19 Excess Death Rate per 100,000 Population, by Country

Panel C: The GHS Index and Age-Standardized Excess Death Rate per 100,000 Population, by Country
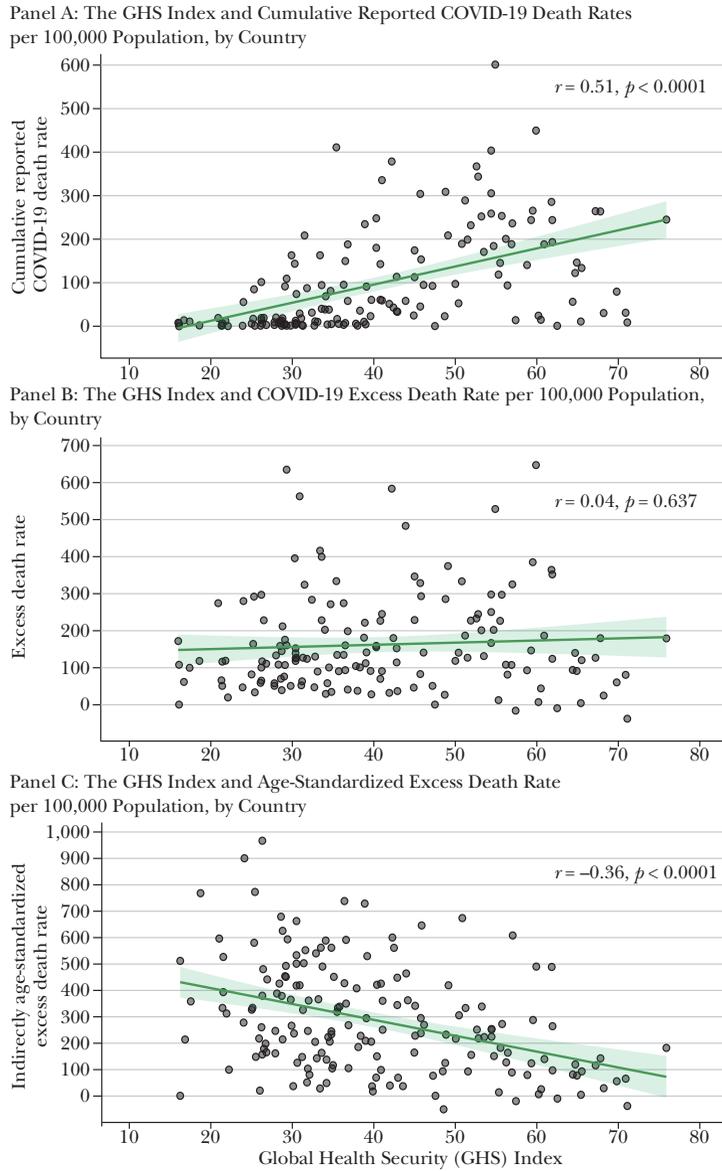
Source: Panel A shows the relationship between the Global Health Security (GHS) Index and cumulative reported COVID-19 death rates per 100,000 population. COVID-19 death data from the John Hopkins COVID-19 data repository, while population counts are from the United Nations (UN) Population Division in 2021. Panel B demonstrates the relationship between the GHS Index and COVID-19 excess mortality per 100,000 population. The excess death data are from the Institute for Health Metrics (IHME) COVID-19 database. Panel C visualizes the relationship between the Global Health Security (GHS) Index and indirectly age-standardized excess death rate per 100,000 population. Excess death rates are from IHME COVID-19 database.

Note: The black points represent countries, while the green line represents linear regression lines with the shaded area visualizing the corresponding 95 percent confidence intervals. Pearson *r* correlations with *p*-values shown in each panel. Some countries removed from panels for clarity. All COVID-19 mortality data are cumulative from January 1, 2020, to December 31, 2021.

Figure 5

Relationship between the Global Health Security (GHS) Index and Indirectly Age-Standardized Excess Death Rate per 100,000 Population for the Most Prepared Countries, 2020–2021
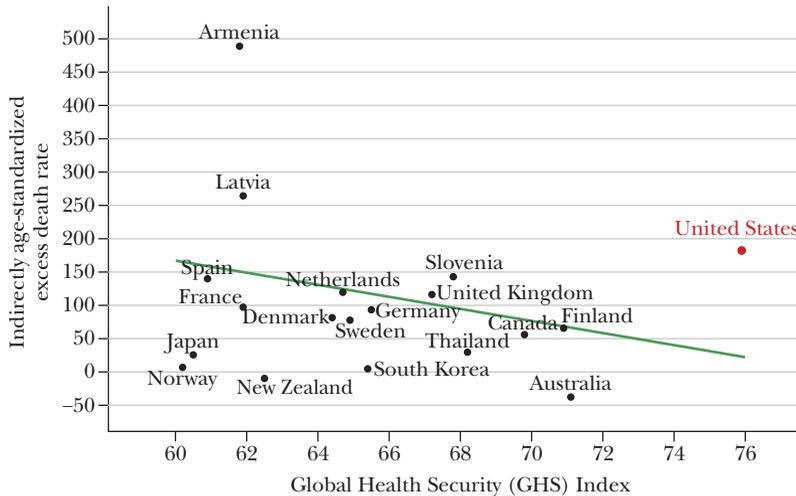
Source: Excess death rate data are from the IHME COVID-19 database.
Note: This figure is analogous to panel C of Figure 4 but limited to countries with Global Health Security (GHS) Index scores of at least 60. The dark green line is the linear regression line from Figure 4 panel C. Considering the association between the GHS Index and age-standardized excess death rate, the age-standardized excess death rate for the United States was expected to be 23 per 100,000 population but the observed rate was 182 per 100,000 population.

how well they use them—or even *whether* they use them. In the case of the United States, specific gaps in its pandemic preparedness may have proven to be especially harmful during COVID, and there was also a failure to use is resources appropriately in responding to the pandemic.

Gaps and Challenges for US Preparedness

Though the United States outranked 194 other countries in the 2019 Global Health Security Index that sought to measure pandemic preparedness, the same report also noted that every country, including the United States, lacked some necessary capacities or possessed risks that could hinder its response to a pandemic. In some cases, existing risks or missing capacities in one area could hinder use of other capacities. For example, a country that has advanced laboratory capacities may not be able to fully realize the benefits of this resource if it lacks a plan to communicate who is at risk of infection and would benefit from being tested.

In 2019, the United States lost points in the Global Health Security Index in a number of areas that may have been especially relevant to its response to COVID-19. One deficiency is that while the US healthcare system has very high spending substantially due to expensive technologies, it does not rank very highly on some basic quantities. Among 60 high-income countries in the GHS Index, the US ranked 38th for its number of physicians per capita and 40th for its number of hospital beds per capita, according to World Bank and World Health organization data. The United States also ranked 175th in the globe for access to healthcare, due to its lack of universal health coverage and high out-of-pocket healthcare costs. Throughout the pandemic, US health facilities struggled to meet demands for care. This pattern likely exacerbated both the direct harms of COVID-19 and the indirect healthcare impacts from the pandemic.

Out-of-pocket costs, and fears of such costs, did affect the pandemic response. Though legislation passed in 2020 sought to remove cost barriers for SARS-CoV-2 testing, testing costs remained and have been cited as a barrier to testing performed in the United States (Kurani et al. 2021). Similarly, while COVID-19 vaccines were provided by the US government free of charge, misperceptions regarding the vaccines' cost were often cited as a reason people delayed getting vaccinated (Hamel et al. 2021).

In addition, some of the United States' existing preparedness capacities may not have been as functional as previously thought. Although the United States had established a national stockpile of medicines, personal protective equipment, and ventilators, when signs of a new outbreak surfaced, calls by federal officials to replenish and augment these supplies were ignored (as reported in Taddonio 2020).

For a federalist country like the United States, it is also important to look beyond the US government's readiness for a pandemic. Whatever efforts the US government makes to amass and exercise plans, resources, tools, and policies, the effects may ultimately be determined by efforts within individual communities. Across the United States, there are more than 3,000 local health departments in 53 states and territories that have a role in responding to crises. In many areas, these local entities have the primary response authority. However, following the recession of 2008–2009, many states reduced public health spending, which translated to cuts of more than 55,000 jobs in local health departments (National Association of County and City Health Officials 2018). Staff shortages left many local health departments unable to adequately conduct case investigations and contact tracing during the early COVID-19 response. To address these shortcomings, part of the American Rescue Plan Act of 2021 was focused on funding COVID-19 testing, contact tracing, and mitigation (Nuzzo and Gostin 2022).

Some factors intrinsic to the United States, and not measured by the Global Health Security Index, likely made it more vulnerable than other countries to the specific epidemiology of the SARS-CoV-2 virus. For example, compared to other countries, the United States has a high percentage of its population living in congregate settings that enable the virus to spread among populations that are particularly susceptible to severe illness from it.

As one prominent example, nursing homes are a congregate setting that may have contributed to high levels of mortality in the United States. Compared to some other high-income countries, the United States had a higher percentage of its 65-and-older population living in nursing homes (Ribbe et al. 1997). For much of the first year of the pandemic, nursing homes accounted for a large share of the total COVID-19 deaths reported in the United States (Shen 2022). Shen (2022) finds that staff neighborhood characteristics were a large and significant predictor of COVID-19 nursing home deaths. Specifically, the author found that nursing homes whose staff came from denser, less white neighborhoods with more public transportation use had significantly larger outbreaks of COVID-19. Another study published in August 2021 calculated that the COVID-19 death rate for seniors living in nursing homes was 23 times that of those who lived outside of these facilities, in part because of lower quality of care and lapses in infection-prevention protocols in nursing homes (Cronin and Evans 2022).

As another example, the United States has one of the highest proportions of its population in prison or jail. Several studies have identified an association between prisons and higher community burden of COVID-19 and COVID-19 mortality (LeMasters et al. 2022; Lofgren et al. 2022; Saloner et al. 2020).

The Potential Disconnect between Preparedness and Response

Preparing for pandemics is a subset of efforts within the larger field of health emergency preparedness, which the World Health Organization (2017) has defined as "the knowledge and capacities and organizational systems developed by governments, response and recovery organizations, communities and individuals to effectively anticipate, respond to, and recover from the impacts of likely, imminent, emerging, or current emergencies." This resource-focused definition of preparedness highlights the difficulty in making direct linkages between countries' preparedness for health emergencies and the actual experience of countries during a pandemic. Preparedness, encompassing the tools that are amassed and the actions taken prior to an emergency, represents only the first half of the equation. The second part entails response—what countries actually do during a health emergency.

In this way, preparedness capacities can be thought of like fire prevention and safety tools. Smoke detectors and fire alarms are important for providing early warning when a fire occurs. Testing whether such capacities will function as planned, such as by checking that fire alarms work and exercising building evacuation plans, also are key to preparing for fires. Policies that reduce risks from fires, such as building codes, help illuminate and reduce the baseline risk for fires. But while these fire safety tools can aid in reducing the frequency and harms from fires, their mere existence is never entirely predictive of how the actual response that will happen when a fire occurs. As with pandemic preparedness, how governments, organizations, and individuals will use their capacities and overcome risk factors will ultimately determine the number of lives that will be saved or lost.

Here, we will sidestep some of the controversial questions about the response to the COVID-19 pandemic, like the effectiveness of mask mandates and school closures. These topics offer important research questions. But at the national level, it is difficult to attribute differences in national mortality to specific policy approaches. Most of the studies rely on observational data, but differential biases in how nations mixed these policies together and how they generated their data relevant to COVID-19 limit what these analyses can tell us. Within the United States, response to COVID-19 was largely implemented at the state and local level and was variable throughout the pandemic. The simultaneous application of multiple policies and other response measures in communities, along with the presence of unmeasured differences between those communities that chose to implement certain measures and those that did not, makes drawing causal inferences difficult. Finally, it would be unwise to extrapolate directly from whether these steps were useful during the COVID pandemic to whether they might be useful in a future pandemic with different epidemiological characteristics.

Despite the limitations of existing evidence for drawing causal inferences, certain factors associated with the US response warrant further investigation as possible causes for its poor outcomes relative to other countries.

First, the United States did not fully make use of its preparedness capacities in its response to the COVID-19 pandemic. A key initial lapse was its failure to quickly make use of its massive network of high-quality laboratories to diagnose and characterize infections. Although the United States has a world-class network of public health and clinical laboratories, which had the capacity to develop their own assays to test for SARS-CoV-2, federal restrictions initially prevented these laboratories from doing so. News outlets have documented strategic and operational missteps the United States made in rolling out test kits to US laboratories, shoring up testing supplies, and harnessing domestic research and development capacity to create rapid tests that could be used outside of the laboratory (Shear et al. 2020). This severely constrained the number of tests the United States could conduct (and likely allowed the virus to spread around the country undetected) until these restrictions were eventually lifted.

Being slow to establish testing relative to other high-preparedness countries, like South Korea, allowed the virus to spread silently in communities, making it difficult to slow or contain the spread (Nuzzo and Gostin 2022). Decreasing the time to identify and isolate infected individuals significantly reduces transmission of the virus (Wang et al. 2022). However, even after testing was established throughout the United States, delays in getting those test results to patients persisted throughout much of the pandemic. Even now, the lack of a national testing strategy and unaddressed shortages in testing supplies continue to limit the country's ability to suppress SARS-CoV-2 (Neelon et al. 2021; Hsiehchen, Espinoza, and Slovic 2020; Shvetsova et al. 2022; Grossman et al. 2020) .

Second, politics seemed to play an important role in the pattern of COVID-19 mortality. Though public support for the health agencies' handling of the pandemic was initially high among members of both political parties, overall support eroded over time, with greater declines seen among Republicans (Kirzinger

et al. 2023). This partisan erosion of support for health agencies' response tracked with partisan lack of compliance with public health recommendations aimed at slowing the spread of the virus.

An association developed between politics and poor health outcomes. One study found that by the end of July 2022, states with a greater fraction of their populations that voted for the 2020 Republican presidential candidate had reported higher COVID-19 cases and deaths (Bollyky et al. 2023)—but not all states struggled equally. This finding builds on evidence from studies that showed partisan differences in testing rates, COVID-19 deaths, and adherence to public health disease control recommendations (Neelon et al. 2021; Hsiehchen, Espinoza, and Slovic 2020; Shvetsova et al. 2022; Grossman et al. 2020). Political affiliation was also identified as a key factor associated with differences in vaccine uptake. About four months after COVID-19 vaccines first became available, the uptake among US counties that voted for Donald Trump and those that voted for Joseph Biden began to diverge. By January 2022, COVID-19 vaccination coverage in counties that voted for Trump was an average of 13 percent lower than in counties that voted for Biden (Kates, Tolbert, and Rouw 2022).

Understanding the interactions of which came first—loss of public support for the COVID-19 response or loss of some political leaders' interest in responding to COVID-19—is important to know. If it is the former, developing and socializing better pandemic response playbooks could be necessary. If it is the latter, better equipping politicians and public health officials to lead during health emergencies may be helpful.

The challenge of minimizing politicization may be especially difficult in the United States. In the Global Health Security Index of pandemic preparedness, the United States was one of only a small number of high-income countries that received the lowest possible score for its level of public confidence in government. Poor confidence in the government can undermine the public's adherence to officially sanctioned or directed disease-control measures. Ultimately, public confidence in government was one measure in the GHS Index that was found to be significantly associated with countries' observed excess mortality (Ledesma et al. 2023).

Mixed communication messages from public health leaders during the pandemic may have also led to politicization and erosion of support. At the start of the pandemic, US officials committed several communication errors. Some pertained to changing scientific evidence. For example, health officials initially discouraged the use of masks, but later mandated them when evidence emerged that the SARS-CoV-2 virus that causes COVID-19 could be transmitted by people without symptoms. Changing science, however, was not the only source of communication error. There were larger challenges in setting expectations for how the pandemic would unfold. Initially health officials discussed the need to shut down public gatherings, schools, and businesses and otherwise implement social distance policies in limited terms—for example, the Trump administration's COVID guidelines spoke of staying home for 15 days to slow the spread (Feuer and Higgins-Dunn 2021). But while knowledge of the virus has changed over the course of the pandemic, there was almost no evidence at the time that

suggested the pandemic was capable of ending within two weeks. Failing to set up the American public adequately for what was in store likely is part of why support for health officials' advice, which was initially high at the pandemic's start, eroded over time (Kirzinger et al. 2023).

The director of the Centers for Disease Control (CDC) recently acknowledged that the agency failed in its response to COVID-19, owing to delayed case reporting and inconsistent communication of social distancing, masking, and vaccination (Tanne 2022). While politicization of the pandemic likely harmed compliance with COVID-19 mitigation strategies, the inconsistent messaging from public health leaders may have also played a role. Improved communication messaging is therefore critical for future public health emergencies. Data from qualitative studies indicate that having messaging that is tailored, culturally congruent, and delivered by trusted messengers while countering misinformation in real-time is most effective (Overton et al. 2021). The need for clear, science-based communication was affirmed by a nationally representative poll of the American public, which found that higher trust in the CDC was tied to a belief that the agency provided evidence-based and protective resources, rather than a belief that the agency did a good job at controlling the outbreak (SteelFisher et al. 2023). Though the CDC's new emphasis on communication, timeliness, collaboration, and accountability will help overcome the limitations of its previous inconsistent messaging and improve health communication strategies, more direct investigations are needed on how the CDC's messaging may have impacted compliance with response strategies.

A third factor that warrants investigation is that the United States struggled to aggregate, analyze, and publish real-time COVID-19 data to help state and local governments, businesses, schools, community groups, and individuals make decisions about how best to protect themselves. Initially, the most complete and up-to-date data were assembled and published by nongovernmental organizations. When Lauren Gardner and Ensheng "Frank" Dong and at Johns Hopkins created the "Hopkins COVID map," it was one of the first sites to publish a map and near-real-time tally of global COVID-19 cases (Dong, Du, and Gardner 2020). Soon after, a volunteer team organized by *The Atlantic* magazine set up a pivotal COVID-19 testing tracking site. The fact that this nongovernmental source of data became the go-to resource for governments, private-sector and other nongovernmental organizations, media, and members of the public is somewhat of an indictment of US governmental surveillance efforts (Donovan 2023; Benadjaoud and Egan 2023).

While nongovernmental efforts could aggregate and display data that state and local government response agencies made public, these data were often of variable quality and not standardized between states. As a result, these data-trackers were only a partial substitute for what should have been a government-led effort to generate high-quality, standardized surveillance data. In the United Kingdom, well-designed surveys were launched to understand the incidence of COVID-19 and demographics of the patients who were being infected—which could then help

government to address difference in who has access to or inclination to get tested or vaccinated. Though many experts called on the US government to develop such surveillance surveys to supplement its incomplete and biased passive surveillance efforts, it did not do so (Dean 2022).

The lack of adequate data likely undermined the ability of the United States to mount responses to the pandemic that were tailored to local circumstances. Efforts to decide which measures were best suited to control the spread of the virus were undermined by lack of data. For example, while debates were raging about whether school closures were affecting disease transmission rates, the US government did not have a national database on school closures (Goldstein 2021). Faced with rapidly rising cases, state and local health departments implemented broad measures, such as state- or county-wide mask mandates or rules on social distancing without knowing whether transmission was being driven by exposures in the environments likely to be affected by the mandates (Nuzzo and Blauer 2021).

Fourth, the United States was late to increase genetic testing needed to track variants of COVID, which likely hindered its response to deadlier and more transmissible variants of the virus (Asgari 2021). When the "delta variant" became of concern in the United States in mid-2021, the United States had only sequenced about 4 percent of cases it had detected by that date (Schuster-Bruce 2021)—far behind the sequencing efforts of the United Kingdom and other high-income countries. Not knowing how widespread the delta variant was made it difficult to mitigate the increase in hospitalizations and deaths that the virus caused.

Finally, the United States did not respond to the pandemic in ways that would counterbalance underlying social vulnerabilities that were likely to increase its COVID-19 tolls. Throughout the pandemic, disparities in stark racial/ethnic composition of reported COVID-19 cases, hospitalizations, and deaths pointed to a need to increase efforts to protect those at greatest risk. For example, an analysis of state-reported testing data showed that while Hispanic/Latino people were more likely to be hospitalized and die from the virus than non-Hispanic/Latino people, there were fewer tests performed for every case identified among Hispanic/Latino people as compared with other racial/ethnic groups (Pond et al. 2022). Insufficient efforts to address income inequality may have been another challenge to the US response. There is evidence that US counties with greater income inequality experienced higher levels of cases and deaths. Furthermore, people with lower incomes were more likely to report life circumstances that impede their ability to reduce their risk of infection, including inability to telework (Papageorge et al. 2021).

The Need for a Consensus on Lessons for Future Health Emergencies

Analyses of why the United States experienced such high levels of COVID-19 mortality relative to its level of preparedness point to several possible reasons. First, while preparedness may be necessary to mount an effective response to a pandemic, the United States may have underutilized its existing pandemic-related

tools, capacities, and policies. Second, while all countries possess intrinsic factors that may make them more vulnerable to pandemics, the United States may have failed to account for these vulnerabilities in how it applied its capacities to respond to COVID-19. Third, the US response to the pandemic was challenged by politics, inconsistent messaging, inadequate data, and inequality.

However, these conclusions are only suggestive. The academic literature and public media accounting of the US experience do not tell us about the decision-making, resource constraints, and operational tradeoffs that government leaders experienced. Without a full audit of the inner workings of the US governmental response, we are left with holes in understanding how one of the best-prepared countries in the world suffered worse from the pandemic than its peers.

Ideally, the US government would pass legislation to create a serious bipartisan audit of its COVID experience—one that opens the books on efforts of the Biden, Trump, and earlier administrations to prepare for and respond to pandemic threats. Much like the passage of a law requiring an inquiry into US missteps leading up to 9/11 (Commission on Terrorist Attacks upon the United States 2004), a thorough record and investigation of governmental efforts to prepare for pandemic threats are needed to understand what the United States did and did not do, and why the country failed to make better use of its prepandemic advantages.

The COVID pandemic has been referred to as a once-a-century crisis, referring back to the Great Influenza Pandemic of 1918. But while these two society-changing events bear some similarities, it is not accurate to assume that pandemics will only come every 100 years. Since 1918, a steady stream of infectious disease emergencies has challenged the United States in different ways. Standout examples include: three influenza pandemics (1957, 1968, and 2009), the HIV pandemic, the reemergence and global spread of Zika virus, and the recent MPOX (formerly monkeypox) outbreak. The steady cadence of these emergencies tracks with data that strongly indicate that the frequency with which new pathogens arise and cause outbreaks has steadily increased. Even accounting for improvements in surveillance, the frequency of emerging infectious disease outbreaks tripled between 1980 and 2010 (Smith et al. 2014). Prudence suggests expecting a rising number of infectious health emergencies in the future and preparing accordingly.

References

Aitken, Tess, Ken Lee Chin, Danny Liew, and Richard Ofori-Asenso. 2020. "Rethinking Pandemic Preparation: Global Health Security Index (GHSI) Is Predictive of COVID-19 Burden, but in the Opposite Direction." *Journal of Infection* 81 (2): 318–56.

Alsan, Marcella, Amitabh Chandra, and Kosali Simon. 2021. "The Great Unequalizer: Initial Health Effects of COVID-19 in the United States." *Journal of Economic Perspectives* 35 (3): 25–46.

Amukele, Timothy, and Mustafa Barbhuiya. 2020. "African Countries Need Cheaper COVID-19 Tests: Here's How to Get Them." *The Conversation*, July 12. http://theconversation.com/african-countries-need-cheaper-covid-19-tests-heres-how-to-get-them-141315.

Arias, Elizabeth, Betzaida Tejada-Vera, Kenneth D. Kochanek, and Farida B. Ahmad. 2022. "Provisional Life Expectancy Estimates for 2021." NVSS Vital Statistics Rapid Release 23.

Asgari, Nikou. 2021. "US Struggles to Track Delta Variant Due to Fragmented Health System." *Financial Times*, July 13. https://www.ft.com/content/f8e83ce5-037f-469d-bee9-aaa37f6e2554.

Benadjaoud, Youri, and Emma Egan. 2023. "What the End of COVID Trackers Means for the Future of the Pandemic." ABC News. February 23, 2023. https://abcnews.go.com/Health/end-covid-trackers-means-future-pandemic/story?id=97399856.

Bollyky, Thomas J., Emma Castro, Aleksandr Y. Aravkin, Kayleigh Bhangdia, Jeremy Dalos, Erin N. Hulland, Samantha Kiernan et al. 2023. "Assessing COVID-19 Pandemic Policies and Behaviours and Their Economic and Educational Trade-Offs across US States from Jan 1, 2020, to July 31, 2022: An Observational Analysis." *The Lancet* 401 (10385): 1341–60.

Cameron, Elizabeth E., Jennifer B. Nuzzo, and Jessica A. Bell. 2019. *Global Health Security Index: Building Collective Action and Accountability*. Washington, DC: Nuclear Threat Initiative. https://www.ghsindex.org/wp-content/uploads/2019/10/2019-Global-Health-Security-Index.pdf (accessed on September 20, 2023).

Centers for Disease Control and Prevention. 2022. "Burden of Flu." https://www.cdc.gov/flu/about/burden/index.html (accessed October 4, 2022).

Choutka, Jan, Viraj Jansari, Mady Hornig, and Akiko Iwasaki. 2022. "Unexplained Post-acute Infection Syndromes." *Nature Medicine* 28 (5): 911–23.

Commission on Terrorist Attacks upon the United States. 2004. *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States*. Washington, DC: US Government Publishing Office.

Cronin, Christopher J., and William N. Evans. 2022. "Nursing Home Quality, COVID-19 Deaths, and Excess Mortality." *Journal of Health Economics* 82: 102592.

Dean, Natalie. 2022. "Tracking COVID-19 Infections: Time for Change." *Nature* 602 (7896): 185.

Dong, Ensheng, Hongru Du, and Lauren Gardner. 2020. "An Interactive Web-Based Dashboard to Track COVID-19 in Real Time." *The Lancet Infectious Diseases* 20 (5): 533–34.

Dong, Ensheng, Jeremy Ratcliff, Tamara D. Goyea, Aaron Katz, Ryan Lau, Timothy K. Ng, Beatrice Garcia et al. 2022. "The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: Data Collection Process, Challenges Faced, and Lessons Learned." *The Lancet Infectious Diseases* 22 (12): e370–76.

Donovan, Doug. 2023. "Johns Hopkins Winds down Pioneering Pandemic Data Tracking." The Hub. February 10, 2023. https://hub.jhu.edu/2023/02/10/coronavirus-resource-center-ending-tracking/.

Feuer, Will, and Noah Higgins-Dunn. 2021. "A Year Later, Trump's '15 Days to Slow the Spread' Campaign Shows How Little We Knew about Covid." *CNBC*, March 16. https://www.cnbc.com/2021/03/16/covid-a-year-later-trumps-15-days-to-slow-the-spread-pledge-shows-how-little-we-knew.html.

Ford, Nicole D., Douglas Slaughter, Deja Edwards, Alexandra Dalton, Cria Perrine, Anjel Vahratian, and Sharon Saydah. 2023. "Long COVID and Significant Activity Limitation among Adults, by Age—United States, June 1–13, 2022, to June 7–19, 2023." *Morbidity and Mortality Weekly Report* 72 (32): 866–70.

Gill, Christopher J., Lawrence Mwananyanda, William B. MacLeod, Geoffrey Kwenda, Rachel C. Pieciak, Laurent Etter, Daniel Bridges et al. 2022. "What Is the Prevalence of COVID-19 Detection by PCR among Deceased Individuals in Lusaka, Zambia? A Postmortem Surveillance Study." *BMJ Open* 12 (12): e066763.

Global Preparedness Monitoring Board. 2019. *A World at Risk: Annual Report on Global Preparedness for Health Emergencies*. Geneva: World Health Organization.

Goldstein, Dana. 2021. "Emily Oster, the Brown Economist, Is Launching a New Data Hub on Schools and the Pandemic." *New York Times*, September 15. https://www.nytimes.com/2021/09/15/education/emily-oster-covid-data-schools.html.

Grossman, Guy, Soojong Kim, Jonah M. Rexer, and Harsha Thirumurthy. 2020. "Political Partisanship Influences Behavioral Responses to Governors' Recommendations for COVID-19 Prevention in the United States." *Proceedings of the National Academy of Sciences of the United States of America* 117 (39): 24144–53.

Haider, Najmul, Alexei Yavlinsky, Yu-Mei Chang, Mohammad Nayeem Hasan, Camilla Benfield, Abdinasir Yusuf Osman, Md. Jamal Uddin et al. 2020. "The Global Health Security Index and Joint External Evaluation Score for Health Preparedness Are Not Correlated with Countries' COVID-19 Detection Response Time and Mortality Outcome." *Epidemiology and Infection* 148: e210.

Hall, Kellie, Francis Higgins, Karla Feeser Beach, Kabaye Diriba, Mandy Sladky, and Timothy C. McCall. 2022. "Disruptions to U.S. Local Public Health's Role in Population-Based Substance Use Prevention and Response during COVID-19." *Substance Abuse Treatment, Prevention, and Policy* 17: 73.

Hamel, Liz, Lunna Lopes, Grace Sparks, Mellisha Stokes, and Mollyann Brodie. 2021. "KFF COVID-19 Vaccine Monitor: April 2021." *KFF*, May 6. https://www.kff.org/coronavirus-covid-19/poll-finding/kff-covid-19-vaccine-monitor-april-2021/.

Harris, Emily. 2023. "Most COVID-19 Deaths Worldwide Were among Older People." *JAMA* 329 (9): 704.

Heuveline, Patrick, and Michael Tzen. 2021. "Beyond Deaths per Capita: Comparative COVID-19 Mortality Indicators." *BMJ Open* 11 (3): e042934.

Hsiehchen, David, M. Espinoza, and P. Slovic. 2020. "Political Partisanship and Mobility Restriction during the COVID-19 Pandemic." *Public Health* 187: 111–14.

Kates, Jennifer, Jennifer Tolbert, and Anna Rouw. 2022. "The Red/Blue Divide in COVID-19 Vaccination Rates Continues: An Update." *KFF*, January 19. https://www.kff.org/policy-watch/the-red-blue-divide-in-covid-19-vaccination-rates-continues-an-update/.

Kim, Jeehyun, Kwan Hong, Sujin Yum, Raquel Elizabeth Gómez Gómez, Jieun Jang, Sun Hee Park, Young June Choe et al. 2021. "Factors Associated with the Difference between the Incidence and Case-Fatality Ratio of Coronavirus Disease 2019 by Country." *Scientific Reports* 11: 18938. https://doi.org/10.1038/s41598-021-98378-x.

Kirzinger, Ashley, Marley Presiado, Isabelle Valdes, Liz Hamel, and Mollyann Brodie. 2023. "The COVID-19 Pandemic: Insights from Three Years of KFF Polling." *KFF*, March 7. https://www.kff.org/coronavirus-covid-19/poll-finding/the-covid-19-pandemic-insights-from-three-years-of-kff-polling/.

Kurani, Nisha, Karen Pollitz, Dustin Cotliar, Giorlando Ramirez, and Cynthia Cox. 2021. "COVID-19 Test Prices and Payment Policy." *Peterson-KFF Health System Tracker*, April 28. https://www.healthsystemtracker.org/brief/covid-19-test-prices-and-payment-policy/.

Ledesma, Jorge Ricardo, Christopher R. Isaac, Scott F. Dowell, David L. Blazes, Gabrielle V. Essix, Katherine Budeski, Jessica Bell, and Jennifer B. Nuzzo. 2023. "Evaluation of the Global Health Security Index as a Predictor of COVID-19 Excess Mortality Standardised for Under-Reporting and Age Structure." *BMJ Global Health* 8 (7): e012203.

LeMasters, Katherine, Shabbar Ranapurwala, Morgan Maner, Kathryn M. Nowotny, Meghan Peterson, and Lauren Brinkley-Rubinstein. 2022. "COVID-19 Community Spread and Consequences for Prison Case Rates." Edited by Steph Scott. *PLOS ONE* 17 (4): e0266772. https://doi.org/10.1371/journal.pone.0266772.

Lofgren, Eric T., Kristian Lum, Aaron Horowitz, Brooke Mabubuonwu, Kellen Meyers, and Nina H. Fefferman. 2022. "Carceral Amplification of COVID-19: Impacts for Community, Corrections Officer, and Incarcerated Population Risks." *Epidemiology* 33 (4): 480–92. https://doi.org/10.1097/EDE.0000000000001476.

Mellis, Alexandra M., Marc N. Potenza, and Jessica N. Hulsey. 2021. "COVID-19-Related Treatment Service Disruptions among People with Single- and Polysubstance Use Concerns." *Journal of Substance Use and Addiction Treatment* 121: 108180.

Munblit, Daniel, Margaret E. O'Hara, Athena Akrami, Elisa Perego, Piero Olliaro, and Dale M. Needham. 2022. "Long COVID: Aiming for a Consensus." *The Lancet Respiratory Medicine* 10 (7): 632–34.

Nabavi, Nikki. 2020. "Long Covid: How to Define It and How to Manage It." *BMJ* 370: m3489.

National Association of County and City Health Officials. 2018. *2018 Forces of Change Survey*. Washington, DC: National Association of County and City Health Officials.

National Institute on Drug Abuse. 2023. "Drug Overdose Death Rates." National Institute on Drug Abuse. https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates (accessed February 9, 2023).

Neelon, Brian, Fedelis Mutiso, Noel T. Mueller, John L. Pearce, and Sara E. Benjamin-Neelon. 2021. "Associations between Governor Political Affiliation and COVID-19 Cases, Deaths, and Testing in the U.S." *American Journal of Preventive Medicine* 61 (1): 115–19.

Nuzzo, Jennifer B., Jessica A. Bell, and Elizabeth E. Cameron. 2020. "Suboptimal US Response to COVID-19 Despite Robust Capabilities and Resources." *JAMA* 324 (14): 1391–92.

Nuzzo, Jennifer B., and Beth Blauer. 2021. "Opinion: If We Must Wear Masks Again, We Need a Smart Approach." *New York Times*, July 29. https://www.nytimes.com/2021/07/29/opinion/mask-mandate-cdc-covid.html.

Nuzzo, Jennifer B., and Lawrence O. Gostin. 2022. "The First 2 Years of COVID-19: Lessons to Improve Preparedness for the Next Pandemic." *JAMA* 327 (3): 217–18.

Nuzzo, Jennifer B., and Jorge R. Ledesma. 2023. "Replication data for: Why Did the Best Prepared Country in the World Fare So Poorly during COVID." American Economic Association [publisher], Interuniversity Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193684V1.

Omberg, Robert Tucker, and Alex Tabarrok. 2022. "Is It Possible to Prepare for a Pandemic?" *Oxford Review of Economic Policy* 38 (4): 851–75.

Overton, D., S. A. Ramkeesoon, K. Kirkpatrick, A. Byron, and E. S. Pak. 2021. "Lessons from COVID-19 on Executing Communications and Engagement at the Community Level during a Health Crisis." Washington, DC: National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/26340.

Panchal, Nirmita, Heather Saunders, Robin Rudowitz, and Cynthia Cox. 2023. "The Implications of COVID-19 for Mental Health and Substance Use." *KFF*, March 20. https://www.kff.org/coronavirus-covid-19/issue-brief/the-implications-of-covid-19-for-mental-health-and-substance-use/.

Papageorge, Nicholas W., Matthew V. Zahn, Michèle Belot, Eline van den Broek-Altenburg, Syngjoo Choi, Julian C. Jamison, and Egon Tripodi. 2021. "Socio-demographic Factors Associated with Self-Protecting Behavior during the Covid-19 Pandemic." *Journal of Population Economics* 34 (2): 691–738.

Pond, Emily N., Lainie Rutkow, Beth Blauer, Angel Aliseda Alonso, Sara Bertran de Lis, and Jennifer B. Nuzzo. 2022. "Disparities in SARS-CoV-2 Testing for Hispanic/Latino Populations: An Analysis of State-Published Demographic Data." *Journal of Public Health Management and Practice* 28 (4): 330–33.

Ribbe, Miel W., Gunnar Ljunggren, Knight Steel, Eva Topinkova, Catherine Hawes, Naoki Ikegami, Jean-Claude Henrard, and Palmi V. Jónsson. 1997. "Nursing Homes in 10 Nations: A Comparison between Countries and Settings." *Age and Ageing* 26 (S2): 3–12.

Saloner, Brendan, Kalind Parish, Julie A. Ward, Grace DiLaura, and Sharon Dolovich. 2020. "COVID-19 Cases and Deaths in Federal and State Prisons." *JAMA* 324 (6): 602. https://doi.org/10.1001/jama.2020.12528.

Schuster-Bruce, Catherine. 2021. "The US Doesn't Really Know How Widespread the Delta Variant Is Because Its Virus Sequencing Is Lagging Far behind Many Other Rich Nations." *Business Insider*, July 13. https://www.businessinsider.com/delta-variant-us-covid-test-sequencing-strain-cases-infections-2021-7.

Shear, Michael D., Abby Goodnough, Sheila Kaplan, Sheri Fink, Katie Thomas, and Noah Weiland. 2020. "The Lost Month: How a Failure to Test Blinded the U.S. to Covid-19." *New York Times*, March 28. https://www.nytimes.com/2020/03/28/us/testing-coronavirus-pandemic.html.

Shen, Karen. 2022. "Relationship between Nursing Home COVID-19 Outbreaks and Staff Neighborhood Characteristics." *PLoS ONE* 17 (4): e0267377.

Shvetsova, Olga, Andrei Zhirnov, Frank R. Giannelli, Michael A. Catalano, and Olivia Catalano. 2022. "Governor's Party, Policies, and COVID-19 Outcomes: Further Evidence of an Effect." *American Journal of Preventive Medicine* 62 (3): 433–37.

Smith, Katherine F., Michael Goldberg, Samantha Rosenthal, Lynn Carlson, Jane Chen, Cici Chen, and Sohini Ramachandran. 2014. "Global Rise in Human Infectious Disease Outbreaks." *Journal of the Royal Society Interface* 11 (101): 20140950.

SteelFisher, Gillian K., Mary G. Findling, Hannah L. Caporello, Keri M. Lubell, Kathleen G. Vidoloff Melville, Lindsay Lane, Alyssa A. Boyea, Thomas J. Schafer, and Eran N. Ben-Porath. 2023. "Trust in US Federal, State, and Local Public Health Agencies during COVID-19: Responses and Policy Implications." *Health Affairs* 42 (3): 328–37.

Taddonio, Patrice. 2020. "Depleted National Stockpile Contributed to COVID PPE Shortage: 'You Can't Be Prepared If You're Not Funded to Be Prepared'" *FRONTLINE*, October 6. https://www.pbs.org/wgbh/frontline/article/depleted-national-stockpile-contributed-to-covid-ppe-shortage/.

Tanne, Janice Hopkins. 2022. "US CDC Announces Major Changes after Criticism of Its Responses to Covid-19 and Monkeypox." *BMJ* 378: o2074.

Wang, Xutong, Zhanwei Du, Emily James, Spencer J. Fox, Michael Lachmann, Lauren Ancel Meyers, and Darlene Bhavnani. 2022. "The Effectiveness of COVID-19 Testing and Contact Tracing in a US City." *Proceedings of the National Academy of Sciences* 119 (34): e2200652119.

World Health Organization. 2017. *A Strategic Framework for Emergency Preparedness*. Geneva, Switzerland: World Health Organization Document Production Services.

World Health Organization African Region. 2021. "Six in Seven COVID-19 Infections Go Undetected in Africa." October 14. https://www.afro.who.int/news/six-seven-covid-19-infections-go-undetected-africa.

The Evolution of Work from Home

José María Barrero, Nicholas Bloom, and
Steven J. Davis

Looking back to 1965, full days worked from home were less than half
a percent of all paid workdays in the United States. As shown in Figure 1,
the work-from-home share rose slowly over the next few decades. In the
1970s, work from home often meant briefcases stuffed with paperwork. By the
1990s, it meant phone calls and floppy disks. In the twenty-first century, the rise of
the internet facilitated collaboration at a distance. By 2019, the work-from-home
share had reached 7 percent, and it seemed reasonable to anticipate a gradual rise
in the years ahead.

Then came the pandemic. Social-distancing mandates and fear of COVID-19
drove a sudden, massive shift to work from home. Much of that shift has endured.
According to data from the US Census Bureau's Household Pulse Survey (2022–
2023) and our own Survey of Working Arrangements and Attitudes (Barrero et al.
2020–2023; https://wfhresearch.com/), full days worked from home account for
28 percent of paid workdays in June 2023—four times the estimated share for 2019.

The pandemic catalyzed the big shift to work from home, but earlier devel-
opments made it possible. Critical tools for remote work include web-based

■ *José María Barrero is Assistant Professor of Finance at Instituto Tecnológico Autónomo de
México (ITAM), Mexico City, Mexico. Nicholas Bloom is William Eberle Professor of Economics
at Stanford University, Stanford, California. Steven J. Davis is a Senior Fellow at the Hoover
Institution, Stanford, California, and William H. Abbott Distinguished Service Professor
of International Business and Economics, University of Chicago Booth School of Business,
Chicago, Illinois. Their email addresses are jose.barrero@itam.mx, nbloom@stanford.edu,
and StevenD5@Stanford.edu.*

Figure 1
Work from Home over Time in the United States
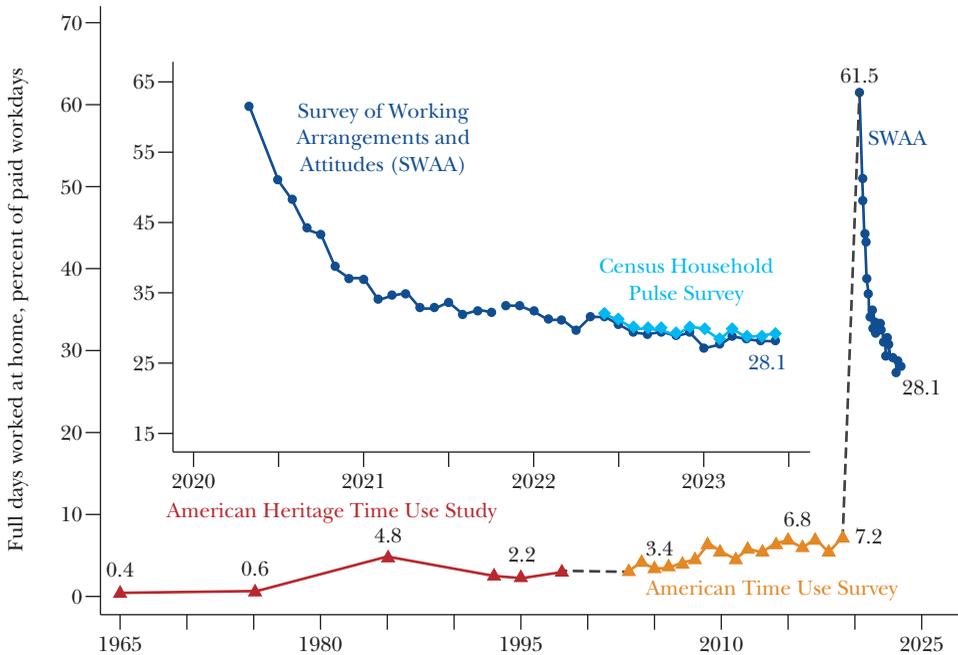
Source: The figure reports estimates for full days worked from home as a percent of all paid workdays for persons 20–64 years of age, drawing on the American Heritage Time Use Study (AHTUS) (Fisher et al. 2018) for 1965, 1975, 1985, 1993, 1995 and 1998; the American Time Use Survey (ATUS) (Flood et al. 2023) from 2003 to 2019 by year; the Survey of Working Arrangements and Attitudes (SWAA) (Barrero et al. 2020–2023) for May 2020 and from July 2020 to June 2023 by month; and the Census Household Pulse Survey (HPS) (US Census Bureau 2022–2023) from May 2022 to June 2023 by month.

Note: We use regression methods to adjust the SWAA data for question design changes. When using SWAA data, we restrict to persons who meet a prior-year earnings requirement of $20,000 in the March 2021 and prior survey waves, and $10,000 in the later waves. We weight the individual-level SWAA data to match the distribution across age-sex-education-earnings cells in the Current Population Survey. When using HPS data, we restrict to persons with household income greater than $25,000 and use the provided population weights. When using the ATUS, we restrict to persons with annual earnings of more than $20,000 in 2019 dollars calculated as 52 times weekly earnings deflated by the GDP deflator for Personal Consumption Expenditures (PCE) (US Bureau of Economic Analysis 2023). In the ATUS, we count paid workdays as ones in which the individual devotes six or more hours of work to their main job, regardless of work location, and we count work-from-home days as ones in which the individual works six or more hours on their main job at home. This approach corresponds to the SWAA-based measure. We use ATUS sample weights and drop the roughly 2 percent of observations flagged as "low quality." We follow the same approach in the AHTUS as in the ATUS, except an absence of data prevents us from imposing the earnings requirement in 1965, 1993, 1995 and 1998. We use AHTUS sample weights and drop the 2–4 percent of observations flagged as low quality.

video-conferencing platforms like Teams, Webex, and Zoom; cloud-based file-sharing services like Box, Drive, and Dropbox; and collaboration software like Asana and Slack. None of them existed in 2000. When the pandemic struck, these

tools were ready for adoption and use at scale. Some of these tools function poorly without high-speed broadband. As it turns out, the share of American adults with broadband service at home rose from zero in 2000 to more than 70 percent in 2018 (Pew Research Center 2021). Thus, the infrastructure that supports home use of remote-collaboration tools also developed greatly before the pandemic struck. These developments created a foundation for the *possibility* of a big shift to work from home.

A different issue is why the shift persisted after distancing mandates ended and after COVID death rates receded.[1] Why have working arrangements not reverted to the prepandemic status quo? To turn the question around: If the big shift was possible (and apparently has been highly valued by many since it happened), why did it not happen sooner and more gradually?

In considering an explanation, start with the obvious: The pandemic triggered a mass social experiment in working arrangements. Now consider two hypotheses. First, all of that experimentation generated a flood of new information, altering perceptions about the practicality and effectiveness of work from home. The simultaneity of experimentation across suppliers, producers, customers, and commercial networks also yielded information and experience that were hard or impossible to acquire before the pandemic. Second, employers and workers reoptimized over working arrangements in light of the new information. In particular, those with favorable experiences in the work-from-home mode opted to stick with higher work-from-home rates after the pandemic than before it struck.

To assess these two hypotheses, we surveyed tens of thousands of workers across many countries. In Barrero, Bloom, and Davis (2021b), we find that most American workers were favorably surprised by their ability to work from home productively during the pandemic. Similarly, in a sample of 27 countries, Aksoy et al. (2022) find that the average worker was favorably surprised by his or her ability to work from home during the pandemic. Both studies also find that the number of work-from-home days that *employers* plan after the pandemic ends rise strongly with *employee* assessments of work-from-home productivity surprises. In short, many workers were favorably surprised by their ability to work remotely, and those favorable surprises led to a large, lasting shift to work from home.

The pandemic also spurred other developments that helped drive a lasting shift to work from home: new investments at home and inside organizations to facilitate remote work and improve its performance, learning-by-doing in the work-from-home mode, pandemic-induced advances in products and technologies that support remote work, much greater social acceptance of work from home, and lingering concerns about infection risks that lead some people to prefer remote

---

[1] COVID-19 deaths fluctuated sharply from February 2020 through March 2022 in the United States, before settling down to comparatively low levels. Weekly deaths per million reached nearly 47 in April 2020, peaked at 70 in January 2021, again rose above 50 in February 2022, and fell sharply in March 2022. Since then, the weekly death rate is less than 10 per million. See "Weekly confirmed COVID-19 deaths per million people" for the United States at https://ourworldindata.org/covid-deaths (accessed August 4, 2023).

work. Aksoy et al. (2022) review the evidence for these developments. They also find higher levels of work from home as of mid-2021 and early 2022—and higher planned levels of work from home after the pandemic ends—in countries and regions with longer, stricter lockdown mandates during the pandemic. Thus, the full story of how the pandemic led to a large, lasting shift to work from home has multiple, reinforcing elements.

How the Big Shift Is Playing Out

Newfound Variety in Working Arrangements

Before the pandemic, most workers had few options when it came to the remote-work intensity of attractive jobs. After the pandemic, choices in this regard exploded, especially for the college-educated. Work from home is now common in many jobs once seen as unsuited to remote work. The flip side of this expanded choice set for workers is a new set of opportunities and challenges for employers. Some organizations have embraced remote work as a means to improve recruitment and retention, moderate pay growth, reduce space needs, and cut overhead costs. Others have resisted remote work, because they see it as detrimental to productivity, innovation, and skills development or antithetical to their workplace cultures.

As of 2023, 59 percent of full-time American employees (20–64 years of age) commute to their employer's worksite or to a client's location each workday. While most employees have resumed a traditional working arrangement, many have not. Twelve percent now work remotely on all or almost all workdays. Nearly 29 percent have a hybrid arrangement, whereby they split the workweek between home and their employer's worksite. Traditional working arrangements continue to prevail in front-line retail jobs, restaurants and bars, hotels, transportation, construction jobs, manufactures, many healthcare jobs, janitorial and cleaning services, onsite security services, and other jobs that require face-to-face contact with customers, clients, and coworkers or the use of specialized equipment and dedicated facilities. Even amidst social-distancing mandates in 2020, people in these types of jobs typically worked at their employer's site or not at all.

As these remarks suggest, task requirements and production technologies sharply limit the possibilities for remote work in some jobs. Often, however, a particular job involves a bundle of tasks that differ in their suitability for remote work. In some jobs, like phone-based customer support, the bundle mainly involves tasks that can be performed effectively in an office setting, a home setting, or a mix of the two. In other jobs, like college professors, the bundle includes some tasks that require presence at the employer's worksite (classroom teaching) and other tasks that do not (preparing lectures, grading assignments). More generally, the tasks associated with a given job fall along a spectrum from less- to more-suitable for remote work.

Unless the tasks for a given job are highly concentrated at the less-suitable end of the spectrum, work-from-home intensity reflects choices about job design, management practices, organizational culture, and life styles. These choices

respond to shifts in perceptions about the productivity of work from home, any stigma associated with the practice, the quality of tools that facilitate remote work, and an organization's capacity to manage remote work. They also respond to shifts in employee desires to work from home, which is perhaps another legacy of the pandemic. When the tasks for a given job cluster near the more-suitable end of the spectrum, even modest shifts in worker preferences, productivity perceptions, stigma, tools for remote collaboration, and manager skill sets can yield big changes in the extent of work from home.[2]

For employees with a hybrid arrangement, a typical pattern is to commute into the office on Tuesday, Wednesday, and Thursday and to focus those workdays on meetings, presentations, training, mentoring, lunches with colleagues and clients, and other in-person tasks. Work from home on Mondays and Fridays can then focus on tasks that require individual effort and intense focus. Hybrid arrangements are common in many professional jobs, especially in middle and senior management positions. They have spread to the healthcare sector, where counseling services, prescription renewals, and routine consultations are now sometimes provided remotely, practices that were rare before the pandemic. Fully remote employees tend to work in information technology support functions, payroll processing, phone-based customer support and sales, and routine administrative functions that require few in-person interactions. Many fully remote jobs are computer-intensive, and most involve limited team work. Typically, the labor inputs in these jobs can be easily monitored, or the individual's work product can be easily assessed.

Table 1 provides information about how working arrangements differ by employer type in the US economy. While 59 percent of full-time employees work onsite each workday, only 33 percent of contractors and gig workers do so, and only 25 percent of other self-employed persons do so. Fully remote work is four times as common for the self-employed as for employees. Given these facts, people with strong desires to work in a fully remote capacity are more likely to choose self-employment, including contract and gig work. Regulations that clamp down on contract and gig work curtail choices and, in doing so, are likely to deter some people from working altogether.

Working arrangements differ by employer size as well. In firms with fewer than 50 employees, two-thirds travel to their employer's site each workday. In firms with 500 to 4,999 employees, the corresponding share is only 51 percent. Hybrid arrangements are most common in firms with 500 to 4,999 employees, and fully remote jobs are most common in firms with 5,000 or more. Among government employees, the distribution of working arrangements is quite similar to that of all employees.

---

[2] Dingel and Neiman (2020) exploit data on the nature of work and work activities to identify occupations that *can be* performed entirely from home. They estimate that 37 percent of American workers held such jobs before the pandemic. In many other jobs, some tasks are suitable for remote work. Thus, their evidence suggests that work-from-home intensity is a choice outcome in half or more of all US jobs.

Table 1

Full-Time Working Arrangements in the United States as of 2023, Percentage Distributions

	Fully onsite (1)	Hybrid arrangement (2)	Fully remote (3)	Percent of all workers (4)
All workers	55.9	28.6	15.5	100
Self-employed, excluding contactors and gig workers	24.9	26.8	48.3	7.4
Contractors and gig workers	32.9	22.7	44.4	3.3
All employees	59.3	29	11.8	89.3
In firms with 1 to 9 employees	67.5	17.6	14.9	7.1
In firms with 10 to 49 employees	68.3	24.1	7.6	14.2
In firms with 50 to 99 employees	57.2	34.1	8.7	13.3
In firms with 100 to 499 employees	56.5	32.4	11.2	19.7
In firms with 500 to 4,999 employees	50.7	37.7	11.6	19.6
In firms with more than 5,000 employees	63.5	18.8	17.8	15.3
Government employees, excluding the armed forces	59.7	27	13.3	3.9

Source: Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023), January to June 2023. $N = 25{,}014$.

Note: This table considers full-time American workers who are 20–64 years old as of the survey, where "full-time" means working for pay five or more days in the survey reference week. "Fully Onsite" refers to those who worked at their employer's worksite (or a client's location) each workday in the reference week. "Fully Remote" refers to those who worked from home on all workdays in the reference week. "Hybrid Arrangement" refers to those who split the workweek between home and their employer's worksite (or client locations). Column 4 reports the sample percentage of persons in the indicated row. "Government Employees" are those who report their industry as "Government," excluding those who say their occupation is "Armed Forces." We sort persons into the row categories based on their main job. We drop respondents who fail our attention-check questions.

Work-from-Home Intensity by Industry

Figure 2 reports work-from-home rates by industry as of 2023. The Information sector has the highest work-from-home rate at 2.6 days per week among employees who work at least five days a week. AirBnB, Upwork, and Yelp are well-known examples of firms in the Information sector that operate with largely remote workforces. Meta (formerly Facebook), Shopify, and Twitter were among the first prominent companies to commit to high levels of remote work after the pandemic struck. They are also in the Information sector. Finance and Insurance and Professional and Business Services have the next highest work-from-home rates.

These three sectors share certain characteristics that facilitate or incentivize work from home: staff are well-paid, jobs are often analytical or computer-intensive in nature, and firms tend to cluster in major cities. Higher earners typically have nicer homes with more room for a home office. They also face higher marginal tax rates, intensifying the tax incentive to work from home. To see this point, suppose your boss offers the following choice: work at the office five days a week and get a 6 percent raise, or split the workweek between home and office for a 2 percent raise. Your cost of the hybrid option is 3 percent of after-tax pay at a 25 percent marginal tax rate but only 2 percent at a 50 percent marginal tax rate. Turning to

Figure 2

US Work-from-Home Rates by Industry Sector in 2023, Employees Who Work Five or More Days per Week
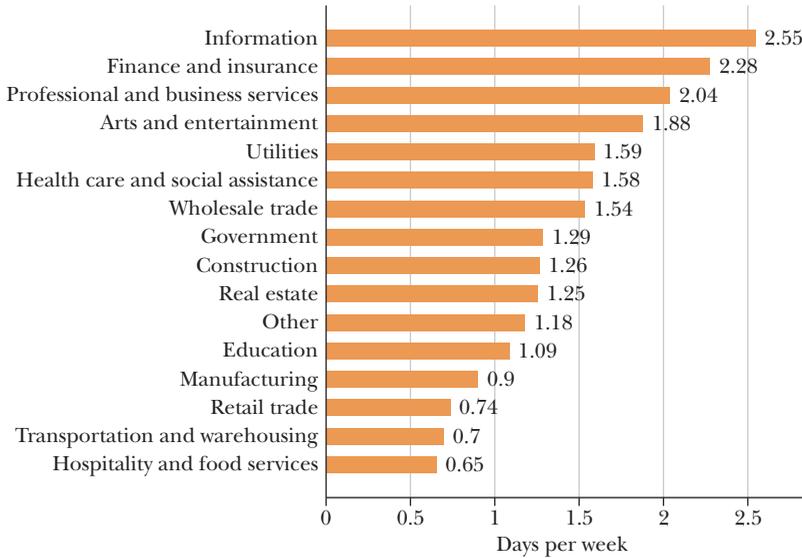
Source: The chart reports mean values for the number of full days worked from home by employees, 20–64 years of age, who worked five or more days in the survey reference week, based on data from the January through June 2023 waves of the US Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023).
Note: Due to small samples, we omit values for Mining and Agriculture. We drop respondents who fail our attention-check questions. $N = 22,341$.

another point, many computer-intensive jobs lend themselves to remote work. And many analytically oriented jobs require periods of intense focus, which is easier to find at home for many workers. Finally, because firms in these sectors cluster in dense urban centers, many of their employees face long commutes to the office. That makes it all the more appealing to avoid the commute, thereby saving time, money, and aggravation.

At the other end of the industry distribution, employees in Retail, Hospitality, Food Services, Transportation, and Manufacturing have low work-from-home rates, ranging from 0.7 to 0.9 days per week. Most jobs in these industries require a physical presence to engage consumers or to work with specialized equipment and dedicated facilities. Because staff work mainly onsite, so do their managers.

There are also some striking differences in work-from-home rates across employers in the same industry. To draw out this point, Hansen et al. (2023) exploit granular data on millions of job vacancy postings. In one of their examples, they consider US automobile manufacturers recruiting for engineers. In 2022, and restricting attention to engineering positions, the share of postings that say the

job allows some remote work was zero for Tesla, 8 percent for Ford, 23 percent for General Motors, and 45 percent for Honda. It was near zero for all four companies in 2019. This example illustrates two of our key messages. First, work-from-home intensity is an outcome of choices about job design, managerial practices, and workplace culture. Organizations can make different choices in this regard, and they do so in the postpandemic economy. Second, the variety of working arrangements on offer has exploded in the wake of the pandemic.

Work from Home and Population Density

Our observations about how and why the work-from-home rate varies by industry suggest that it also varies systematically with local population density. That turns out to be the case, and powerfully so. To develop this point, we first assign zip codes to individual workers based on where they live (home location) or where their employer locates (job location). Next, we use data on residential population to sort zip codes (and workers) into population density bins. Finally, we compute full work-from-home days as a percent of all paid workdays in each bin.

Figure 3 displays the results, using blue dots when sorting by home location and red dots when sorting by job location. The same pattern emerges either way. The figure also highlights selected zip codes. For example, zip code 84720 falls into the second-lowest density bin and is one of two zip codes for Cedar City, Utah. Zip code 11225 falls into the top density bin and is one of more than 100 zip codes for Brooklyn, New York. The most sparsely populated zip codes, accounting for one-fifth of US population, are to the left of the dashed line.

In these sparsely populated areas of the United States, the average work-from-home rate is 25 percent and varies little with population density. In the rest of the country, the work-from-home rate rises strongly with population density. In the top density bin, for example, work-from-home days account for more than 40 percent of all paid workdays. Industry mix, occupation mix, and local workforce demographics (age, sex, education) account for half of the density-related differences in work-from-home rates highlighted by Figure 3 (Buckman et al. 2023). Longer commutes also encourage high work-from-home rates in dense urban areas. High-quality internet service, which is more readily and widely available in urban settings, facilitates work from home and raises its productivity (Barrero, Bloom, and Davis 2021a).

Other types of data tell a consistent story. As of July 2022, foot traffic in central business districts had recovered to nearly prepandemic levels in cities with fewer than 150,000 employees, but to only 60 percent of those levels in cities with 1.5 million or more employees (Monte, Porcher, and Rossi-Hansberg 2023). As of December 2022, the residential price premium for living near downtown returned to prepandemic levels in small cities, but remained considerably smaller than before the pandemic in large cities (Monte, Porcher, and Rossi-Hansberg 2023). From February 2020 to November 2022, Zillow's US home price index rose 40 percent in exurban areas, 32 percent in suburban areas, 18 percent in high-density areas, and only 9 percent in city centers (Ramani and Bloom 2022). As of May 2023, US public

Figure 3
Work from Home Rises with Population Density, US Data
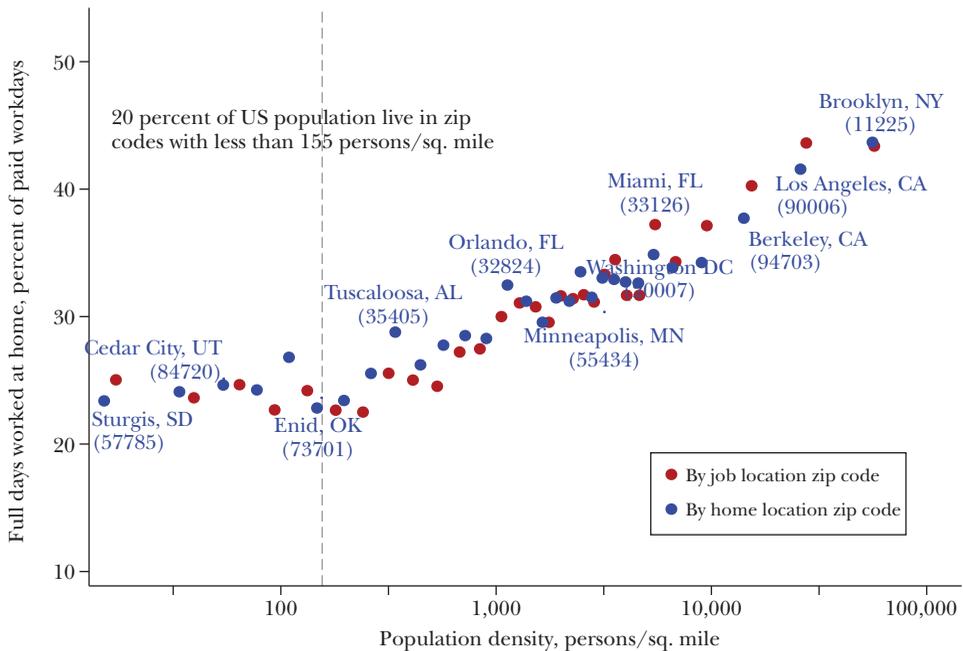
Source: The data are from the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023), January 2021 through June 2023, restricted to persons 20–64 years of age with prior-year earnings of at least $10,000.
Note: The chart shows full days worked from home as a percent of all paid workdays for population density bins. Each density bin is a collection of zip codes and the workers in those zip codes based on their job location (red) or home location (blue). $N = 119,922$ (Job location data). $N = 122,714$ (home location data).

transit ridership was only 70 percent of its May 2019 level.[3] All of these statistics confirm that the big shift to work from home is concentrated in high-density areas.

Work from Home and Demographic Characteristics

Education. Work-from-home intensity rises steeply with education in the postpandemic economy. Among Americans 20–64 years of age with a high school education or less, full days worked from home account for 20 percent of all paid workdays. The corresponding figure is 26 percent for those with some college, 34 percent for those with a four-year college degree, and 36 percent for persons

---

[3] See US Bureau of Transportation Statistics, Public Ridership Transit, retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/TRANSIT, accessed on August 29, 2023.

Figure 4
Work-from-Home Intensity Peaks among Persons in Their Thirties
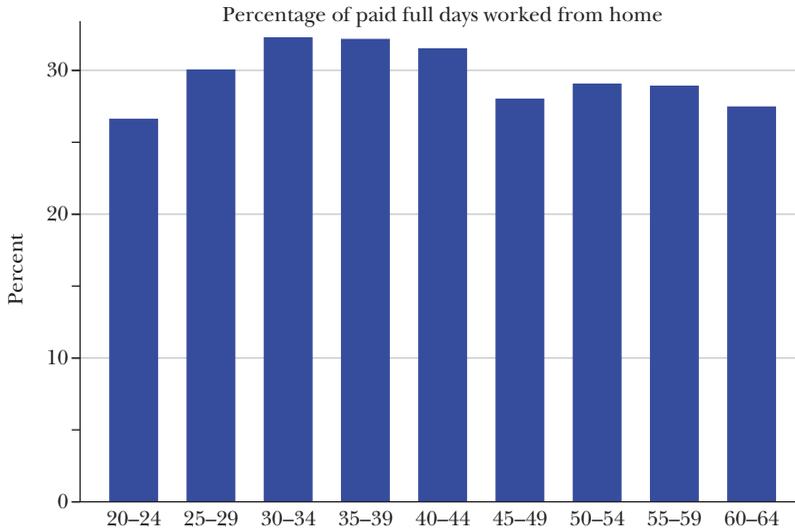
Source: The chart reports full days worked from home as a percent of all paid workdays by age group in the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023).
Note: We drop respondents who fail our attention-check questions. The sample runs from January 2022 through June 2023. $N = 71,000$.

with a graduate degree.[4] As one might anticipate from this pattern, work-from-home intensity also rises steeply with earnings (Barrero, Bloom, and Davis 2021b).

In line with our remarks about jobs as task bundles and the evidence in Dingel and Neiman (2020), occupation and industry are strong predictors of whether someone works from home. Highly educated workers are concentrated in the industry sectors of Information, Finance & Insurance, and Professional & Business Services. These sectors—and the people who work in them—have high work-from-home intensities for the reasons discussed above.

Age. Figure 4 shows how work-from-home intensity varies by age in the United States. It is lowest among people in their early 20s and peaks among those in their 30s. People in their 20s have high returns to professional networking, on-the-job training, and mentoring—activities that benefit greatly from in-person interactions. Young workers may also place more value on socializing at the workplace or nearby. They are more likely to live in small or shared apartments, which reduces the appeal of work from home. People in their 30s and early 40s are more likely to live with children and face long commutes, raising the appeal of

[4] These statistics reflect data from the January–June 2023 waves of the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023).

work from home. Older employees may be less keen to work from home because they no longer have childcare responsibilities, or simply because they like to socialize at the workplace.

In unreported results, we investigate the role of children and marital status in the life-cycle pattern of Figure 4. Consider the raw gap of 5.7 percentage points in work-from-home intensity between workers who are 30–34 years old and those who are 20–24. When we use regression methods to control for the presence of children under 14 in the household, this gap shrinks by one percentage point. Controlling for marital status has no effect on the gap, and marital status itself is statistically insignificant. Similarly, controlling for the presence of children shrinks the raw gap of 4.8 points between workers who are 30–34 and those who are 60–64 by about 1.6 points, with no impact of marital status. In short, the timing of children over the life cycle partly explains the pattern in Figure 4.

Sex. Figure 5 shows, perhaps surprisingly, that American women work from home at only modestly greater rates than men. In the first half of 2023, full days worked from home are 29.3 percent of paid workdays among women, 20–64 years old, as compared to 27.0 percent among men. This gap is statistically significant, but small compared to the gap between college and noncollege workers and to the changes over time. Higher levels of education among women push up their work-from-home rates, other things equal. When we use regression methods to control for educational attainment, the gap between women and men shrinks to 1.3 points. When we also include industry and occupation fixed effects in the regression specification, the gap between men and women is 1.5 percentage points. In other research, Le Barbanchon, Rathelot, and Roulet (2021) find that women place more value than men on reductions in commuting time. That aspect of their preferences pushes in the direction of higher work-from-home intensity for women relative to otherwise similar men.

Parents. People who live with children value the ability to work from home more highly. In a cross-section of advanced and middle-income countries, workers living with children under 14 express a greater willingness-to-pay for the option to work from home two or three days a week, as compared to observationally similar workers without children (Aksoy et al. 2022). The effect holds for men and women and is pervasive across countries.

Moving from preferences to outcomes, we find that people with children do indeed work from home at higher rates, as foreshadowed by our discussion of Figure 4. Table 2 develops this point more fully, again drawing on data from the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023). All specifications control for five-year age bins and month fixed effects. The coefficient on "Female" in column 2 says that full days worked from home (as a percent of all paid workdays) are 1.0 percentage points higher for women than men, the omitted group. The other coefficient in column 2 says that full days worked from home are 4.5 percentage points higher for workers who live with children under 14. Column 3 adds a term to capture the interaction between "Female" and "Children under 14." The –2.3 coefficient on this term is statistically significant at the

Figure 5
Women and Men Work from Home at Similar Rates, US Data

	2022	*2023 H1*
All	30.2	28.1
Female	31.3	29.3
Male	29.4	27
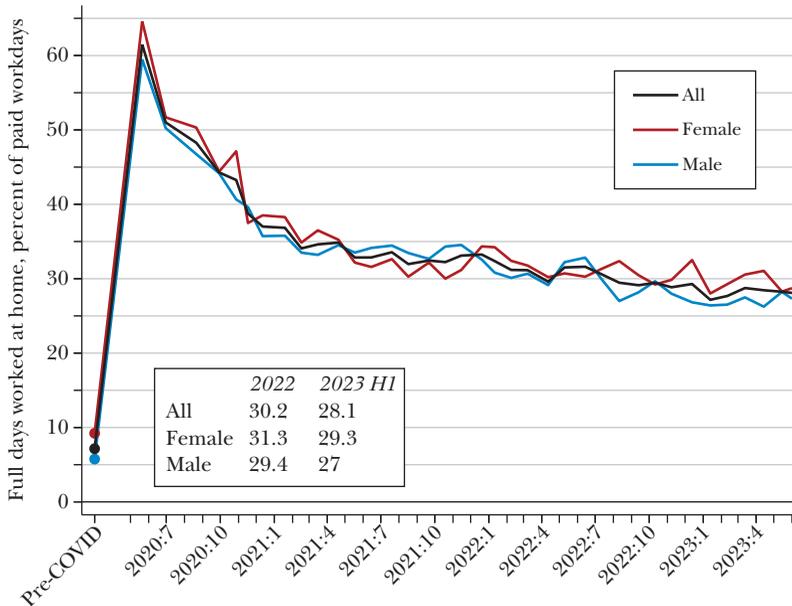
Source: The chart reports full days worked from home as a percent of all paid workdays for men and women in the Survey of Working Arrangements and Attitudes (SWAA) (Barrero et al. 2020–2023) and 2019 American Time Use Survey (ATUS) (Flood et al. 2023).
Note: The SWAA data cover May 2020 and each month from July 2020 through June 2023. We restrict samples to workers 20–64 years of age and 2019 earnings of at least $20,000 before March 2021 and then transition to a $10,000 threshold in 2019 or the prior year. We reweight the SWAA to match the distribution across age-sex-education-earnings cells in the Current Population Survey. We use regression methods to adjust for changes in the question design over time, as detailed in https://wfhresearch.com/tracking-wfh/. We obtain the pre-COVID numbers from ATUS by determining whether respondents worked for at least six hours in a day, and whether the location of that work was their home. $N = 1,885$ (2019 ATUS) and $N = 143,410$ (SWAA). 2023 H1 denotes the first half of 2023.

5 percent level. In other words, living with children is associated with a larger marginal increase in work-from-home intensity for men than women. The coefficient on the main effect for "Children under 14" is now 5.5 percentage points.

Finally, when we add controls for the worker's education, industry, and occupation in columns 4 and 5, we continue to find higher work-from-home rates among those who live with children. The results in columns 4 and 5 also confirm that education is a powerful predictor of work-from-home intensity. (Here, the omitted group is persons with no postsecondary education.) However, the coefficients on the main and interaction effects for women are no longer statistically significant.

Table 2
How Work-from-Home Rates Covary with Individual Characteristics

	Full days worked at home as percent of paid workdays				
	(1)	(2)	(3)	(4)	(5)
1(Female)	0.9*	1.0**	1.9***	–0.2	0.6
	(0.5)	(0.5)	(0.6)	(0.6)	(0.6)
1(Lives with child under 14)		4.5***	5.5***	2.6***	1.6**
		(0.5)	(0.7)	(0.7)	(0.7)
1(Female) × 1(Lives with child under 14)			–2.3**	–0.0	0.5
			(1.0)	(0.9)	(0.9)
1(one to three years of college)				7.0***	5.1***
				(0.7)	(0.6)
1(four-year college degree)				16.5***	11.4***
				(0.6)	(0.7)
1(Graduate degree)				19.1***	13.4***
				(0.7)	(0.8)
Industry and occupation fixed effects					Y
N	48,244	48,244	48,244	48,244	48,244
R^2	0.01	0.01	0.01	0.04	0.11

Source: Survey of Working Arrangements and Attitudes, www.wfhresearch.com (Barrero et al. 2020–2023)
Note: We use data from the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023) covering October 2021 to October 2022 (inclusive) and regress full days worked at home as percent of paid workdays on indicators for sex, for whether the respondent lives with a child under 14, and education categories. All columns include monthly survey wave fixed effects, and fixed effects for five-year age bins (e.g. 25 to 29, 30 to 34, etc.). 1(. . .) denotes the indicator function. The sample includes respondents who worked during the reference week, pass our attention-check questions, and have nonmissing data on occupation and industry of the current or most recent job. We report standard errors in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Let us summarize the demographic patterns. First, work-from-home intensity rises strongly with the education of individual workers. This pattern is highly robust to controls for other demographic characteristics, marital status, and the presence of children at home. The presence of children is also robustly associated with more work from home. Women work from home slightly more than men, but this already-small difference shrinks further or disappears altogether when controlling for women's greater education. Finally, the small R-squared values in the Table 2 regressions warrant attention, because they imply that many other factors also influence who works from home and how much.

Other Countries

The extent of work from home differs greatly across countries, although it is hard to say exactly how much. One useful source of information is Google

Workplace Mobility, which, until October 2022, tracked the frequency of workplace visits by country and month. As of October 2022, workplace visits in the United States were down about 25 percentage points relative to the January 2020 benchmark. That drop roughly matches the corresponding jump (relative to 2019) in work-from-home days shown in Figure 1. Canada and the United Kingdom show similar-sized drops in workplace visits (again, as of October 2022, and relative to the prepandemic benchmark). A selection of European countries (France, Germany, Italy, Spain, the Netherlands, Sweden, and Portugal) shows a smaller drop of 21 percentage points. Elsewhere, the drops are smaller yet. For example, workplace mobility is down only nine points in Australia as of October 2022, and by even less in Japan, South Korea, Singapore, and Taiwan.

Along with coauthors in Aksoy et al. (2022, 2023b), we have developed another source of data on work-from-home intensity that covers many countries: The Global Survey of Working Arrangements (https://wfhresearch.com/gswa-data), which samples full-time employees who have at least a secondary education. According to these data, full-time employees worked from home 1.4 days a week in the United States as of March-April 2023, more than any other country except Canada (1.7 days) and the United Kingdom (1.5 days). By way of comparison, the average across 15 countries in continental Europe is only 0.8 days a week, with a maximum of 1.0 days in Germany and The Netherlands. The average across six Asian countries (China, Malaysia, Japan, Singapore, South Korea, and Taiwan) is 0.7 days a week.

Why does the United States have a higher work-from-home intensity than most other countries? We think several factors are at work. First, residences tend to be larger in the United States than in Europe and Asia, making it easier to accommodate a home office. Second, the United States has an unusually high share of workers in Information, Finance and Insurance, and Professional and Business Services. As discussed earlier, these sectors share several characteristics that facilitate or encourage work from home. Third, managerial practices around performance measurement and evaluation tend to be more advanced at US firms than at firms in other countries (for example, Scur et al. 2021). Strong practices in this regard are especially beneficial when managers (and coworkers) cannot observe employees directly as they perform their work-related activities.

Finally, most countries in Asia and many in Europe were more successful than the United States in mitigating COVID-related deaths and hospitalizations. And they often did so without tight, extended clampdowns on commercial activities. South Korea and Taiwan are especially striking examples in this regard. Because of their less dire COVID experiences—and, in many cases, their more limited clampdowns on commercial activity—people and organizations in other countries experimented less with remote work, made fewer investments in systems and management practices that support remote work, had less learning-by-doing in work-from-home mode, and had smaller shifts in attitudes and norms around remote work. In short, several factors that reinforced the big shift to work from home in the United States acted with less force in most other countries.

What the Big Shift Means for Pay, Productivity, and Innovation

Labor Costs and the Wage Structure

The rise of remote work affects labor costs and wages in several ways. For one thing, it can have important effects on productivity, which feeds into labor costs and wages. For the moment, and for the sake of analytical clarity, let us consider how remote work affects labor costs and wages net of any productivity effects. The rise of remote work makes it easier for firms situated in high-wage areas to recruit and employ staff in areas with lower wages. There is also evidence that quit rates and turnover costs fall when a firm lets its employees adopt hybrid working arrangements (Bloom, Han, and Liang 2023). And standard economic models imply that the rise of remote work puts downward pressure on real wages through labor supply effects.

To see this last point, start by observing that most people prefer to work remotely part of the week. On average, Americans value the option to work from home two or three days a week at 8 percent of pay, according to data from the Survey of Working Arrangements and Attitudes (Barrero et al. 2020–2023). The experimental study by Mas and Pallais (2017) finds similar estimates for a narrower group of workers. In the textbook model of a competitive labor market, a new-found job amenity raises the supply of labor at any given wage. In addition, when employees work from home, they devote part of their commute-time savings to their jobs (Aksoy et al. 2023a). That also adds to labor supply. Finally, the big shift creates new job opportunities for people who live in remote locations and economically depressed areas, couples facing joint-location constraints, parents who want to be near their children while working, people with mobility impairments that make it hard to commute, and those who find it hard to cope with face-to-face interactions in the workplace. New job opportunities afforded by work from home will draw some of these people into the labor force, expanding labor supply. For any given labor demand curve, outward shifts in labor supply reduce the equilibrium real wage. In an inflationary environment, this real wage drop can come about through slower nominal wage growth.

Bargaining models also imply that the rise of remote work puts downward pressure on wages. In this class of models, the employer and employee each get a share of the value generated by the employment relationship. Initially, employees captured the full amenity value of the big shift to work from home, because wages reflected bargains made before the surprise onset of the pandemic. Over time, as compensation packages adjust, employers capture a portion of the amenity-value gains associated with the rise of remote work. Employer benefits take the form of slower wage growth during the transition to a new equilibrium with pay packages that reflect higher remote work levels.

Has the rise of remote work actually exerted a material influence on wages and other labor costs? If so, how much? Thus far, these questions have attracted scant attention from researchers. In Barrero et al. (2022), we provide some initial evidence. We surveyed more than 600 business executives at US firms in April and

May 2022 as part of the Survey of Business Uncertainty. We asked each executive the following question: "Over the past twelve months, has your firm expanded the opportunities to work from home (or other remote locations) as a way to keep employees happy and to moderate wage-growth pressures?" If the executive says "Yes"—and about four in ten do so—we then ask, "What is your best estimate for how much expanded remote-work opportunities have moderated wage-growth pressures at your firm in the past twelve months?" We also ask analogous questions about work-from-home plans in the next twelve months. When we aggregate over all responses to the look-back and look-ahead questions, we obtain a cumulative wage-growth moderation of 2.0 percentage points over two years, centered on spring 2022. In computing this average, we include the majority of executives who say work from home has not impacted their firm's wages.

An effect of 2.0 percentage points is a material restraint on nominal wage growth. It came during a critical period for monetary policy, as the Fed sought to curtail the strong inflationary pressures that emerged in the latter part of 2021 and 2022. The same survey also uncovers evidence that the rise in remote work is associated with more reliance on part-time employees (at 26 percent of firms), more use of independent contractors (23 percent), more use of leased workers (11 percent), more offshoring of jobs (8 percent), and more employment of physically-challenged persons. These developments also tend to reduce labor costs.

The big shift to work from home is likely to alter the structure of wages as well. To see how, recall that the shift is concentrated among college-educated workers and in certain industry sectors (as shown in Figure 2). The amenity-value benefits associated with the big shift are concentrated among these same workers and sectors. A long line of thinking in economics says that wages are lower, other things equal, in jobs with amenity attributes that workers like.[5] Thus, we hypothesize that the big shift shrank the college wage premium and put more downward pressure on wages in sectors with larger increases in work-from-home intensity.

While the available evidence is limited, it favors this hypothesis. In their survey of business executives, Barrero et al. (2022) find smaller wage-growth moderation effects in sectors with few jobs that are suitable for remote work and larger moderation effects in sectors with many such jobs. Autor, Dube, and McGrew (2023) document a large and "unexpected compression" in the US wage distribution after the pandemic struck, including a reduction in the college wage premium. Their explanation stresses the pandemic's effects on labor market tightness and wage markdowns, but they observe that amenity-value shocks may also play a role.

There is room for much more research into how the big shift affects labor costs and the wage structure. In addition to the channels we have stressed, work from home has the potential to influence profoundly the extent of spatial competition in labor markets, including monopsony power in local labor markets. Especially in

---

[5] Rosen (1986) offers a classic statement of this theory of "equalizing differences" or "compensating differentials." As he notes, the basic idea originates in *The Wealth of Nations* by Adam Smith. For a recent discussion of compensating differentials in this journal, see Lavetti (2023).

fully remote jobs, competition from workers in other countries can exert a powerful influence on wages. For evidence on this theme using data from a web-based job platform with global reach, see Brinatti et al. (2022).

Productivity Effects

The productivity implications of the big shift to work from home have sparked vigorous debate among business leaders, researchers, and pundits. Managers differ greatly in their views about how remote work affects productivity in their own organizations (Bloom et al. 2023a). Worker perceptions also differ greatly in this regard (Barrero, Bloom, and Davis 2021b).

The controversies and disagreements surrounding the productivity effects of work from home reflect the complex nature of the issue. Jobs and tasks differ greatly in their suitability for remote work, as do workers, managerial styles, and workplace cultures. Thus, there is no sound reason to expect the productivity effects of remote work to be uniform across jobs, workers, managers, and organizations. In addition, communications, performance evaluations, and management practices must adapt to new working arrangements, if they are to work well. Adaptation often requires new skills, perhaps especially for managers. The adaptation process itself involves a good deal of trial and error and learning-by-doing. Thus, the productivity effects of work from home are likely to unfold over months or years.

With these cautions in mind, we now offer several remarks on how productivity relates to working arrangements. For us, the chief question of interest is the following: How will a large, lasting shift to work from home affect productivity outcomes when working arrangements are a matter of choice, rather than necessity, and when organizations and individuals have adapted to their preferred working arrangements? Previous studies differ in whether, and how clearly, they throw light on the answer to this question.

Productivity concepts. From a manager's perspective, it makes sense to think of labor productivity as work output per paid time unit. This labor productivity concept corresponds to what statistical agencies typically seek to measure. For employees, in contrast, it makes sense to count commuting time as part of the total time devoted to their jobs. This approach also makes sense from a societal perspective, because commuting time is part of the resources expended in producing market output.

This conceptual distinction matters. Recall from Figure 1 that full days worked from home rose by 21 percentage points from 2019 to June 2023. That corresponds to 1.05 fewer commuting days per week. When working from home (rather than on employer premises), the average daily savings in commuting and grooming time is 65 minutes for American workers. So, the big shift generates time savings of 68 minutes per week per worker, which is about 2.8 percent of a 40-hour workweek. Standard productivity measures published by statistical agencies miss this source of productivity gains associated with the big shift.[6]

---

[6] A fuller analysis of the missing productivity gain would account for three other factors. First, the big shift is concentrated among workers with higher earnings, who also tend to have longer commutes. Adjusting

Productivity perceptions. On balance, workers hold more favorable views about productivity in work-from-home mode than managers. Much of this difference in perceptions turns on the same conceptual distinction. Consider someone who works eight paid hours a day, lives 30 minutes from the office, and accomplishes the same amount whether working from home or the office. In this example, total time devoted to work is nine hours per day when commuting, and eight hours per day when working from home. So, the worker perceives (correctly) that he or she accomplishes the same amount in 11 percent less time when working from home—a big productivity boost! From the manager's vantage point, however, the productivity effect of work from home is nil in this example. Thus, managers and workers can quite reasonably hold different views about the productivity effects of work from home.

According to data from the Survey of Working Arrangements and Attitudes covering January to June 2023, 43 percent of workers able to work from home say they are more productive when working from home, and 14 percent say they are less productive. The rest say their productivity is about the same either way. Probing further, workers attribute the self-assessed productivity benefits of work from home mainly to the savings in commute time (Barrero, Bloom, and Davis 2021b). When we net out the productivity benefit that workers attribute to commute time-savings, the manager-worker gap in perceptions about the productivity effects of work from home shrinks to a small difference.

There is a broader lesson here: Disagreements about the productivity effects of work from home turn partly on what counts as productivity. Workers regard commute time-savings as a source of productivity gains, while managers do not. Workers may be blind to the managerial challenges of remote work. Managers, especially senior ones, are probably more concerned about its implications for workplace culture. Remote work may let seasoned employees accomplish more in the near term while, at the same time, detracting from the transmission of their knowledge to younger coworkers—with detrimental effects on the organization's productivity over the longer term.

Fully Remote Work. Several studies find that fully remote work yields lower productivity than on-site work. Emanuel and Harrington (2023) analyze data from a Fortune 500 firm that, before the pandemic, operated call centers with both remote and on-site employees in the same jobs. In response to the pandemic, the firm shifted all employees in these jobs to fully remote work. Productivity among formerly onsite employees fell 4 percent relative to that of already-remote employees. Emanuel and Harrington also find evidence that the closure of physical call centers reduced call quality, especially among less experienced employees.

---

for these patterns would raise the 2.8 percent figure, especially on an earnings-weighted basis. Second, workers devote between 30 and 40 percent of their commute time-savings to their jobs. Adjusting for that would lower the 2.8 percent figure. Third, hours worked are observed with error in the data that statistical agencies use to measure labor productivity, and the observational errors may be correlated with the big shift itself. See Davis (2022) for additional discussion.

These findings are noteworthy, in part, because they involve a firm with prior experience in managing fully remote call-center workers. Presumably then, the firm had already adapted its systems and practices to manage fully remote workers.

Gibbs, Mengel, and Siemroth (2023) study productivity outcomes for skilled professionals at a large Indian technology services company. In March 2020, the company abruptly shifted all employees to fully remote work in response to the pandemic. Immediately after the shift, average worktime rose by 1.5 hours per day and output fell slightly according to their primary performance measure. They estimate that the shift to remote work lowered average labor productivity (output per hour worked) by 8 to 19 percent. They also provide evidence that greater communication and coordination costs drove much of the measured productivity drop. In particular, time spent on meetings and coordination activities rose, crowding out time devoted to a concentrated focus on work tasks.

The productivity concept matters here, as well. Table 2 of Gibbs, Mengel, and Siemroth (2023) reports an average two-way commute time of 1.3 hours per day. That is, the savings in commute time largely offsets the extra 1.5 hours per day that employees put into their jobs after the transition to fully remote work. From the worker's perspective, and from a societal perspective as well, the company's shift to remote work had small effects on productivity. The larger point is that the commute time-savings from remote work can offset sizable drops in productivity, as conventionally measured.

Gibbs, Mengel, and Siemroth (2023) also find that employees "narrowed the scope of their networks" after shifting to fully remote work, engaging in fewer contacts inside and outside the organization. Other studies also find scope-narrowing and siloing effects associated with remote work. For example, Yang et al. (2022) find that communications among 61,000 Microsoft employees became more asynchronous after a pandemic-induced shift to remote work, and collaborations became more static and siloed. Battiston, Blanes i Vidal, and Kirchmaier (2021) find that police dispatchers work faster, especially in busy periods, when they sit in the same room as other dispatchers. These three studies suggest, in various ways, that remote work can slow communications, impede the diffusion of knowledge within an organization, and narrow the scope of collaborative efforts.

Many managers and workers also express concerns about mentoring and learning in remote settings. According to a survey by the Pew Research Center, 36 percent of American employees who work from home all, most, or some of the time believe that teleworking hurts their opportunities to be mentored at work "a lot/a little" (Parker 2023). However, 10 percent think it helps. With respect to "Chances of getting ahead in their job" and "Whether they are given important assignments," large majorities think that working from home neither helps nor hurts, and the rest are evenly split between helps and hurts. One interpretation of these results is that the impact of work from home on mentoring and job advancement depends on the person, the job, and the way the organization manages its remote workers.

In a study of mentoring practices and team relations, Emanuel, Harrington, and Pallais (2023) consider software engineers at a large technology firm. Some engineering teams were housed in the same building, and other teams were split across two adjacent buildings. Before the pandemic, employees housed in the same building as teammates received 21 percent more comments on their code from coworkers. These comments provide suggestions on how to improve code, and they play an important role in employee learning and performance gains. When the pandemic struck, all employees shifted to work from home. The comment rate difference between colocated and other teams vanished, and overall comment volume fell by almost half. The volume drop was largest for junior employees, reinforcing concerns that remote work impedes knowledge transmission to younger workers.

Other concerns about the productivity effects of remote work center on motivation and self-control. As one gag line has it, the three enemies of working from home are the bed, the refrigerator, and the television. Indeed, college students sometimes study in libraries as a commitment device—even though grades, self-esteem, and future career prospects would seem to provide powerful incentives already. So, it is perhaps unsurprising that some people struggle with motivation when working in a remote capacity. Armed with a bit of self-knowledge, they may also choose to work at their employer's site despite the attractions of work from home.

A fuller analysis of why organizations employ fully remote workers, even when it means lower labor productivity, would look to other cost savings. We discussed how remote work can lower labor costs. In addition, fully remote employees do not require office space and the overhead costs that come with a physical footprint. These cost savings must be considered along with the productivity consequences, now and in the future, when optimizing over working arrangements. At the organizational level, this optimization problem is a complex one, because a shift in working arrangements also requires new skills and work habits, new managerial practices, and new organizational capabilities if remote work is to be effective.

Hybrid arrangements. Studies of hybrid working arrangements often find productivity gains (relative to traditional arrangements) or no discernable effect. In an early study, Bloom et al. (2015) consider a field experiment with 250 call-center employees at cTrip.com, a large Chinese travel agent. Employees were randomly allocated between two groups. One group continued working five days a week in the office, and the other group switched to a hybrid arrangement with four home workdays and one office workday each week. Average daily output rose by 13 percent among switchers. Nine percentage points of this output gain arose from extra worktime due to shorter breaks and less sick leave. The other four points reflect greater efficiency in the form of more calls handled per work hour.

In another prepandemic experiment, Choudhury, Foroughi, and Larson (2021) study the productivity of patent examiners employed by the US Patent and Trademark Office. Starting in 2006, several hundred examiners were allowed to shift to a hybrid arrangement with up to four home workdays per week. The shift took place in a staggered manner, with exogenous timing at the worker level, facilitating

the estimation of causal effects. Patent actions rose by 5 percent, on average, after examiners switched to the hybrid arrangement. The switch had no impact on the incidence of rework, indicating that quality did not suffer. Later, starting in 2012, patent examiners were allowed to live anywhere in the contiguous United States with periodic travel to headquarters at their own expense. Patent actions rose by an additional 8 percent among those who switched from the hybrid to the work-from-anywhere arrangement. Choudhury et al. provide evidence that these productivity gains reflected greater effort by patent examiners.

Three other aspects of the setting in this study warrant attention. First, patent examinations are performed by individuals, not teams. Second, examiners were required to spend at least two years working in the office before transitioning to the hybrid arrangement—and more time in that arrangement before transitioning to a work-from-anywhere arrangement. Third, individual examiners chose whether to switch from a traditional working arrangement to the alternative arrangements. Thus, the selection process into the alternative arrangements reflects individual preferences and self-knowledge about ability to work remotely.

Other recent studies analyze more modest shifts from traditional to hybrid working arrangements. Choudhury et al. (2022) consider an organization in Bangladesh that randomly assigned administrative staff to three groups: high work-from-home intensity (more than 40 percent of workdays), intermediate intensity (23-40 percent), and low intensity (0–23 percent). Staff in the intermediate group sent more emails to more people, drafted more complex emails, reported better job satisfaction, felt more connected to colleagues, and received stronger performance evaluations from managers. Bloom, Han, and Liang (2023) conduct a randomized control trial with 1600 highly educated employees (software engineers and marketing and finance professionals) at cTrip.com. Those born on even-numbered dates continued to work onsite five days a week, while the rest had the option to work from home on Wednesday and Friday. The experiment ran for six months, yielding zero or small productivity gains from the hybrid arrangement. Employees highly valued the hybrid arrangement, except for managers. Angelici and Profeta (2023) consider a nine-month experiment that injected flexibility into the working arrangements of full-time employees at a large Italian firm. The control group stuck to a traditional arrangement that prescribes time and place of work throughout the week. In the treated group, white-collar employees chose where and when to work—and blue-collar employees chose when—one day a week. Productivity rose 10 percent, on average, among the treated relative to the controls, as measured by self-assessments and by supervisors. The treated group also reported higher levels of well-being and work-life balance. Greater flexibility for treated workers had no apparent effect on the productivity of coworkers on the same team.

All three of these studies suggest that working from home one or two days a week improves productivity and leads to happier employees. This pattern supports the view that, in many jobs, some tasks are well suited for remote work. The productivity gains associated with hybrid working arrangements in these studies could

reflect greater effort levels by happier employees, quieter work environments at home, and the time savings that employees put back into their jobs.

The Pace of Innovation

Face-to-face interactions in the workplace can foster the diffusion of knowledge and the generation of new ideas. Looking beyond a single workplace, cities have long functioned as hubs for knowledge spillovers across workers, firms, and industries and as centers of invention, innovation, and entrepreneurship.[7] In this light, the big shift to work from home raises concerns about its potential to slow the pace of innovation and the growth rate of productivity.

In one study that speaks to these concerns, Brucks and Levav (2022) designed and fielded commercial innovation experiments in cooperation with employers in Finland, Hungary, India, Israel, and Portugal. More than 800 engineering teams were tasked with suggesting new uses for an existing product, with teams randomly assigned to meet in person or via videoconference. In-person teams proved more effective, in that their product ideas received higher external ratings. When it came to selecting ideas (rather than developing them), teams performed as well, or even better, when they met by videoconference. These results suggest that brainstorming activities benefit from in-person meetings, but some other aspects of the innovation process do not. In a similar spirit, Lin, Frey, and Wu (2023) develop evidence that collaboration at a distance is especially challenging in the early stages of research "when an idea is hard to articulate and knowledge is tacit." They show that inventor teams that collaborate remotely make fewer breakthroughs, as measured by patent citations. The same pattern holds for scientific publications. Lin, Frey, and Wu also show that collaboration in dispersed teams is more focused on technical tasks, including data collection and analysis, rather than idea generation and research design. Both studies suggest that in-person communications are particularly valuable in the early stages of the innovation process but less so in later stages and in technical tasks.

Profit-oriented firms have strong incentives to recognize and respond to the internal costs and benefits of working arrangements. However, if those arrangements yield idea spillovers beyond the boundaries of the firm, outcomes need not be efficient or socially desirable. Monte, Porcher, and Rossi-Hansberg (2023) pursue this theme. In their calibrated model, a city can settle at a low or high work-from-home intensity. The low-intensity equilibrium yields more spillovers, greater productivity, and higher social welfare in their large-city calibration. Still, that equilibrium can be disrupted by a shock that, for a period time, renders it infeasible or unattractive to commute to the worksite. Even after the shock ends, workers and firms can settle at a new equilibrium with less commuting (a benefit), lower productivity (a cost), and lower social welfare. Their model offers an alternative explanation for the enduring character of the big shift to work from home after the pandemic. Although Monte,

---

[7]The literature is vast and diverse. See Carlino and Kerr (2015) and Combes and Gobillon (2015) for reviews.

Porcher, and Rossi-Hansberg focus on how work-from-intensity affects the level of productivity, a similar point applies to the pace of innovation. In other words, it is possible that a shock-induced shift to work from home could slow the pace of innovation by undermining the idea-generating capabilities of cities.

Despite these reasons for concern, we are optimistic about the outlook for the pace of innovation, at least as it relates to working arrangements. The same developments that facilitated a big shift to work from home also created greater reach and higher quality in one-way and two-way communications at a distance. These include the rise of the internet, better broadband infrastructure, better videoconferencing, the emergence of the cloud, and better software tools for remote collaboration.

Chen, Frey, and Presidente (2022) also study the relationship of remote collaboration to the impact of scientific articles, as reflected in citations. Before 2010, remote collaboration produced articles that were more incremental in nature and less likely to yield "disruptive" advances, echoing the findings in Lin, Frey, and Wu (2023). However, Chen, Frey, and Presidente also show that the quality discount on articles written by dispersed teams shrinks over time, vanishes around 2010, and then becomes a premium. A plausible explanation is that advances in remote-collaboration technologies made it easier and cheaper to coordinate a broader range of specialized and geographically scattered complementary inputs. In the model of Becker and Murphy (1992, Section 6) such a fall in coordination costs raises the innovation rate.

In an earlier study, Forman and van Zeebroeck (2012) examine patenting activity inside firms before and after the adoption of internet technologies, which happened at different times in different firms. Adoption led to more patent-producing collaborations by geographically dispersed teams within the firm. Evidently, internet access lowered coordination costs among geographically dispersed scientists in the same firm. In later work, Forman and van Zeebroeck (2019) find that the spread of internet connectivity increased knowledge flows across locations, as reflected in between-location patent citations within firms.

In conversations with work-from-home skeptics, we often hear some version of the following claims: Working from home means a loss of serendipitous encounters in and around the workplace. Or, knowledge flows more readily among people who work in the same location. Another claim is that people generate more and better ideas, and bring them to fruition more quickly, when they work in close proximity. Even if these claims are true, they miss much.

First, there is an opportunity cost to chatting with your coworker in the next-door office: You could be collaborating with your far-away coauthor via Zoom. It is hardly obvious that serendipitous encounters in the workplace foster innovation better than planned encounters selected from a much larger universe. Second, remote-collaboration technologies tremendously expand the opportunities to form teams that are optimized for particular research projects and questions. This advantage in team formation is especially valuable in projects that call for the cooperation of many scientists with distinct skills and types of expertise. Third, if some aspects of the

innovation process truly benefit from in-person communications, then researchers have strong incentives to focus on those aspects when they are in the same location. Hybrid working arrangements, in particular, do not prevent in-person meetings. They only require more forethought and coordination to make them happen.

Looking Ahead

At the outset of this essay, we explained how the pandemic catalyzed a big shift to work from home and why it will endure. Statistical evidence confirms this conclusion. According to Figure 1, the US work-from-home rate in 2023 has settled at about four times the 2019 level. Office occupancy statistics from Kastle Security Systems (n.d.) also indicate that work-from-home intensity has stabilized in 2023.

As part of the July 2023 Survey of Business Uncertainty, fielded by the Atlanta Federal Reserve, we asked US business executives about the work-from-home outlook at their own firms. The survey responses cover about 500 firms distributed widely across industries, states, and firm size categories. Specifically, we asked: "Looking forward to five years from now, what share of your firm's full-time employees do you expect to be in each category [fully in person, hybrid, fully remote] in 2028?" We asked a parallel question about the firm's current working arrangements. Compiling the results, executives anticipate modest increases over the next five years in both the fully remote share and the hybrid share (Bloom et al. 2023b). This finding holds whether we weight each firm equally or in proportion to its number of employees.

Another reason to anticipate steady or slowly rising work-from-home rates over the next several years is that organizations will continue to adapt their practices to manage hybrid and fully remote workers more effectively. That will raise productivity in work-from-home mode. Where experience teaches that remote work is unsuitable, organizations will revert to traditional arrangements, if they not have done so already.

Yet another reason involves the innovation incentives created by the big shift. A growing market provides incentives for investments in innovations to serve that market (Schmookler 1966). The US market for technologies and products that support remote work is now four times as large as in 2019. It has also become much larger in the rest of the world. To assess the force of the Schmookler effect, Bloom, Davis, and Zhestkova (2021) consider the monthly flow of newly filed applications for US patents. They use automated text readings to determine which ones claim to advance technologies in support of video conferencing, telecommuting, remote collaboration, and work from home. Patent applications that advance these technologies double as a share of all newly filed US patent applications from January to September 2020. In ongoing research with Mihai Codreanu, we find that this redirection of innovation efforts has continued through at least early 2022. So, it is reasonable to anticipate that remote-collaboration technologies and tools will continue to advance at a rapid pace for some years to come, further reinforcing the shift to remote work.

We see the outlook for fully remote jobs as less secure than the one for hybrid arrangements. Cost pressures encourage firms to shift fully remote jobs to countries with lower wages, including countries with many English-proficient workers. Of course, that process has been underway for decades. The new element is that firms now have better information about which jobs and tasks are well suited for remote work. The automation of call-center jobs and routine administrative tasks may also reduce the number of fully remote jobs in the United States.

Work-from-home rates will continue to differ sharply across industries, occupations, cities, and worker education levels. One corollary is that an economic downturn concentrated in remote-intensive sectors could lower the overall work-from-home rate for a time. Another corollary is that the amenity-value benefits of the big shift will also remain unevenly distributed across sectors, occupations, and demographic groups. As a result, we think the big shift will have lasting effects on the structure of wages. It may also profoundly influence the nature and intensity of competition in the labor market.

References

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate. 2022. "Working from Home around the World." *Brookings Papers on Economic Activity* 53 (2): 281–330.

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate. 2023a. "Time Savings When Working from Home." *AEA Papers and Proceedings* 113: 597–603.

Aksoy, Cevat Giray, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate. 2023b. "Working from Home around the Globe: 2023 Report." EconPol Policy Brief 53.

Angelici, Marta, and Paola Profeta. 2023. "Smart Working: Flexibility without Constraints." *Management Science*. https://doi.org/10.1287/mnsc.2023.4767.

Autor, David, Arindrajit Dube, and Annie McGrew. 2023. "The Unexpected Compression: Competition at Work in the Low Wage Labor Market." NBER Working Paper 31010.

Barrero, Jose Maria, Nicholas Bloom, Shelby Buckman, and Steven J. Davis. 2020–2023. "WFH Code and Data: May 2020 to August 2023." WFH Research. https://wfhresearch.com/data-restrict/ (accessed on September 2, 2023).

Barrero, Jose Maria, Nicholas Bloom, Shelby Buckman, and Steven J. Davis. 2023. "SWAA February 2023 Updates." https://wfhresearch.com/wp-content/uploads/2023/02/

WFHResearch_updates_February2023.pdf.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis. 2021a. "Internet Access and Its Implications for Productivity, Inequality, and Resilience." In *Rebuilding the Post-Pandemic Economy*, edited by Melissa S. Kearney and Amy Ganz, 18–51. Washington, DC: Aspen Institute Press.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis. 2021b. "Why Working from Home Will Stick." NBER Working Paper 28731.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis. 2023. "Replication data for: The Evolution of Work from Home." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193587V1.

Barrero, Jose Maria, Nicholas Bloom, Steven J. Davis, Brent Meyer, and Emil Mihaylov. 2022. "The Shift to Remote Work Lessens Wage-Growth Pressures." NBER Working Paper 30197.

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier, 2021. "Face-to-Face Communication in Organizations." *Review of Economic Studies* 88 (2): 574–609.

Becker, Gary S., and Kevin M. Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge." *Quarterly Journal of Economic,* 107 (4): 1137–60.

Bloom, Nicholas, Jose Maria Barrero, Steven J. Davis, Brent Meyer, and Emil Mihaylov. 2023a. "Research: Where Managers and Employees Disagree about Remote Work." *Harvard Business Review,* January 5. https://hbr.org/2023/01/research-where-managers-and-employees-disagree-about-remote-work.

Bloom, Nicholas, Jose Maria Barrero, Steven J. Davis, Brent Meyer, and Emil Mihaylov. 2023b. "Survey: Remote Work Isn't Going Away—And Executives Know It." *Harvard Business Review,* August 28. https://hbr.org/2023/08/survey-remote-work-isnt-going-away-and-executives-know-it.

Bloom, Nicholas, Steven J. Davis, and Yulia Zhestkova. 2021. "COVID-19 Shifted Patent Applications toward Technologies That Support Working from Home." *AEA Papers and Proceedings* 111: 263–66.

Bloom, Nicholas, Ruobing Han, and James Liang. 2023. "How Hybrid Working from Home Works Out." NBER Working Paper 30292.

Bloom, Nicholas, James Liang, John Roberts, and Zhichun Jenny Ying. 2015. "Does Working from Home Work? Evidence from a Chinese Experiment." *Quarterly Journal of Economics* 130 (1): 165–218.

Brinatti, Agostina, Alberto Cavallo, Javier Cravino, and Andres Drenik. 2022. "The International Price of Remote Work." NBER Working Paper 29437.

Brucks, Melanie S., and Jonathan Levav. 2022. "Virtual Communication Curbs Creative Idea Generation." *Nature* 605: 108–12.

Buckman, Shelby, Jose Maria Barrero, Nicholas Bloom, and Steven J. Davis. 2023. "The Demographics of Work from Home." Unpublished.

Carlino, Gerald, and William R. Kerr. 2015. "Agglomeration and Innovation." In *Handbook of Regional and Urban Economics,* Vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange. Amsterdam: North-Holland.

Chen, Chinchih, Carl Benedikt Frey, and Giorgio Presidente. 2022. "Disrupting Science." Oxford Martin Working Paper Series on Technological and Economic Change 2022-4.

Choudhury, Prithwiraj, Cirrus Foroughi, and Barbara Zepp Larson. 2021. "Work-from-Anywhere: The Productivity Effects of Geographic Flexibility." *Strategic Management Journal* 42 (4): 655–83.

Choudhury, Prithwiraj, Tarun Khanna, Christos A. Makridis, and Kyle Schirmann. 2022. "Is Hybrid Work the Best of Both Worlds? Evidence from a Field Experiment." Harvard Business School Working Paper 22-063.

Combes, Pierre-Philippe, and Laurent Gobillon. 2015. "The Empirics of Agglomeration Economies." In *Handbook of Regional and Urban Economics*, Vol. 5, edited by Gilles Duraton, J. Vernon Henderson, and William C. Strange, 247–348. Amsterdam: North-Holland.

Davis, Steven J. 2022. "The Big Shift to Working from Home." Slides prepared for the NBER Macroeconomics Annual Conference, April 1. https://static1.squarespace.com/static/5e2ea3a8097ed30c779bd707/t/62474133a4e30576426c57c3/1648836921090/Big+Shift%2C+NBER%2C+1+April.pdf.

Dingel, Jonathan, and Brent Neiman. 2020. "How Many Jobs Can Be Done at Home." *Journal of Public Economics* 189: 104325.

Emanuel, Natalia, and Emma Harrington. 2023. "Working Remotely? Selection, Treatment, and the Market for Remote Work." Federal Reserve Bank of New York Staff Report 1061.

Emanuel, Natalia, Emma Harrington, and Amanda Pallais. 2023. "The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today?" Unpublished.

Fisher, Kimberly, Jonathan Gershuny, Sarah M. Flood, Joan Garcia Roman, and Sandra L. Hofferth. 2018. "American Heritage Time Use Study Extract Builder: Version 1.2." Minneapolis, MN: IPUMS. https://doi.org/10.18128/D061.V1.2 (accessed on August 15, 2023).

Flood, Sarah M., Liana C. Sayer, Daniel Backman, and Annie Chen. 2023. "American Time Use Survey Data Extract Builder: Version 3.2." College Park, MD: University of Maryland and Minneapolis, MN: IPUMS. https://doi.org/10.18128/D060.V3.2 (accessed on August 3, 2023)

Forman, Chris, and Nicolas van Zeebroeck. 2012. "From Wires to Partners: How the Internet Has Fostered R&D Collaborations within Firms." *Management Science* 58 (8): 1549–68.

Forman, Chris, and Nicolas van Zeebroeck. 2019. "Digital Technology Adoption and Knowledge Flows within Firms: Can the Internet Overcome Geographic and Technological Distance." *Research Policy* 48 (8): 103687.

Gibbs, Michael, Friederike Mengel, and Christoph Siemroth. 2023. "Work from Home and Productivity: Evidence from Personnel and Analytics Data on Information Technology Professionals." *Journal of Political Economy Microeconomics* 1 (1): 7–41.

Hansen, Stephen, Peter J. Lambert, Nick Bloom, Steven J. Davis, Raffaella Sadun, and Bledi Taska. 2023. "Remote Work across Jobs, Companies, and Space." NBER Working Paper 31007.

Kastle Security Systems. n.d. "Getting America Back to Work." www.kastle.com/safety-wellness/getting-america-back-to-work/-workplace-barometer (accessed on August 31, 2023).

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading Off Commute against Wage." *Quarterly Journal of Economics* 136 (1): 381–426.

Lin, Yiling, Carl Benedikt Frey, and Lingfei Wu. 2023. "Remote Collaboration Fuses Fewer Breakthrough Ideas." arXiv: 2206.01878.

Mas, Alexandre, and Amanda Pallais. 2017. "Valuing Alternative Work Arrangements." *American Economic Review* 107 (12): 3722–59.

Monte, Ferdinando, Charly Porcher, and Esteban Rossi-Hansberg. 2023. "Remote Work and City Structure." NBER Working Paper 31494.

Parker, Kim. 2023. "About a Third of U.S. Workers Who Can Work from Home Now Do So All the Time." *Pew Research Center*, March 30. https://pewrsr.ch/3nwMOap.

Pew Research Center. 2021. "Internet/Broadband Fact Sheet." https://www.pewresearch.org/internet/fact-sheet/internet-broadband/.

Ramani, Arjun, and Nicholas Bloom. 2022. "The Donut Effect of Covid-19 on Cities." NBER Working Paper 28876.

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." In *Handbook of Labor Economics*, Vol. 1, edited by Orley C. Ashenfelter and Richard Layard, 641–92. Amsterdam: North-Holland.

Schmookler, Jacob. 1966. *Invention and Economic Growth*. Cambridge, MA: Harvard University Press.

Scur, Daniela, Raffaella Sadun, John Van Reenen, Renata Lemos, and Nicholas Bloom. 2021. "The World Management Survey at 18: Lessons and the Way Forward." *Oxford Review of Economic Policy* 37 (2): 231–58.

US Bureau of Economic Analysis. 2023. "Personal Consumption Expenditures: Chain-Type Price Index [PCEPI]." FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/PCEPI (accessed August 17, 2023).

US Census Bureau. 2022–2023. "Household Pulse Survey Public Use Files (PUFs): Week 46 (June 1 to June 13, 2022) to Week 58 (June 7 to June 19, 2023)." https://www.census.gov/programs-surveys/household-pulse-survey/datasets.html (accessed on June 29, 2023).

Yang, Longqi, David Holtz, Sonia Jaffee, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce et al. 2022. "The Effects of Remote Work on Collaboration among Information Workers." *Nature Human Behavior* 6: 43–54.

COVID-19, School Closures, and Outcomes

Rebecca Jack and Emily Oster

One of the earliest and clearest facts that emerged at the start of the COVID-19 pandemic was that children were among the groups least affected in terms of health outcomes. Despite this, the pandemic imposed enormous costs on children: loss of parents and caregivers, illness, disruptions to daily routine, and loss of opportunities for social contact. Among the most significant disruptions, and the focus of this paper, was the closure of schools.

In March 2020, when the pandemic began, schools closed in nearly every country worldwide. Many expected schools to reopen after a few weeks, but it quickly became clear that school closures were going to extend through the end of the school year. Before the following 2020–2021 school year began, school and education leaders had to decide whether to reopen for in-person education or continue with virtual schooling (or adopt a hybrid approach). At this stage of the pandemic, there was relatively little data to rely on. Because of the uncertainty and various political climates, reopening was slow and partial. In some locations, schools stayed closed for months or years. Students learned at home with support from family, curriculum materials, online collaboration tools like Zoom, or even through the radio. But these methods worked poorly, or worse, for a number of students.

In the end, a number of commentators have pointed to prolonged school closures, and their attendant effects on children, as one of the greatest negative impacts of the pandemic (Leonhardt 2022; *The Economist* 2022). This article details what we know about the patterns of closures and their effects.

■ *Rebecca Jack is a PhD student in Economics, University of Nebraska-Lincoln, Lincoln, Nebraska. Emily Oster is Professor of Economics, Brown University, Providence, Rhode Island. Their email addresses are bjack2@unl.edu and emily_oster@brown.edu.*

First, we begin with the question of data and our perspective on the importance of public, accessible, and contemporaneous data in the face of a public crisis. Then, we present data on the extent of school closures, both globally and within the United States. We describe the available data on the degree of these closures, which will provide a set of resources for studying longer-term consequences as they emerge. We also highlight what we know about the demographic patterns of school closures.

Second, we discuss the emerging estimates of the short-term impacts of school closures. This section will highlight possible identification strategies for estimation of the effects, but the analytical task of identifying effects of closures from other economic and health consequences of the pandemic is challenging. A central finding throughout our discussion is that school closures during the pandemic tended to increase inequality, both within and across countries. School closures were more common in less developed countries and, within the United States, more common for lower-income students. Estimates of short-term impacts were larger for these groups, too.

We also emphasize that fully understanding the long-run impact of COVID-related school closures on students will take time—by definition, long-term consequences can only be measured in the long run—and will surely be influenced by events and policies in the next few years. In particular, estimates of early-grade disruptions on high school graduation rates, college-going, and labor market outcomes are a decade or more away.

Data during a Public Crisis

In March 2020, virtually every school in the United States (and worldwide) closed for in-person instruction. In the United States, schools remained closed through the end of the 2019–2020 school year. After these initial school closures, schools had only the short summer months to decide whether or not to return to in-person school in late summer or fall of 2020.

At this stage of the pandemic, the evidence on how (and whether) to safely reopen schools was limited. Several European countries reopened their schools during the summer of 2020, but it was unclear whether the reassuring lessons learned there would translate to the United States. With relatively little data-based guidance, US school district reopening decisions were varied. Some schools chose to remain virtual while others opened part- or full-time with various mitigation efforts, including masking, social distancing by three or six feet, and extensive cleaning or air purification measures.

As some schools began opening in the late summer of 2020, it became clear that there was an opportunity to learn from the schools that did reopen, which in turn could inform those schools which had not yet made reopening decisions. Unfortunately, there was no coordinated effort by any federal agency to assemble data on school reopening or on COVID cases in schools. We embarked on an effort to collect these data, which we briefly describe below.

Initial COVID-19 Case Data

We began collecting COVID-19 case data in fall 2020. We collected data on student and staff cases from any school district that volunteered their data, and from states that were collecting and reporting district level data (New York, Texas, and Massachusetts). The dataset was eventually expanded to cover 30 states, and as many as 12.9 million students, approximately 24 percent of the US student population (COVID-19 School Response Dashboard 2021).

In the midst of a public health crisis, it was important to us that these data be made available quickly. We published data on a biweekly basis, showing COVID case rates in schools as they evolved over time. These data allowed us, initially, to show that COVID cases in school populations were limited and were reflective of case rates in the overall population.

Later in the 2020–2021 school year, we were able to leverage these data to influence policy. Collecting this data and working with data partners from Massachusetts allowed us to show that there were no statistical differences in caseloads between schools who implemented three feet of distancing compared to schools who implemented six feet of distancing, ultimately contributing to the decision by the Centers for Disease Control to abandon these distancing requirements (van den Berg et al. 2021).

School Learning Models

As we collected information on COVID cases, we also collected information on school reopening modes. This allowed us to generate some early evidence on disparities in student experience. For example, Black and Hispanic students were more likely to be in districts that only provided virtual schooling compared to white students, who were more likely to have the option of attending in person (Oster et al. 2021). These initial disparities made it clear that, in order to eventually understand the long-term consequences of what students experienced during COVID, it would be important to have universal data on school closures and school learning modes during the 2020–2021 school year.

Real-time data on school closures was somewhat haphazard during the 2020–2021 school year. Similar to the COVID data, this information was not tracked in a systematic way by any federal agency.

The most complete data available in real time came from private companies. The Burbio K–12 School Opening Tracker (Burbio 2021) covered up to 1,200 (of 13,000) school districts and approximately 47 percent of the US public school enrollment. Another private company, SafeGraph, made data available on cell phone traffic to public schools, allowing for a measure to approximate school closures by looking for drops in typical cell phone traffic (Parolin and Lee 2021). Following the end of the 2020–2021 school year, better data became available through state-level education agencies. These agencies often had to collect these data for reporting to the Pandemic Electronic Benefits Transfer (P-EBT) program, in which schoolchildren from qualifying lower-income families were given a preloaded card that allowed them to access nutritional benefits. However, these state-level data were generally not published or accessible without contacting each state-level agency individually.

We were involved in creating the COVID-19 School Data Hub, financed by several foundations, which provides data on opening patterns by school district, from official state agency records, for 48 states and the District of Columbia. Data are available at the district- and school-level (CSDH 2023a, b). The data have been used in a number of analyses (for example, Jack et al. 2023; Cohodes and Pitts 2022; Heise 2023; Singer et al. 2022; Weber and Baker 2023).

In this paper, we use these data in the analysis of the impacts of school closures on test scores. Long term, they may be useful for thinking about the further effects of school closures during the COVID year.

Important Lessons from Data Collection during COVID

Based on our experience collecting data both on COVID cases and learning model data, we highlight three lessons for collecting data during times of crisis and, more broadly, for collecting data intended to affect policy.

First, when possible, make aggregated data public, not proprietary: Public data allows for real-time feedback and correction. Making data public quickly also allows media, policymakers, and researchers to use the data early on, which led to faster policy change during an important time of upheaval.

Second, there is a need for systems to collate data that are already being collected for other purposes: Much of our data in the COVID-19 School Data Hub came from the data states collected for the Pandemic Electronic Benefits Transfer (P-EBT) program. While states had comprehensive data on district policies, without collating this data, it was unavailable for researchers to use to understand the impacts of the pandemic.

Third, when collecting data, it is important to keep both short-run and long-run in mind: During initial shut-downs, it was important to understand how COVID spread in schools in order to determine what the best policies were regarding schools. For long-term analysis, it is most important to know what students experienced. Keeping both of these perspectives in mind influenced our decision to collect learning model data along with caseload data.

Patterns and Experiences of School Closures

Worldwide

Virtually every country in the world closed its schools in March 2020, when the full force of the pandemic hit. According to the Oxford COVID-19 Government Response Tracker, a global database of pandemic policies, only three countries did not require any school closures in 2020: Burundi, Tajikistan, and Turkmenistan (Hale et al. 2021).

During this initial period of school closures in spring 2020, countries used a variety of approaches to continue education. TV programming and government-supported online platforms were the most widely used formats of education delivery (78 and 74 percent of countries, respectively), but the majority of countries also

Figure 1
Education Delivery Methods by Region, March–May 2020
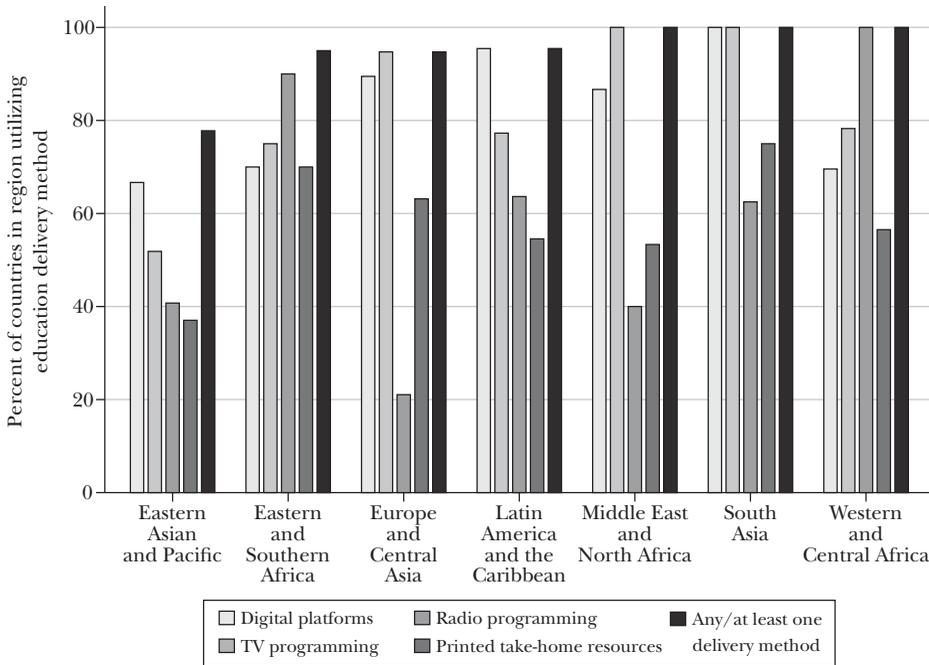
Note: We present regional data from UNICEF showing the percent of countries in each region utilizing selected education delivery methods. "Any/ at least one delivery method" is calculated using UNICEF data on the percent of countries in each region utilizing none of the education delivery methods in the UNICEF dataset.

used radio programming, printed resources, and mobile and social media methods. Many of these methods provided students resources for self-learning, rather than synchronous learning with a teacher (UNICEF 2021).

Figure 1 illustrates the cross-country variation in technologies used in the absence of in-person school (UNICEF 2021). While digital platforms, including both government-run platforms and others, were commonly used in Europe and South Asia, in other regions (Africa, in particular) radio was more common. Some regions, including the Middle East and North Africa, South Asia, and Western and Central Africa, achieved universal coverage with at least one type of education delivery method, while in other regions, including the Eastern Asia and Pacific region, not all countries provided these resources during the initial shutdowns.

Schools around the world gradually reopened starting later in 2020, depending on a combination of school calendars and COVID restrictions. Munro et al. (2023) document the full range of school experiences using data from the Oxford COVID-19 Policy Tracker. The coding of these openings and closings at the national level is

challenging, because in many cases there is within-country variation. With this caveat, the data make clear there was considerable variation, even within countries with comparable income levels. Some countries (for example, Iceland) reopened as early as May 2020, with some alterations. Others (for example, Germany) were not fully open until mid-2021. At the tail end is Saudi Arabia, with school closures lasting well into 2021.

We summarize these data by region in Figure 2. North America had the highest average number of days of school closed at 535 days in the two years between January 2020–December 2021, which is around 73 percent of the two-year period, while sub-Saharan Africa had the least with fewer than 300 days closed, or around 38 percent of the two-year period. Thus, students in countries with the longest closures spent more than twice the amount of time without traditional in-person schooling compared to students in countries with the shortest closures.

United States

Relative to the global picture, the United States was on the upper end in terms of length of school closures. The database in Munro et al. (2023) estimates that schools in the United States were either closed or partially closed for around 667 days over the two-year period from January 2020 to December 2021, inclusive of typical school and summer breaks.

This overall number masks enormous variation across the United States. Virtually all schools closed in March 2020. When it came time to decide about school reopening in the fall 2020, while there was state-level pressure on both sides, the ultimate decision about whether to open and in what way was made by each of the roughly 13,000 school districts individually.

School districts chose one of three primary paths, although with significant variations within each approach. First, some school districts opened for full-time, in-person schooling for all students at the start of the 2020–2021 school year (which varies in the United States from late July to early September). These districts engaged in varying levels of COVID mitigation, including additional ventilation and masking, surveillance testing, and quarantine procedures. Second, some school districts adopted hybrid schooling models, where students came to school in-person some of the time, but either not all days of the week or not all hours in a day. These models had a huge range—from partial days, to partial weeks, to complicated schedules like eight days on, eight days off. In many cases, these hybrid options were created to ensure that the social distancing recommended by the Centers for Disease Control could be maintained in school (at least in principle). Finally, some school districts opened in a fully virtual mode. In these cases, students often "attended" via some form of distance learning, like Zoom. An important note is that during the school year virtually *all* school districts, regardless of their opening mode, offered fully virtual learning options to students, either within the district or in state-led virtual academies.

During the 2020–2021 school year, schools moved between modes. Some schools which had opened in the early fall closed for short periods around the winter surge. More importantly for overall trends, in the later winter, after vaccines

Figure 2

Average Number of Days with School Closed, by Region, January 2020–December 2021
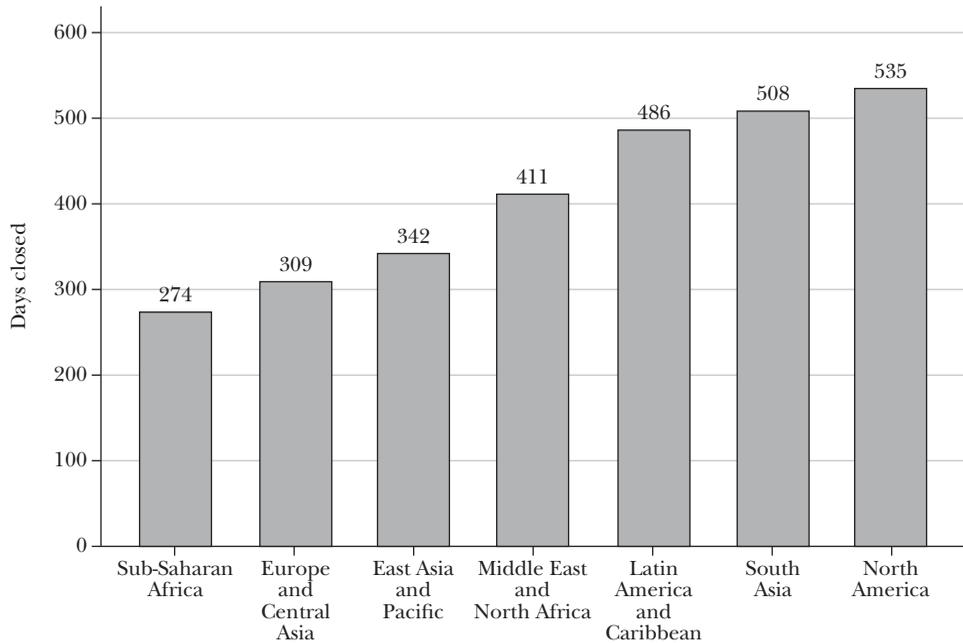
Source: Regional closures aggregated from country-level closures using data from the Oxford COVID-19 Policy Tracker (Hale et al. 2021). A closure is defined as a Tracker score of 2 or 3 following the method used for Figure 2 in Munro et al. (2023). Data for regional categorization of countries from World Bank (2023).
*Note:* For each country, we find the total number of days with an Oxford COVID-19 Policy Tracker score of 2 or 3 from January 2020–December 2021. These totals include standard school closures such as summer vacations and holidays. We then calculate the mean number of days closed by region as defined by the World Bank.

became available, there was a stronger push for school reopening. To get a general sense of reopening, in September 2020, 32 percent of the districts in the COVID School Data Hub spent a majority of the month in virtual schooling and 40 percent of districts spent a majority of the month with in-person schooling available. By the end of the school year, less than 5 percent of districts had a majority of May with only virtual schooling, and 59 percent of districts had a majority of May with in-person schooling available.

There is significant variation in opening by demographics, both across and within states. Figure 3 shows correlations between various district and county characteristics and the average number of days spent in virtual school, weighted by district enrollment, and calculated based on the US average of 180 days per school year. We include data on the following characteristics: the share of the school population

Figure 3

Average Days in Virtual School by Local Characteristics during the 2020–2021 School Year
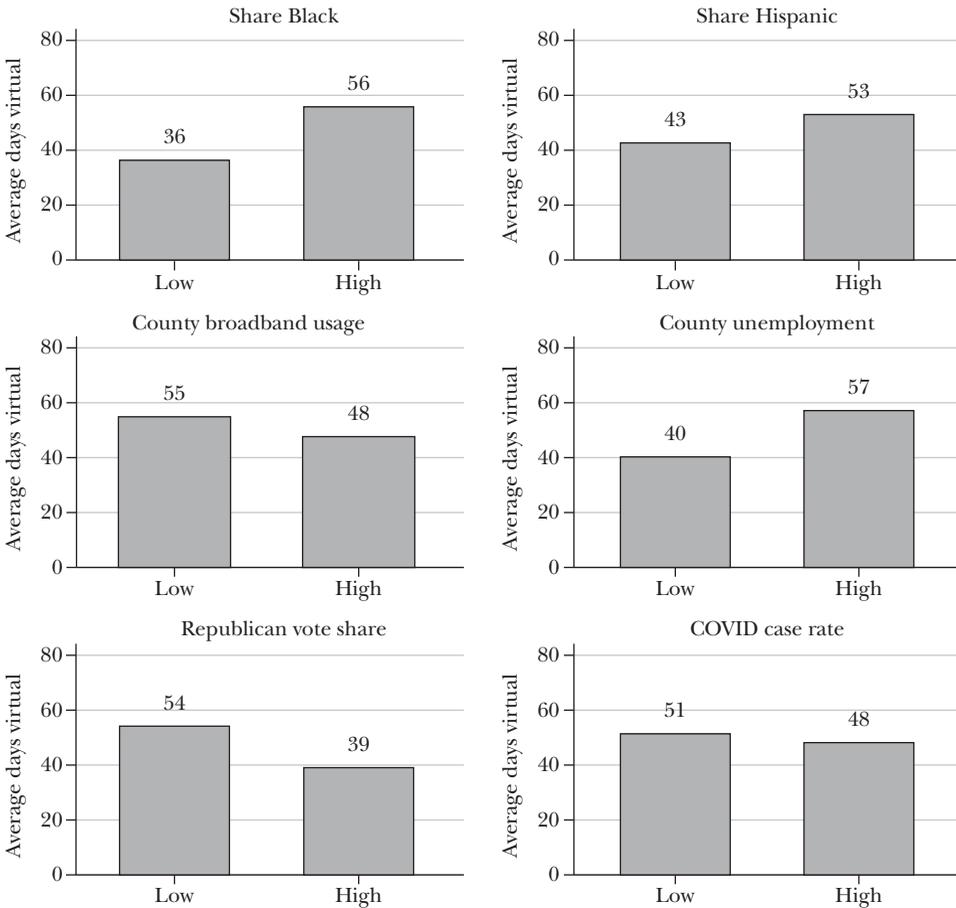
Source: Demographic variables are from the US Department of Education's Common Core of Data (US Department of Education 2022) for the most recent year available and include students who are Black or Hispanic (2020–2021 school year). County broadband usage data come from the US Broadband Usage Percentages Dataset (Kahan 2020); county unemployment rate data come from the US Bureau of Labor Statistics (2021); Republican vote share data come from the United States General Election Presidential Results by County Dataset (McGovern 2009–2020); and COVID-19 case rate data come from the COVID-19 School Data Hub (CSDH 2021).
Note: For each demographic characteristic, we calculate the median for each state in our data, and then assign "high" to school districts with levels above their state median and "low" to school districts with levels below their state median. For each high and low group, we calculate the mean share of the school year spent in virtual schooling, weighting by school district enrollment, and then use 180 days, the United States average number of school days, to calculate the average number of days spent in virtual schooling.

that is Black or Hispanic during the 2020–2021 school year; county broadband usage rates; county-level unemployment rates; the Republican vote share in the 2020 presidential election; and county-level data on the COVID-19 case rate. We assign high and low categories for each characteristic by state, assigning "high" to school districts with levels above the state median for each characteristic and "low" to school districts with levels below the state median. This assignment allows the results to be driven by variation within states, rather than across states. Of course, these are descriptive correlations, not causal arguments.

Figure 3 shows that, within state, school districts with higher shares of Black students and Hispanic students spent more of the 2020–2021 school year with only virtual schooling compared to districts with lower shares of Black students and Hispanic students. In addition, school districts located in counties with lower broadband usage and school districts located in counties with higher unemployment rates had more of the 2020–2021 school year in virtual schooling. Virtual schooling was also correlated with the political leanings of an area, with more Republican-leaning areas having fewer days of virtual school on average. Days of schooling mode had a very low correlation with county-level COVID case rates.

Taken together, access to in-person education was unequal across demographic and county characteristics, with more disadvantaged districts spending more of the school year in virtual schooling. These results echo existing work done with other reopening data (Hodgman et al. 2021; Oster et al. 2021).

The COVID-19 School Data Hub focuses on US public schools, which are attended by approximately 91 percent of children in the United States (Irwin et al. 2022). Although the data for private schools are less systematic, a parent survey from November 2020 found that private school students were more likely to have access to in-person instruction as compared to their public school peers (Henderson, Peterson, and West 2020). This pattern likely reinforces the income inequality in school opening, as private schools generally serve a population which comes from higher-income families (Murnane et al. 2018).

It would be interesting to compare specific methods of virtual learning. Reports during 2020–2021 indicate that district approaches to virtual learning varied widely and included paper packets sent home with students, Zoom sessions with teachers and classmates, and tools provided by means of various digital platforms (Kamenetz 2020). However, there is a little or no systematic data on the precise approaches districts took. Similarly, hybrid learning environments varied tremendously and there is limited data to be precise about the variation.

Consequences of School Closures

Many of the questions about the effects of school closures are only likely to be answered as consequences appear in the long term. Thus, we focus here on the emerging short-term evidence. An analytical challenge arises in estimating causal impacts of school closures. Generally, papers in this small but growing literature

have used variation in school closures across areas within a country, region, or state and related them to changes in student outcomes. Although this approach is naturally subject to concerns that there may be a reason places had varying closures, appropriately chosen comparison groups and pre-trend adjustments may be able to address these concerns. We focus primarily here on test scores, which have been the best studied outcome, and then briefly turn to non-test-score outcomes like enrollment and mental health.

Prepandemic Literature on School Closures and Learning Loss

Before the pandemic, there was a modestly sized economic literature on how school closures in the past affected learning and earnings, with data typically drawn from short-term episodes. For example, Jaume and Willén (2019) show that teacher strikes in Argentina, which closed schools for an average of 88 days across provinces, reduced long-term labor earnings of males by 3.2 percent and females by 1.9 percent.

Several studies look at the aftermath of natural disasters that disrupted schooling. Andrabi, Daniels, and Das (2020) look at the aftermath of an earthquake in Pakistan in 2005 and find that students in regions affected by the quake lost 1.5–2 years of schooling compared to students in other regions—and a substantial part of that loss happened after the earthquake, when students who had fallen behind had a hard time catching up. In a study of the long-term academic outcomes of students evacuated from their neighborhood and forced to switch schools by Hurricanes Katrina and Rita, Sacerdote (2012) found that while such students experienced test-score losses in the first year after the hurricanes, they began to experience academic *gains* by the third and fourth year, with the gains concentrated among students who had previously been the lowest-performing. Other papers considering the aftereffects of hurricanes on education include Pane et al. (2008) on the Louisiana hurricanes of 2005 and Lamb, Gross, and Lewis (2013) on the effects of Hurricane Katrina in Mississippi. Harmey and Moss (2021) consider a group of 15 papers, including hurricanes, earthquakes in New Zealand and Japan, and a school district in rural Illinois that was closed for eight days in preparation for flooding. Along with evidence of learning loss, these authors emphasize the challenges of adapting teaching and curriculum for the disrupted students.

It has also been widely observed that student time out of school in the summer causes learning loss, and more so for low-income students. For example, McCombs et al. (2011) argue that the average student loses the equivalent of one month of education during summer break, but with a wide distribution around that average. Alexander, Pitcock, and Boulay (2016) edit a collection of essays on this literature, and von Hippel, Workman, and Downey (2018) emphasize some difficulties of measurement (for example, the tests taken during fall term of one grade are often not the same as those taken spring of the preceding grade).

This literature is certainly suggestive of possible effects of the COVID-19 school closures on learning. However, the pandemic school disruptions went far beyond what had been experienced in prior episodes. The scope for inference about the pandemic period from this earlier work is limited.

Test Scores during the COVID Pandemic: Losses

Test score impacts have been most studied in the US context, likely because of the longer period of school closures and the widely available test score data. Comparisons from before and during the pandemic can be useful for understanding the size of the learning loss problem, but viewed strictly from the standpoint of the causal effects of school closings, such comparisons will inevitably mix together the effects of school closures with other aspects of the pandemic like economic and health disruptions for many households, social isolation, mental stress, and so on. When looking at variation across whether schools were open, closed, or in a hybrid arrangement during the pandemic, the existence of the pandemic is in some sense held constant across these groups. However, it remains important to consider whether the school districts that remained open are systematically different in other ways that might affect education from those that closed.

Overall, US students experienced an historic decline in test scores during the pandemic period. The National Assessment of Educational Progress (NAEP) is given to a nationally representative sample of fourth-, eighth-, and twelfth-grade students every few years. It is sometimes called the "Nation's Report Card," because it offers a unified way of looking at student progress over time. Nationally, NAEP test-score declines from 2019 to 2022 were most significant in math, but also evident in reading (NCES 2022a, b), with declines large enough overturn the gradual but clear progress of the last two to three decades. Testing by the private research organization Northwest Evaluation Association (NWEA), which compares spring 2021 scores to prepandemic outcomes in spring 2019 (using its "MAP growth assessments"), show similar overall declines (Lewis et al. 2021).

At least two papers have specifically looked at the impact of remote schooling on test-score declines. Our work in Jack et al. (2023) uses state-level assessment data from twelve states that have district-level data available (Colorado, Connecticut, Florida, Massachusetts, Minnesota, Nevada, Ohio, Rhode Island, Virginia, West Virginia, Wisconsin, and Wyoming) to estimate the effect of schooling mode, measured at the district level, on test scores. Our findings show that within even small geographic areas—like a county or commuting zone—a greater degree of remote schooling during the 2020–2021 school year was associated with a larger decline in student test scores between spring 2019 and spring 2021. Our estimates suggest that moving a district from fully in-person to fully remote learning would predict a 13 percentage points larger decline in student pass rates in math and an 8 percentage points larger decline in English/Language Arts. Moving to fully hybrid learning for the school year had about half of that effect in the data.

Goldhaber et al. (forthcoming) use individual-level test score data from the NWEA for 2.1 million students in 10,000 schools in 49 states to relate student academic *growth* to the schooling mode available. They find that remote and hybrid learning environments led to lower student achievement growth, while, for example, the dispersion in math scores remained much the same during the pandemic in schools that mostly remained in-person.

Figure 4

Average Change in Pass Rates on State Standardized Assessments in Spring 2021 versus Spring 2018–2019
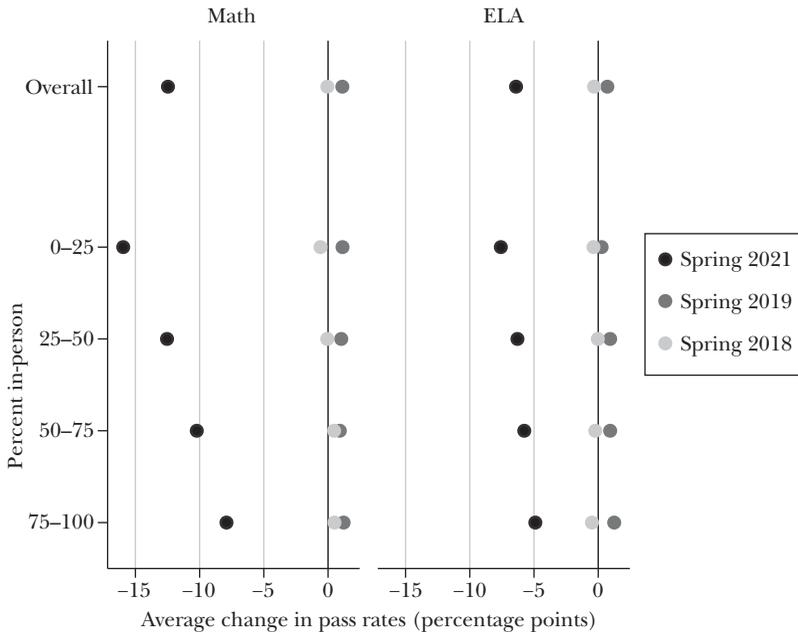
Source: Data for this figure comes from the final data set from the replication materials for Jack et al. (2023).
Note: The above figure uses district-level pass rate data from twelve state standardized assessments, which are typically taken yearly in the Spring. Standardized assessments were canceled in 2020 due to the COVID-19 pandemic. We calculate the enrollment-weighted mean change in pass rates year-to-year overall. We also use schooling mode data from the COVID-19 School Data Hub to calculate the district-level share of the 2020–2021 school year with in-person learning available to calculate the average change in pass rates based on the amount of time spent in-person.

Importantly, both of these papers show larger negatives of remote learning for less advantaged schools or students. Goldhaber et al. (forthcoming), for example, demonstrate a much larger impact on math scores for students in high- or mid-poverty districts. In both cases, the equity impacts are twofold. First, schools with lower resources (and students with fewer resources) are less likely to have access to in-person schooling, as noted in the previous section. Second, the *consequences* of that lack of in-person schooling are larger. This may be due to fewer resources invested by the schools in such students with remote learning, or to greater challenges that such students face with at-home learning environments.

Moving to the worldwide picture, there have been efforts in a number of countries using to study the impact of spring 2020 school closures on learning. In general, these papers either exploit variation across regions or simply look at how the test data from during the pandemic compares to prior years. Generally, studies

of both types show significant test-score losses. For example, using data for the Netherlands, Engzell, Frey, and Verhagen (2021) found that elementary students experienced learning losses equivalent to approximately one-fifth of a school year following eight weeks of school closures in spring 2020.

Using data for Italy, Contini et al. (2021) compare the pre-COVID and COVID cohorts of second-grade students, and find the losses are largest for high-performing children with less-educated parents. Looking at students in the last year of primary school in the Dutch-speaking Flemish region of Belgium, Maldonado and De Witte (2021) find sizeable learning losses, again largest for more disadvantaged students. With data on fifth-graders in Baden-Württemberg, Germany, Schult et al. (2022) find considerable learning loss, and again particularly for low-achieving students. In a number of German-speaking cantons in Switzerland, the schools use a computer-based system that provides feedback in grades three through nine, which allows Tomasik, Helbling, and Moser (2021) to look at what students were learning in the eight weeks of school before the pandemic compared to the eight weeks that schools moved to virtual learning. They found: "In the 8 weeks before the school closures, learning in primary schools took place rather uniformly and with hardly observable differences between single pupils, but during the school closures, inter-individual differences skyrocketed." Their dataset did not have socioeconomic data for students, but other surveys suggested that students from higher-income families were more likely to receive tutoring when the schools were closed.

Outside high-income countries and using data for South Africa—where the school year spans four terms between mid-January through mid-December—Ardington, Wills, and Kotze (2021) find that learning losses for students equaled roughly 60 percent of a year of learning for students in grades two and four. Given the very large disruptions in schooling throughout much of the developing world, it seems likely that there will be devastating long-term consequences for student learning. However, data is not yet available to test that hypothesis.

The most important implication of these facts is the COVID-19 school closures most negatively affected the students who were most vulnerable. The result was a widening of the already large inequality gap.

Test Scores: Recovery

To help support schools respond to and recover from COVID-19, policymakers established the Elementary and Secondary School Emergency Relief (ESSER) Fund in 2020 through significant investments that were part of broader federal relief packages: the Coronavirus Aid, Relief, and Economic Security Act , known as the CARES Act, signed into law March 27, 2020; the Coronavirus Response and Relief Supplemental Appropriations Act of 2021, known as CRRSAA, signed into law on December 27, 2020; and the American Rescue Plan Elementary and Secondary School Emergency Relief, known as ARP ESSER, signed into law on March 11, 2021 (Skinner, Fountain, and Dortch 2023).

The first two funding allocations to the ESSER Fund from CARES and CRRSAA (ESSER I and ESSER II) totaled $13 billion and $54 billion, respectively. ESSER I

and ESSER II allowed school districts to have flexibility in how funds were allocated, but generally targeted COVID-19 preparedness and response initiatives. The third federal relief package, however, designated over $122 billion dollars to the ESSER Fund as part of the American Rescue Plan (ARP ESSER). These funds again could support an array of initiatives, but had two important requirements: Districts needed to (a) develop a plan for students to safely return to in-person instruction in schools, and (b) allocate at least 20 percent of their funding specifically to address learning loss in an effort to support student recovery.

The results of these ESSER investments, and the extent of the hoped-for recovery of test scores in general, remains unclear as of summer 2023. In Halloran et al. (2023), we use state-level test score data through spring 2022 to illustrate large variation in the extent of test score recovery across 21 states over the 2021–2022 school year. Kuhfeld and Lewis (2022) explored learning loss and recovery at a national level from the 2021–2022 school year based on NWEA data and found that, overall, student achievement continued to lag relative to a typical year and that declines were greater in math compared to reading. In comparing spring 2019 to spring 2022 outcomes across districts on state assessments, Fahle et al. (2022) found that test-score declines were greater among districts with more remote learning during 2020–2021, but that this was not the main factor and that substantial variation was observed among districts.

Overall, the literature on recovery of test scores is still underdeveloped. Unfortunately, this is another example where limited data infrastructure may affect our ability to learn from the data. Despite the very large federal expenditures, little effort has been made to document how these funds are being spent. As a result, it may be difficult (in the short or even the long term) to point to a particular approach to recovery which has worked better.

Enrollment

Another concern with school closures and extended amounts of time without in-person education is a drop in school enrollment. A few studies have tried to look at global enrollment changes. In a literature review, Moscoviz and Evans (2022) review 14 studies that cover 13 middle- and low-income countries looking at dropout rates (along with 29 studies looking at learning loss, with some studies doing both). They find a wide variation in dropout rates during the pandemic. Among the research they survey, the highest dropout rates are from Nigeria and Brazil; for example, school enrollment in Nigeria dropped from 90 percent in 2019 to 82 percent after schools reopened in 2020. In contrast, reported changes in dropout rates in a number of African countries were only a few percent, as in Liberia, Malawi, Senegal, Sierra Leone, South Africa, and Uganda. However, even in the countries with low overall dropout rates, certain groups like those who had been about to leave school, girls, and those from households with lower incomes were often more affected. Better understanding of global school enrollment changes will be important for addressing the long-term effects of the pandemic.

In the United States, public school enrollment was affected through several channels, including movement to private schools and homeschooling, delays in enrollment for young students, and potential dropouts for older students. With schools closed, parents may have looked for other educational options for their children, including private school and homeschooling. Looking at Michigan public school enrollment, Musaddiq et al. (2022) found enrollment declines in fall 2020 that averaged 3 percent among K–12 students and 10 percent among kindergarten students compared to the prior year—roughly similar to evidence across the country. They highlighted the heterogeneous responses based on schooling mode. Specifically, students in districts with more virtual schooling switched to private schools at higher rates, while students in districts with more in-person schooling switched to homeschooling at higher rates, reflecting parental preferences.

Because of the uncertainty around schooling modes and the pandemic in general, some parents chose to delay kindergarten, which meant that a lot of the enrollment changes are concentrated in kindergarten. Combining federal data from the Common Core collected by the US Department of Education with state-level enrollment data, Dee et al. (2022) find that overall public school enrollment in the United States fell by 1.1 million students in the 2020–2021 school year. They also found that school districts offering remote-only instead of in-person instruction saw their enrollment reduced by 1.1 percentage points (from –2.6 to –3.7 percent), and these drops in enrollment were largest for kindergarten compared to other grades. Of course, kindergarten is easier to delay because it is not universally mandatory and the practice of "redshirting" (that is, starting students a year later than they were eligible to enter) is common.

We examine this correlation between drops in enrollment and schooling mode in Figure 5 by comparing fall 2020 enrollment with fall 2019 enrollment using data from the US Department of Education's Common Core of Data (US Department of Education 2022). We calculate the share of the 2020–2021 school year with in-person learning offered for each district and then divide districts into eleven groups based on this share, shown by the eleven dots in each panel. Finally, we can see the correlation between enrollment and schooling mode by comparing the percent change in enrollment for each of these groups. In panel A, there were overall enrollment declines across the board, but districts that offered more in-person schooling during the 2020–2021 school year had smaller drops in enrollment compared to districts that had more virtual schooling.

We also examine the correlation between only kindergarten enrollment and schooling mode in panel B of Figure 5 because we might expect that virtual schooling would be most difficult for young students, giving parents of kindergartners a stronger incentive to look for other schooling options. We find a stronger correlation between in-person schooling share and enrollment drops across school districts. Districts with the highest amounts of virtual schooling experienced around a 10 percent decrease in kindergarten enrollment for the 2020–2021 school year compared to the previous school year, while the districts that offered the most in-person schooling only saw a decrease of around 3 to 4 percent.
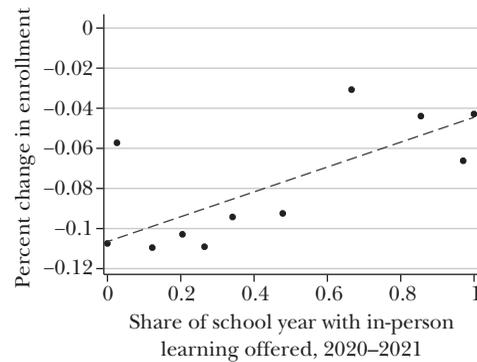
Figure 5
Percent Change in District Enrollment by Schooling Mode

Panel A. Percent change in overall district
enrollment, Fall 2019–Fall 2020
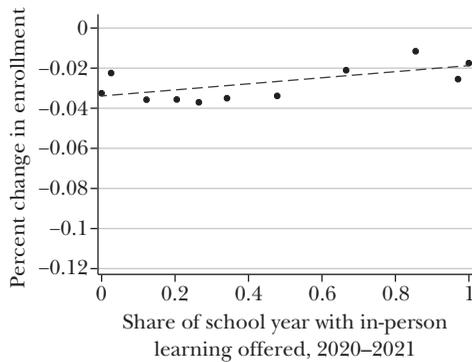
Panel B. Percent change in kindergarten
enrollment, Fall 2019–Fall 2020

Source: Schooling mode data accessed via the COVID-19 School Data Hub (CSDH 2023b). District enrollment data accessed via the US Department of Education's Common Core of Data (US Department of Education 2022) from the Urban Institute's Data Explorer.
Notes: We compare fall 2020 enrollment with fall 2019 enrollment for both school district-level overall enrollment and kindergarten enrollment. We calculate the share of the 2020–2021 school year with in-person learning offered for each district. We then group districts into eleven bins based on the share of the school year spent in-person. Finally, we calculate the enrollment-weighted mean change in enrollment for each bin to examine the correlation between the share of time spent with in-person schooling and the changes in overall enrollment and kindergarten enrollment.

Mental Health

National data has shown a rise in youth reports of poor mental health over the past decade (CDC 2020). However, the pandemic appears to have accelerated the decline in students' reported mental health, with the number of high school students reporting persistent feelings of sadness or hopelessness reaching 44 percent in 2021, an increase of over 7 percentage points in just two years (up from 36.7 percent in 2019) (Jones et al. 2022; CDC 2020), along with an increase in adolescents receiving care for eating disorders (Hartman-Munick et al. 2022).

The evidence on the extent to which student mental health is linked to school closures—as opposed to other stresses of the pandemic or rising engagement with social media—is limited and mixed. For example, school closures may have reduced school bullying. About one-fifth of high school students report being bullied in a given year, but Bacher-Hicks, Goodman, and Mulhern (2021) use data on search patterns from Google Trends to suggest that remote learning *decreased* bullying in spring 2020, with bullying patterns returning to normal as in-person schooling returned in some cases in fall 2020. In contrast, Hawrilenko et al. (2021) use nationally representative survey data from parents that used what is called the Strengths and Difficulties Questionnaire to assess their children in the categories of emotional problems, peer problems, conduct, and hyperactivity. In this data, school closures

as opposed to in-person learning were associated with reduced mental health; in particular, the costs to mental health seemed largest for high school children.

Slightly further afield, Gassman-Pines et al. (2022) collected daily data from 645 service workers paid hourly wages, who also have a young child aged two to seven, in a large US city from February to April 2020. They study a link between childcare disruptions and child behavior and parental mood. They find that childcare disruptions were more common for families with children in remote schooling, and these disruptions worsened child behavior and parental mood.

There is a lot more work to do to understand the full impacts of the pandemic on students' mental health. It is difficult to collect data on mental health and to disentangle all of the various things that impact mental health, but it remains an important area to understand as we work to understand the full impacts of school closures on children.

The Current Situation

Although it can be tempting to take short-term estimates of COVID-related learning loss or estimates from other non-COVID disruptions and project them into long-term consequences of COVID, such temptations should be resisted. Such extrapolation is potentially so distant—say, effects on long-term wages of current middle-school students—as to be ridiculous. Moreover, how the short term translates to the long term depends on what happens in the middle term. This length of school closure is unprecedented and the recovery efforts are likely to be, as well.

That said, it seems clear from the emerging evidence that the pandemic in general and school closures in particular imposed costs of learning loss, attendance, and mental health. In the medium term, there is a continued need for research and policy to engage on remediating these consequences, both globally and within the United States.

References

Alexander, Karl, Sarah Pitcock, and Matthew C. Boulay, eds. 2016. *The Summer Slide: What We Know and Can Do about Summer Learning Loss.* New York: Teachers College Press.

Andrabi, Tahir, Benjamin Daniels, and Jishnu Das. 2020. "Human Capital Accumulation and Disasters: Evidence from the Pakistan Earthquake of 2005." RISE Working Paper 20/039.

Ardington, Cally, Gabrielle Wills, and Janeli Kotze. 2021. "COVID-19 Learning Losses: Early Grade Reading in South Africa." *International Journal of Educational Development* 86: 102480.

Bacher-Hicks, Andrew, Joshua Goodman, and Christine Mulhern. 2021. "Inequality in Household Adaptation to Schooling Shocks: Covid-Induced Online Learning Engagement in Real Time." *Journal of Public Economics* 193: 104345.

Burbio. 2021. "Burbio's K–12 School Opening Tracker." https://about.burbio.com/school-opening-tracker (accessed on August 16, 2023).

Centers for Disease Control and Prevention (**CDC**)**.** 2020. *Youth Risk Behavior Survey: Data Summary and Trends Report, 2009–2019.* Atlanta, GA: CDC.

Cohodes, Sarah, and Christine Pitts. 2022. *A New COVID-19 Data Source for Answering Emerging Pandemic Questions.* Seattle, WA: Center on Reinventing Public Education.

Contini, Dalit, Maria Laura Di Tommaso, Caterina Muratori, Daniela Piazzalunga, and Lucia Schiavon. 2021. "The COVID-19 Pandemic and School Closure: Learning Loss in Mathematics in Primary Education." IZA Discussion Paper 14785.

COVID-19 School Data Hub (CSDH). 2021. "Community Case Rate Data by District." Data Resources. https://www.covidschooldatahub.com/data-resources (accessed on April 21, 2022).

COVID-19 School Data Hub (CSDH). 2023a. "District Monthly Shares (Version 3/8/2023)." Data Resources. https://www.covidschooldatahub.com/data-resources (accessed on March 13, 2023).

COVID-19 School Data Hub (CSDH). 2023b. "District-Yearly Shares." Data Resources. https://www.covidschooldatahub.com/data-resources (accessed on March 13, 2023).

COVID-19 School Response Dashboard. 2021. https://covidschooldashboard.com/ (accessed August 16, 2023).

Dee, Thomas S., Elizabeth Huffaker, Cheryl Phillips, and Eric Sagara. 2022. "The Revealed Preferences for School Reopening: Evidence from Public-School Disenrollment." *American Educational Research Journal.* https://doi.org/10.3102/00028312221140029.

Engzell, Per, Arun Frey, and Mark D. Verhagen. 2021. "Learning Loss Due to School Closures during the COVID-19 Pandemic." *Proceedings of the National Academy of Sciences* 118 (17): e2022376118.

Fahle, Erin, Thomas J. Kane, Tyler Patterson, Sean F. Reardon, and Douglas O. Staiger. 2022. "Local Achievement Impacts of the Pandemic." Education Recovery Scorecard. https://educationrecoveryscorecard.org/wp-content/uploads/2022/10/Education-Recovery-Scorecard_Key-Findings_102822.pdf.

Gassman-Pines, Anna, Elizabeth Ananat, John Fitz-Henley II, and Jane Leer. 2022. "Effects of Daily School and Care Disruptions during the COVID-19 Pandemic on Child Mental Health." NBER Working Paper 29659.

Goldhaber, Dan, Thomas J. Kane, Andrew McEachin, Emily Morton, Tyler Patterson, and Douglas O. Staiger. Forthcoming. "The Educational Consequences of Remote and Hybrid Instruction during the Pandemic." *American Economic Review: Insights**.*

Hale, Thomas, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster et al. 2021. "A Global Panel Database of Pandemic Policies (Oxford COVID-19 Government Response Tracker)." *Nature Human Behaviour* 5 (4): 529–38.

Halloran, Clare, Claire E. Hug, Rebecca Jack, and Emily Oster. 2023. "Post COVID-19 Test Score Recovery: Initial Evidence from State Testing Data." NBER Working Paper 31113.

Harmey, Sinéad, and Gemma Moss. 2021. "Learning Disruption or Learning Loss: Using Evidence from Unplanned Closures to Inform Returning to School after COVID-19." *Educational Review* 75 (4): 637–56.

Hartman-Munick, Sydney M., Jessica A. Lin, Carly E. Milliren, Paula K. Braverman, Kathryn S. Brigham, Martin M. Fisher, Neville H. Golden et al. 2022. "Association of the COVID-19 Pandemic with Adolescent and Young Adult Eating Disorder Care Volume." *JAMA Pediatrics* 176 (12): 1225–32.

Hawrilenko, Matt, Emily Kroshus, Pooja Tandon, and Dimitri Christakis. 2021. "The Association between School Closures and Child Mental Health during COVID-19." *JAMA Network Open* 4 (9): e2124092.

Heise, Michael. 2023. "The Distribution of In-Person Public K–12 Education in the Time of COVID: An Empirical Perspective." *Journal of Empirical Legal Studies* 20 (2): 305–38.

Henderson, Michael B., Paul E. Peterson, and Martin R. West. 2020. "Pandemic Parent Survey Finds Perverse Pattern: Students Are More Likely to Be Attending School in Person Where Covid Is Spreading More Rapidly." *Education Next* 21 (2): 34–48.

Hodgman, Sarah, Jordan Rickles, Joanne Carminucci, and Mike Garet. 2021. *District Approaches to Instruction in 2020–21: Differences in Instructional Modes and Instruction Time across Contexts.* Arlington, VA: American Institutes for Research.

Irwin, Véronique, Josue De La Rosa, Ke Wang, Sarah Hein, Jijun Zhang, Riley Burr, Ashley Roberts et al. 2022. *Report on the Condition of Education 2022 (NCES 2022-144)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.

Jack, Rebecca, Clare Halloran, James Okun, and Emily Oster. 2023. "Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts." *American Economic Review: Insights* 5 (2): 173–90.

Jack, Rebecca, Clare Halloran, James Okun, and Emily Oster. 2023. "Replication data for: Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E168843V1.

Jack, Rebecca, and Emily Oster. 2023. "Replication data for: COVID-19, School Closures, and Outcomes." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193523V1.

Jaume, David, and Alexander Willén. 2019. "The Long-Run Efects of Teacher Strikes: Evidence from Argentina." *Journal of Labor Economics* 37 (4): 1097–1139.

Jones, Sherry Everett, Kathleen A. Ethier, Marci Hertz, Sarah DeGue, Vi Donna Le, Jemekia Thornton, Connie Lim, Patricia J. Dittus, and Sindhura Geda. 2022. "Mental Health, Suicidality, and Connectedness among High School Students during the COVID-19 Pandemic—Adolescent Behaviors and Experiences Survey, United States, January–June 2021." *Morbidity and Mortality Weekly Report* 71 (S3): 16–21.

Kahan, John. 2020. "United States Broadband Usage Percentages Dataset." Github Repository. https://github.com/microsoft/USBroadbandUsagePercentages (accessed on March 27, 2023).

Kamenetz, Anya. 2020. "5 Things We've Learned about Virtual School in 2020." *NPR News*, December 4. https://www.npr.org/2020/12/04/938050723/5-things-weve-learned-about-virtual-school-in-2020.

Kuhfeld, Megan, and Karyn Lewis. 2022. *Student Achievement in 2021–2022: Cause for Hope and Continued Urgency*. Portland, OR: Northwest Evaluation Association.

Lamb, John, Sarah Gross, and Mark Lewis. 2013. "The Hurricane Katrina Effect on Mathematics Achievement in Mississippi." *School Science and Mathematics* 113 (2): 80–93.

Leonhardt, David. 2022. "Not Good for Learning." *New York Times*, May 5. https://www.nytimes.com/2022/05/05/briefing/school-closures-covid-learning-loss.html.

Lewis, Karyn, Megan Kuhfeld, Erik Ruzek, and Andrew McEachin. 2021. *Learning during COVID-19: Reading and Math Achievement in the 2020–21 School Year*. Portland, OR: Northwest Evaluation Association.

Maldonado, Joana Elisa, and Krsitof De Witte. 2021. "The Effect of School Closures on Standardised Student Test Outcomes." *British Educational Research Journal* 48 (1): 49–94.

McCombs, Joan Sloan, Catherine H. Augustine, Heather L. Schwartz, Susan J. Bodily, Brian McInnis, Dahlia S. Lichter, and Amanda Brown Cross. 2011. *Making Summer Count: How Summer Programs Can Boost Children's Learning*. Santa Monica, CA: RAND Corporation.

McElrath, Kevin. 2020. "Nearly 93% of Households with School-Age Children Report Some Form of Distance Learning during COVID-19." US Census Bureau, August 26. https://www.census.gov/library/stories/2020/08/schooling-during-the-covid-19-pandemic.html.

McGovern, Tony. 2009–2020. "United States General Election Presidential Results by County from 2009 to 2020." Github Repository. https://github.com/tonmcg/US_County_Level_Election_Results_08-20 (accessed on April 19, 2022).

Moscoviz, Laura, and David K. Evans. 2022. "Learning Loss and Student Dropouts during the COVID-19 Pandemic: A Review of the Evidence Two Years after Schools Shut Down." Center for Global Development Working Paper 609.

Munro, Alasdair, Danilo Buonsenso, Sebastián González-Dambrauskas, Robert C. Hughes, Sunil S. Bhopal, Pablo Vásquez-Hoyos, Muge Cevik, Maria Lucia Mesa Rubio, and Damian Roland. 2023.

"In-Person Schooling Is Essential Even during Periods of High Transmission of Covid-19." *BMJ Evidence-Based Medicine* 28 (3): 175–79.

Murnane, Richard J., Sean F. Reardon, Preeya P. Mbekeani, and Anne Lamb. 2018. "Who Goes to Private School? Long-Term Enrollment Trends by Family Income." *Education Next* 18 (4): 58–66.

Musaddiq, Tareena, Kevin Stange, Andrew Bacher-Hicks, and Joshua Goodman. 2022. "The Pandemic's Effect on Demand for Public Schools, Homeschooling, and Private Schools." *Journal of Public Economics* 212: 104710.

National Center for Education Statistics (NCES). 2022a. *2022 NAEP Mathematics Assessment: Highlighted Results at Grades 4 and 8 for the Nation, States, and Districts.* Washington, DC: NCES.

National Center for Education Statistics (NCES). 2022b. *2022 NAEP Reading Assessment: Highlighted Results at Grades 4 and 8 for the Nation, States, and Districts.* Washington, DC: NCES.

Oster, Emily, Rebecca Jack, Clare Halloran, John Schoof, Diana McLeod, Haisheng Yang, Julie Roche, and Dennis Roche. 2021. "Disparities in Learning Mode Access among K–12 Students during the COVID-19 Pandemic, by Race/Ethnicity, Geography, and Grade Level—United States, September 2020–April 2021." *Morbidity and Mortality Weekly Report* 70 (26): 953–58.

Pane, John F., Daniel F. McCaffrey, Nidhi Kalra, and Annie J. Zhou. 2008. "Effects of Student Displacement in Louisiana during the First Academic Year after the Hurricanes of 2005." *Journal of Education for Students Placed at Risk (JESPAR)* 13 (2–3): 168–211.

Parolin, Zachary, and Emma K. Lee. 2021. "Large Socio-economic, Geographic and Demographic Disparities Exist in Exposure to School Closures." *Nature Human Behaviour* 5 (4): 522–28.

Sacerdote, Bruce. 2012. "When the Saints Go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita." *American Economic Journal: Applied Economics* 4 (1): 109–35.

Schult, Johannes, Nicole Mahler, Benjamin Fauth, and Marlit A. Lindner. 2022. "Did Students Learn Less during the COVID-19 Pandemic? Reading and Mathematics Competencies before and after the First Pandemic Wave." https://doi.org/10.1080/09243453.2022.2061014.

Singer, Jeremy, Julie A. Marsh, David Menefee-Libey, Jacob Alonso, Dwuana Bradley, and Hanora Tracy. 2022. "The Politics of School Reopening during COVID-19: A Multiple Case Study of Five Urban Districts in the 2020–21 School Year." *Educational Administration Quarterly* 59 (3): 542–93.

Skinner, Rebecca R., Joselynn H. Fountain, and Cassandria Dortch. 2023. *Education Stabilization Fund Programs Funded by the CARES Act, CRRSAA, and ARPA: Background and Analysis.* Congressional Research Service (CRS) Report R47027. Washington, DC: CRS.

The Economist. 2022. "Covid Learning Loss Has Been a Global Disaster." July 7. https://www.economist.com/international/2022/07/07/covid-learning-loss-has-been-a-global-disaster.

Tomasik, Martin J., Laura A. Helbling, and Urs Moser. 2021. "Educational Gains of In-Person vs. Distance Learning in Primary and Secondary Schools: A Natural Experiment during the COVID-19 Pandemic School Closures in Switzerland." *International Journal of Psychology* 56 (4): 566–76.

UNICEF. 2021. "National Education Responses to COVID-19: UNICEF Global Tracker, May 21 (COVID-19 Page, COVID-19 Resources)." Tableau Public. https://public.tableau.com/app/profile/unicef-data/viz/EduViewv1_0/home (accessed on August 11, 2023).

US Bureau of Labor Statistics. 2021. "County-Level Unemployment Data (la.data.64.County)." US Department of Labor. https://download.bls.gov/pub/time.series/la/ (accessed on December 4, 2021).

US Department of Education. 2022. "Common Core of Data." Urban Institute Education Data Portal (Version 0.16.0). https://educationdata.urban.org/documentation/ (accessed November 16, 2022).

van den Berg, Polly, Elissa M. Schechter-Perkins, Rebecca S. Jack, Isabella Epshtein, Richard Nelson, Emily Oster, and Westyn Branch-Elliman. 2021. "Effectiveness of 3 versus 6 ft of Physical Distancing for Controlling Spread of Coronavirus Disease 2019 among Primary and Secondary Students and Staff: A Retrospective, Statewide Cohort Study." *Clinical Infectious Diseases* 73 (10): 1871–78.

von Hippel, Paul T., Joseph Workman, and Douglas B. Downey. 2018. "Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, of 'Are Schools the Great Equalizer?'" *Sociology of Education* 91 (4): 323–57.

Weber, Mark, and Bruce D. Baker. 2023. "Does School Funding Matter in a Pandemic? COVID-19 Instructional Models and School Funding Adequacy." Annenberg Institute at Brown University EdWorkingPaper 23-736.

World Bank Group (World Bank). 2023. "World Development Indicators". https://databank.worldbank.org/source/world-development-indicators (accessed on April 10, 2023).

Changes in the Distribution of Black and White Wealth since the US Civil War

Ellora Derenoncourt, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick

The vast majority of Black Americans were enslaved until the 1860s, and thus contributed to building the nation's prosperity while being barred from accumulating wealth themselves. As a result, the wealth gap between Black and white Americans at the time of Emancipation in 1863 was enormous. Recent estimates put the average wealth gap at 56:1 at the eve of the US Civil War (Derenoncourt et al. 2023a). Thus, Black Americans embarked on freedom with extremely low levels of wealth compared to white Americans.

This paper assembles the available historical evidence on the evolution of Black and white wealth since Emancipation. Despite some pioneering research on Black wealth accumulation in the post-Emancipation South by Du Bois (1901) and later by economic historians like Margo (1984) and Higgs (1982), little has been known about the long-term trajectory of the wealth gap between Black and white households. We understand now that the wealth gap between Black and white Americans narrowed in the first century after Emancipation, despite grave obstacles to Black wealth accumulation. However, Black-white wealth convergence proceeded more slowly than in a counterfactual world with equal conditions for

■ *Ellora Derenoncourt is Assistant Professor of Economics, Princeton University, Princeton, New Jersey. Chi Hyun Kim is Postdoctoral Researcher, University of Bonn, Bonn, Germany. Moritz Kuhn is Professor of Economics, University of Mannheim, Mannheim, Germany. Moritz Schularick, is President of the Kiel Institute for the World Economy, Kiel, Germany, and Professor of Economics at Sciences Po Paris, Paris France. Their email addresses are ellora.derenoncourt@princeton.edu, ckim@uni-bonn.de, mokuhn@uni-mannheim.de, and moritz.schularick@sciencespo.fr.*

the accumulation of wealth (Derenoncourt et al. 2023a). In the post-slavery period, Black Americans faced discrimination in labor and housing markets, exclusion from financial markets, and outright destruction of Black wealth that hindered continuous convergence in their wealth. On top of this, the rising importance of capital gains since the 1980s has reversed the racial wealth convergence process altogether: the average wealth gap today is larger than at the end of the civil rights era of the 1960s and 1970s.

The analysis in the paper proceeds along three dimensions. We first summarize what we have learned from research on the evolution of aggregate Black and white wealth since the Civil War and discuss the evolution of racial wealth differences based on the average wealth of both groups. We present evidence of how the per capita, or average, white-Black wealth ratio started at an extreme level of 56 to 1 on the eve of the US Civil War, fell to a ratio of 10 to 1 by 1920, and to 7 to 1 during the 1950s. However, the white-Black wealth gap then stalled and started to increase again in the last third of the twentieth century.

Of course, such averages can hide substantial heterogeneity in the distribution of wealth and therefore miss important changes in the situation of the majority of Black households. Thus, a second step of our analysis is to explore the median wealth gap, which represents the situation of a typical Black and white household; for example, the median wealth gap between white and Black households was 10 to 1 in 2020, and therefore much higher than the average wealth gap of 6 to 1 that year. However, studying the median and the distribution of wealth is even more challenging than studying the average wealth differences between Black and white Americans, due to the lack of data availability. We provide new estimates for the evolution of the share of Black and white households with (measurable) positive net wealth. The key result from this analysis is that the median Black wealth level was mostly zero during the post-Emancipation period and only turned positive between 1940 and 1950. Combining this information with microdata on wealth for the post-1950 period allows us to construct the median racial wealth gap since 1860. Strikingly, the zero-wealth level for the median Black household from 1860 to 1940 implies an infinite median wealth gap for this time relative to the median white wealth that we estimate to be positive throughout the period from 1870 to 1940.

Our third step focuses on the post-1950 period and explores the tails of the wealth distribution. The wealth distribution is highly skewed, so that wealth growth at the top can exert a large influence on the average while there is little movement for households at the bottom or in the middle of the distribution. Thus, we also look into the distribution of wealth shares and population shares over time, as well as the rank gap of the median and 90th percentile of the Black and white wealth distributions. All evidence points to the persistent and in some aspects rising over-representation of Black Americans at the bottom of the wealth distribution, as well as their declining share in total US wealth in post–World War II America. These new facts enrich the evidence on the secular evolution of the racial wealth gap and the historical wealth-building of the typical Black household.

"Landless, homeless . . . without money or tools": Black Wealth and the Racial Wealth Gap at the Time of the Civil War

Immediately before the US Civil War, nearly 4 million Black Americans out of a total Black population of 4.4 million were enslaved. These enslaved individuals were considered property and deprived of the right to own property themselves. Reconstruction-era proposals for providing land to freed persons failed, and the vast majority of the formerly enslaved embarked on freedom "landless, homeless . . . without money or tools" and in circumstances where "starvation or practical re-enslavement awaited them" (Du Bois 1901). In the first step of our discussion of racial wealth differences, we begin with the evidence from historical Census data on the distribution of Black and white wealth. We then discuss research that has worked with estimates for wealth totals and population numbers to determine the average wealth of Black Americans over time. The reason is that for the pre-1950 period, there is no consistent microdataset that would allow researchers to document the distribution of wealth by racial group after 1870.

Per capita Black wealth amounted to just $0.13 on the eve of the Civil War in 1860,[1] but the bottom 90 percent of the Black Americans, who were enslaved, had no measurable wealth at all, as illustrated in panel A of Figure 1. What wealth was owned was concentrated in a small group within the free Black population. Historians of Black property holding have documented that this group consisted primarily of planters in the Lower South, craftsmen and entrepreneurs in the Upper South, and merchants and real estate owners in the North (Berlin 1975; Walker 1983, 1986; Schweninger 1989, 1990).

The remaining millions of Black Americans toiled under lifetimes of bondage, with no recourse to earn, save, or invest the fruits of their labor. Emancipation saw some of the formerly enslaved defy all odds to purchase land, facilitated by fallen land prices in the wake of the Civil War, giving rise to an emergent property-owning class among the formerly enslaved, particularly in more urban areas of the Upper South (Schweninger 1990).[2] Yet, our estimates show that in 1870, a few years after Emancipation, over 80 percent still had no measurable wealth.

While the abolition of slavery rocked the economy of the US South to its core and eliminated the South's wealth advantage in the country (Dray, Landais, and Stantcheva 2023), the median white household was far less directly connected to the institution. There were fewer than 500,000 slaveholders in the United States on the eve of the Civil War, and slaveholding was itself highly unequally distributed (Ager, Boustan, and Eriksson 2021). Thus, the typical white southerner and all northerners had no direct wealth in the enslaved to lose as a consequence of the Civil War.
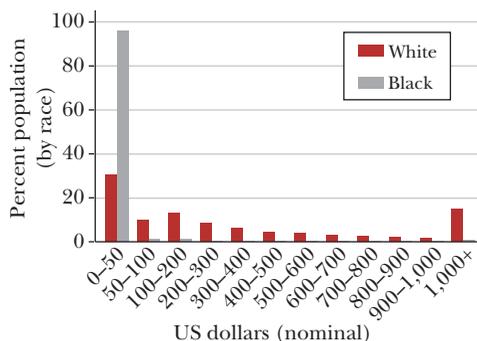
---

[1] Estimate from Derenoncourt et al. (2023a). All datasets used in this paper are available in Harvard Dataverse (Derenoncourt et al. 2023b).
[2] See Ager, Boustan, and Eriksson (2021) for estimates of the decline in southern land prices after the Civil War.
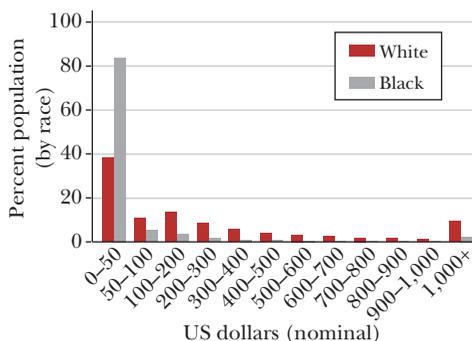
Figure 1
The Distribution of White and Black Wealth in 1860 and 1870

Panel A. Wealth distribution in 1860 by race | Panel B. Wealth distribution in 1870 by race

Source: US Census (Derenoncourt et al. 2023b).
Note: Wealth distribution of Black and white wealth in 1860 and 1870 (in current US dollars).

As shown in Figure 1, over 60 percent of white Americans owned some degree of wealth both before and after the Civil War, though an increase in wealthlessness in the white distribution is apparent between 1860 and 1870. Nevertheless, research suggests that even those slaveholders most affected by the abolition of slavery recovered their socioeconomic position within a generation (Ager, Boustan, and Eriksson 2021).

After the 1870s, it becomes difficult to trace the full distribution of Black and white wealth due to the lack of microdata. Instead, we summarize insights from our recent work on the average racial wealth gap before presenting new estimates of positive wealth holding among median Black households.

A More Comprehensive Picture of Black Wealth since Emancipation

Obtaining a more comprehensive picture of Black and white wealth since Emancipation has proven challenging. Starting in 1983, the Federal Reserve started carrying out the Survey of Consumer Finances, providing high-quality data on the financial situation of US households every three years. These data allow researchers to study the distribution of Black and white wealth over the decades after 1980.

For the period since 1950, national surveys similar to the Survey of Consumer Finances were taken as early as 1947, directed by the Economic Behavior Program of the Survey Research Center of the Institute for Social Research at the University of Michigan. The surveys were taken annually between 1947 and 1971, and then again in 1977. Kuhn, Schularick, and Steins (2020) have linked the archival survey data to the modern Survey of Consumer Finances by harmonizing and reweighting the historical data to make them as compatible as possible with the post-1983 data.

We will refer to this data as the augmented Survey of Consumer Finance, or SCF+. This novel dataset adds four decades of household-level microdata and effectively doubles the time coverage of the post-1983 Survey of Consumer Finances, without compromising on the detailed information with respect to socioeconomic and other characteristics. The historical waves come with direct measurements of assets and debt of Black and white households, as well as additional information to stratify households by other characteristics such as educational attainment, family size, and marital status.

For the pre-1950 period, there is no consistent microdataset that would allow us to document the distribution of wealth by racial group after 1870. We therefore have to follow the example of the fragmentary earlier research and rely on a variety of different sources, employ different methods, and sometimes connect different concepts of measuring inequality. We follow our work in Derenoncourt et al. (2023a) and quantify the growth of average Black wealth in the long run, relying on a variety of sources to approximate the trend. From 1870 to 1929, we extrapolate aggregate Black wealth in the 1870 census using growth rates estimated from state-level data on assessed property and tax payments. In the nineteenth and early-twentieth centuries, an intricate system of taxation for real and personal property existed in every US state, in which state auditor, treasurer, or comptroller offices regularly published reports on the finances of their state, including assessed property, taxes, and revenue collected. We used the assessed wealth of Black taxpayers in six southern states (Arkansas, Georgia, Kentucky, Louisiana, North Carolina, and Virginia), as these were the only states that reported property or taxes separately by race. Of course, the underlying assumption is that the national growth rate of Black wealth is accurately represented by the growth rate of Black wealth in the aforementioned states. After estimating total Black wealth, non-Black wealth is calculated as the difference between total national wealth (Department of Commerce 1924) and Black wealth.

For the period from 1930 to 1940, we obtained estimates of total Black wealth from Monroe Nathan Work's "The Negro Year Book." This collection of annual reports focused on the economic progress of Black individuals and covered aspects such as business, education, wealth, and social progress. Estimates of national Black wealth are available for 1930 and 1936 from these reports. The underlying method to generate these estimates is not explicitly described in these reports. However, Derenoncourt et al. (2023a) show that Work referenced Black real and personal property valuations from Georgia, North Carolina, and Virginia auditor reports, similar to our method for the 1870–1929 estimation. We combined Work's estimates with national wealth estimates from Saez and Zucman (2016) by subtracting Black wealth from total wealth and dividing the non-Black and Black wealth by their respective populations, allowing us to calculate estimates of per capita wealth.

While confidence intervals are sometimes wide, we believe that the broad trends from these calculations provide an accurate account of the evolution of racial wealth differences in the United States. Panel A of Figure 2 tracks Black

Figure 2
Historical Evolution of Black Wealth, 1870–2020

Panel A. Black wealth relative to total wealth

Panel B. Real growth of Black and white wealth
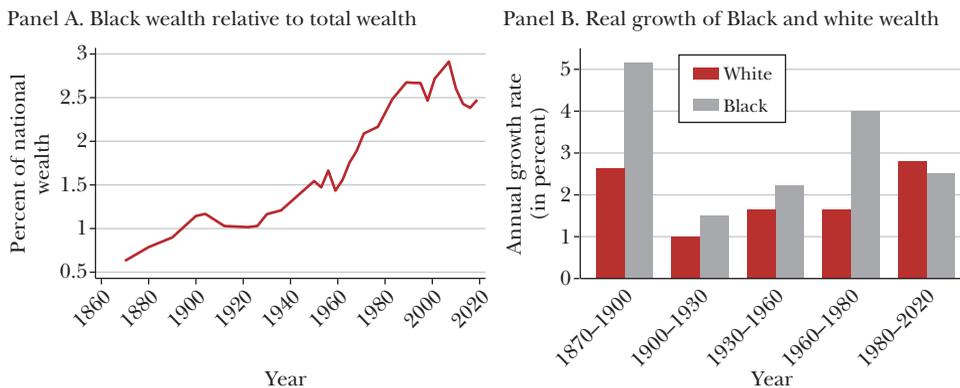
Source: Final wealth gap series are from Derenoncourt et al. (2023a). Primary data sources are 1860 and 1870 complete count censuses (Ruggles et al. 2021), Southern state auditor reports 1866–1929, Work (1932), Work (1938), Saez and Zucman (2016), and US Census Bureau's "Wealth, Public Debt, and Taxation report" (Department of Commerce 1924).
Note: Panel A presents the authors' series of the Black share of national wealth from 1860 to 2020. Black share is computed as total Black wealth as share of national wealth over time. Panel B presents the annual growth rate of Black and white wealth during 1870–2020.

wealth relative to total US household wealth. While the share of Black wealth in total wealth has also increased from less than 1 percent in the nineteenth century to around 2.5 percent today, the increase has stalled in recent decades. Despite a Black population share of close to 14 percent, the Black wealth share has remained stuck at 2.5 percent since the 1980s, ending its continuous upward trajectory since the nineteenth century.

In panel B of Figure 2, we plot the inflation-adjusted growth rates of Black wealth since Emancipation. The nominal wealth growth series are deflated by consumer price indices taken from the Jordà-Schularick-Taylor Macrohistory Database that is available for 1870–2020 (Jordà, Schularick, and Taylor 2017). The growth rate of Black wealth was the highest during the first three decades after Emancipation (1870–1900), with an annual growth rate exceeding 5 percent. Such high growth rates are not surprising: after all, the majority of Black Americans started with wealth levels close to zero, which allowed high relative increases in their wealth compared to their white counterparts. Black wealth also experienced high growth during the civil rights era from 1960–1980, when the nation experienced dramatic changes in the landscape of racial progress and discrimination. Compared to this time period, during 1900–1960 Black wealth growth was significantly slower compared to the first decades after Emancipation, yet higher than white wealth growth. After 1980, however, Black wealth growth became smaller than white wealth growth.

To put this stagnation of the Black wealth share in recent decades in perspective, it is useful to remember that the share of national wealth held by the top 0.1 percent of the population (which is predominantly white) has risen substantially in recent decades.[3] The top 0.1 percent of households held roughly 10 percent of total national wealth from 1950 into the mid-1980s, but since then, the share of the top 0.1 percent has risen to about 18 to 19 percent of total wealth (based on the SCF+ data; see also Piketty, Saez, and Zucman 2018). Hence, 0.1 percent of US households own a share of total household wealth that is an order of magnitude larger than the share of wealth that the entire US Black population owns.

Combining Wealth and Population: Average Wealth

Combining the estimates for total Black and white wealth from Derenoncourt et al. (2023a) with population data allows us to track the average wealth gap between Black and white Americans. Figure 3 plots the resulting long-run time series (red solid line), which exhibits a hockey-stick pattern. The fastest convergence of Black and white wealth occurred in the first decades after Emancipation. The wealth gap started at 56:1 at the eve of the US Civil War, fell to 10:1 by 1920, and to 7:1 during the 1950s. Wealth convergence stalled in the last third of the twentieth century, mirroring the stagnation of the Black wealth share since the 1980s, and the average wealth ratio was at 6:1 in 2019.

The pace of per capita wealth convergence was particularly fast in the early decades after Emancipation. In 1860, the average Black American owned less than $0.02 for every white dollar of wealth. The racial wealth gap dropped sharply between 1860 and 1870, the first post-Emancipation observation, with the gap falling to a level of 23:1, or a more than 50 percent decrease relative to 1860. According to US Census data, the reason was strongly growing Black wealth. Black per capita wealth tripled between 1860 and 1870, while white wealth grew by only 18 percent.

The Civil War eliminated the wealth that slaveholders held in enslaved individuals through the abolition of slavery. It also resulted in the depreciation of southern land values and afforded the formerly enslaved an opportunity to accumulate wealth for the first time. How much of the decrease in the wealth gap in the decade of the Civil War can be attributed solely to the elimination of slave wealth?[4] Using an estimate of total slave wealth from the *Historical Statistics of the United States* (Sutch 1988), slave wealth likely made up around 15 percent of total wealth in 1860. Subtracting slave wealth from white wealth in 1860, the wealth gap "mechanically" falls from 56:1 to 47:1. Thus, eliminating slave wealth accounted for about 25 percent of the total drop in the wealth gap. While important, the elimination of slave wealth alone

---

[3]For an illustration of this pattern, see online Appendix Figure A.1.4. According to the SCF+, in 2019 the share of Black in the wealthiest 0.1 percent is less than 1 percent. Note that besides Black and non-Hispanic white, the SCF+ provides an additional racial category defined as "Others", which includes among others Asians and Hispanics. This group's share in the top 0.1 percent is around 7 percent in the same year.

[4]Data on slave wealth are available at https://hsus.cambridge.org/HSUSWeb/toc/showTable.do?id=Bb209-218.

Figure 3
White-Black Per Capita Wealth Ratio: 1860–2020
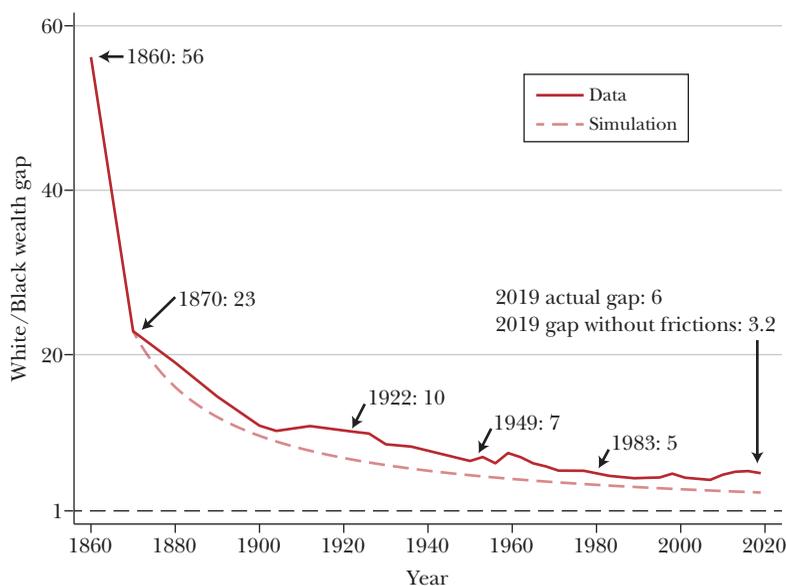
Source: Racial wealth gap series are from Derenoncourt et al. (2023a). Primary data sources are 1860 and 1870 complete count censuses (Ruggles et al. 2021), Southern state auditor reports 1866–1929, Work (1932), Work (1938), Saez and Zucman (2016), and the US Census Bureau's "Wealth, Public Debt, and Taxation report" (Department of Commerce 1924).
Note: The red solid line presents the white-to-Black per capita wealth ratios from 1860 to 2020. The light red dashed line presents the simulated wealth gap series under the scenario that after Emancipation, Black and white Americans had the same saving rates and capital gains. The only difference across Black and white Americans is coming from their initial starting point of wealth and income growth rates. Full details on the construction of this series are available in Derenoncourt et al. (2023a).

does not account for the bulk of the reduction in the wealth gap from 1860 to 1870. There was also a sharp drop in southern land prices and wartime destruction of southern property (Ager, Boustan, and Eriksson 2021). In combination, these forces depressed wealth growth of the white population over these decades.

Despite the ongoing presence of Jim Crow legislation hindering equal opportunity for Black Americans, the racial wealth gap continued to fall in the decades after 1870, reaching about 10:1 by 1920. During the Great Depression, the wealth gap declined further, despite the fact that New Deal era relief and social insurance policies tended to exclude regions or sectors with a large representation of Black workers (Katznelson 2005). Rather, the decline in asset prices during the Great Depression, especially the drop in valuation of stocks and business equity that were predominantly in the hands of white Americans, seems to have resulted in a

substantial drop in white wealth and a narrowing of the wealth gap. The decades after World War II until 1980 saw dramatic progress along a number of dimensions of racial justice. Starting from a level of 7:1 in 1949, the wealth gap reached a level of 5:1 in 1983 (almost 1 percent rate of convergence per annum). However, the wealth gap left its convergence path in the last quarter of the twentieth century. The most recent numbers in 2019 shows that the wealth gap two decades into the twenty-first century has returned to a level of 6:1.

To understand the drivers of long-run racial wealth convergence, Derenoncourt et al. (2023a) use a stylized framework of wealth accumulation, where the racial differences in wealth accumulation depend on their (1) initial conditions after Emancipation and (2) different wealth accumulating conditions such as saving rates and capital gains. The light red dashed line in Figure 3 presents the simulated wealth gap series under the scenario that Black and white Americans have had identical saving rates and capital gain rates since Emancipation. Interestingly, the convergence follows the same hockey stick shape, thus highlighting that the convergence path can be mainly explained by huge differences in initial conditions of Black and white Americans, and not because of their differences in savings rates and capital gains. In addition, the simulation shows that even under equal wealth accumulation conditions, the historical differences in 1870 still leave their on today's wealth gap: white Americans would, on average, own $3 per $1 of wealth of Black Americans today.

By contrast, the actual white-to-Black per capita wealth ratio in 2019 is 6:1, which suggests that Black-white differences in savings rates and in capital gains contributed to slower convergence. In Derenoncourt et al. (2023a), we argue further that the stagnation of the Black-white wealth gap since the 1980s is largely due to the ways in which white and Black capital gains became less favorable to wealth convergence during this time.

Wealth of the Median Black Household over Time

The average wealth gap might hide substantial heterogeneity and in this way miss the evolution of the financial situation of the typical household. Therefore, in this section we compile the first estimates of the wealth of the median Black household and compare this to median white wealth. There is an age-old argument on whether the mean or the median offers a better representation of a right-skewed distribution. For assessing the situation of the majority of Black households, looking at the median of the wealth distribution, rather than the mean, can provide important additional insights. At the time of Emancipation, the median black household had zero wealth. Studying when that median value turns positive helps us understand whether the gains in average wealth shown in the previous section were broadly shared across the majority of Black households, or whether wealth gains for Black Americans were mainly associated with a smaller group at the top of the Black wealth distribution.

Before 1940: When Did the Majority of Black Households Have Positive Net Wealth?

To obtain a more comprehensive picture of the evolving distribution of Black wealth over time, the first step is to determine what share of the Black population possessed nontrivial and measurable wealth over time. By definition, we can only study the difference in wealth between the median Black and white households from the time when at least half of the Black population owned some marketable assets. In essence, we will go through the evolution of the main assets on the household balance sheet to determine when the majority of Black households possessed different types of assets. We will focus on homeownership and personal property that includes financial assets such as stocks, bonds, mortgages, and notes, as well as nonfinancial assets other than housing, such as livestock, plates, jewels, and furniture.

Homeownership is often the largest single asset for US households, and information on homeownership by race is available from 1860 to 2020 from US Census data. Our data are extracted from complete-count Census for the time period from 1860 to 1940.[5] For the time period from 1960 through 2019, we add further data from the American Community Survey conducted by the Census Bureau. As a robustness check, we also construct homeownership rates with the augmented SCF+ data for 1950–2019.

Figure 4 presents the results of this exercise. Homeownership rates of white households have been higher than for Black households throughout the 160-year period. At each point in time up through 1940, around 50 percent of white households owned a house. White homeownership trended down slightly until 1940 (and especially during the Great Depression decade of the 1930s), but then surged strongly until the 1960s in the postwar housing boom, partly aided by policies like the Servicemen's Readjustment Act of 1944, better known as the "GI Bill" (Fetter 2013). Homeownership rates for white families continued to trend upwards until the financial crisis of 2008. Thus, based on homeownership alone, the median white household held positive wealth since 1860.

For Black households, homeownership rates increased substantially in the first 30 years after Emancipation. Starting from almost zero before Emancipation, in 1900 around 20 percent of Black households were homeowners. Between 1900 and 1940, however, Black homeownership rates flat-lined. In the decades after World War II, Black homeownership rates increased strongly between 1940 and 1960, from 20 percent to almost 40 percent, mirroring the developments observed for white Americans, albeit on a persistently lower level. Black homeownership also dropped much more strongly during the global financial crisis of 2007–2009. In 2019, homeownership rates of Black households remain at around 45 percent, roughly the same level as in 1970. Despite the increase in homeownership rates in the last 160 years, Black homeownership rates never exceeded 50 percent and only came close

---

[5] The resulting series during this time period can be compared to the series compiled by Collins and Margo (2011).

Figure 4
White and Black Homeownership Rates, 1860–2020
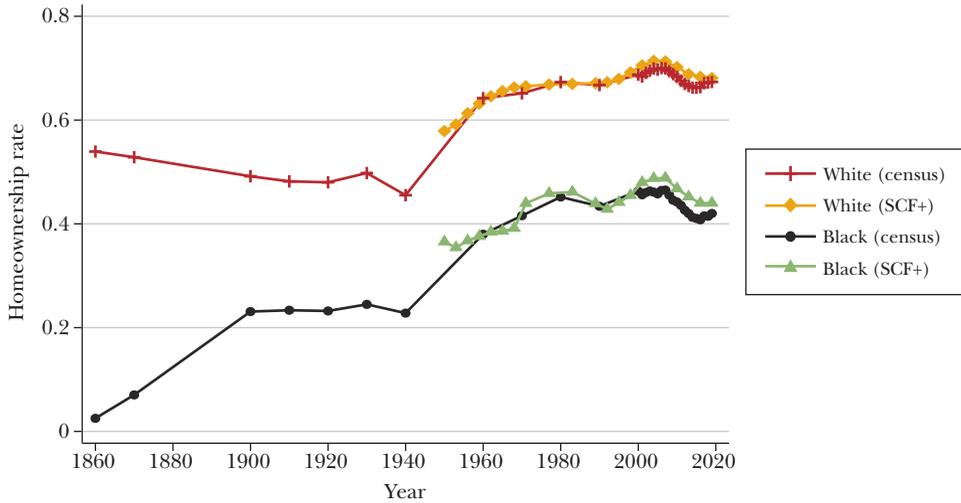
Note: The line with cross-marks plots white homeownership rates from the census, and the line with dots shows Black homeownership rates from the census. The squares and triangles show white and Black homeownership rates, respectively, estimated from the SCF+ microdata.

for a short moment in the housing boom of the 2000s. Thus, homeownership alone would not be enough for the median Black household to have positive wealth.

What about other personal property other than housing? Detailed individual-level data on holdings of personal wealth are scarce. We have information on certain elements of personal property for 1860 and 1870 coming from the Census, including bonds, notes, and wealth in enslaved persons in 1860. But after that, a complete picture of wealth by asset class is only available starting from 1950 onwards with the augmented SCF+ data. In the 1870 Census data, around 30 percent of Black households reported some form of wealth, which was substantially higher than the roughly 10 percent of Black households that owned a house that year. In the 1950 SCF+ data, slightly more than 50 percent of Black households reported positive wealth, which is higher than the approximately 30 percent of Black households that owned houses at that time. Thus, at some point between 1870 and 1950 the majority of Black households accumulated measurable wealth. But when?

In light of the data gap between 1870 and 1950, we can only approximate this transition indirectly. We impute the shares of the Black and white population with positive wealth by exploiting sociodemographic characteristics of households with personal wealth first by projecting forward from the 1860 and 1870 censuses and then by projecting backward from the 1950 SCF+ survey wave. Details of our calculations are available in the online Appendix. Our basic approach was to begin

by investigating the sociodemographic characteristics of a household head *i* that had personal property, who lives in state *s* in year *t* and does not possess housing wealth. We included sociodemographic characteristics such as family size, gender, literacy, urban status, and age (and its square), as well as labor market characteristics such as labor force status and occupational characteristics (farmer/laborer/ professionals). We first implement forward-looking predictions from the 1860 and 1870 Census data.[6] We use our estimated coefficients to impute the share of positive wealth holdings, including homeownership for the decadal years up to 1940. In effect, we are asking how the share of the Black population with personal wealth would have changed if the relationship between sociodemographic characteristics and nonhousing wealth remained fixed, and if only the sociodemographic characteristics shifted over time

Of course, we are aware of the fact that the sociodemographics of property owners may have changed over time and the 1860/1870 characteristics may not be representative for the mid-twentieth century. Therefore, in a second step, we also compute backward-looking predictions using the same approach and sociodemographic characteristics based on the 1950 data of the SCF+, and then work back to 1940. Hence, we ask what share of white and Black households had wealth in 1940 if the relationship between sociodemographic characteristics and wealth holding from 1950 also applied in 1940.

Figure 5 presents the data for 1860, 1870, and from 1950 onward, together with our imputation results for the period from 1880 to 1940. In 1860 and 1870, we observe opposite dynamics in the Black and white shares with positive wealth. In 1860, only around 10 percent of the Black population possessed wealth, while more than 85 percent of white had positive wealth holdings. In the aftermath of Emancipation, Black shares increased drastically to a level of around 32 percent, which is not surprising as Black Americans were finally allowed to possess wealth. In comparison, white shares decreased to 78 percent, with the strong decrease in white shares mostly coming from southern states. This pattern is unsurprising because Emancipation led to the total nullification of all slave wealth, which hit wealthy Southerners the most (Ager, Boustan, and Eriksson 2021).[7]

After Emancipation, our forward imputations (dashed red line for white and dashed black line for Black) show that white shares with positive wealth in 1900 recovered slightly from 79 percent to 83 percent, and remain quite stable throughout the next 40 years. Compared to this, we observe for the Black population a stronger increase in 1900, with around 40 percent of the total Black population owning positive wealth (in 1870, the share was 32 percent). Nevertheless, the share of the Black population with positive wealth remained stable over the time period from 1900 to 1940 and never reached the 50 percent threshold. Hence, our estimation predicts

---

[6] All census data were obtained from Ruggles et al. (2021).

[7] In panels A and B of Figure 1, we also present the distribution of Black and white wealth during 1860 and 1870. In line with our results with the shares with positive wealth, we also observe a larger spread in Black wealth distribution after Emancipation, while the white wealth distribution contracts.

Figure 5
**Black and White Shares with Positive Wealth, 1860–2020**
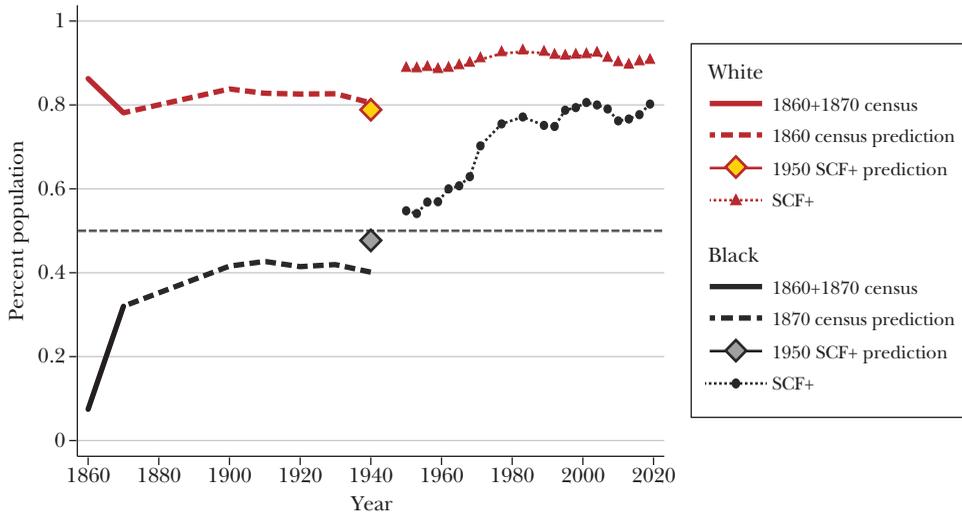
Source: Census, SCF+ (Derenoncourt et al. 2023b), and authors' calculations.
Note: The solid black and red line present Black and white shares with positive wealth, respectively, during 1860–1870 using census data. The solid red line with cross marks plots white shares with positive wealth, and the black line with dots shows Black shares during 1950–2020 coming from data of the SCF+. The dashed black and red lines present our forward imputations of Black and white shares with positive wealth, respectively. The yellow diamond plots our backward imputation of white shares in 1940, while the grey diamond plots the results for Black shares. Finally, the grey dashed line visualizes the 50 percent threshold.

zero wealth for the median Black household until the mid-twentieth century. When we impute the 1940 values using backward-looking imputation based on the 1950 data, we obtain similar results for white (yellow diamond) and corroborate the zero-wealth result for the Black median. Using the 1950 data, the share of positive wealth holdings of Black Americans is slightly higher than the share imputed with the 1870 estimates (forward-looking imputation, grey diamond), but still below the 50 percent threshold. This finding of zero wealth for the median Black household circa 1940 aligns with the fact that the early-twentieth century is characterized by high racial oppression, such as Jim Crow laws, the revival of the Ku Klux Klan during the early 1920s, and various events of violent destruction or expropriation of their property (Albright et al. 2021; Cook 2014; Messer, Shriver, and Adams 2018). The Great Depression during the early 1930s and its aftermath further suppressed the accumulation of wealth.

Our analysis suggests that the median Black household started to possess measurable levels of wealth only after 1940. From Emancipation until the mid-twentieth century, despite the overall growth in total and average Black wealth

Figure 6
The Racial Wealth at the Median

Panel A. Wealth gap

Panel B. Growth rates in median wealth by racial group
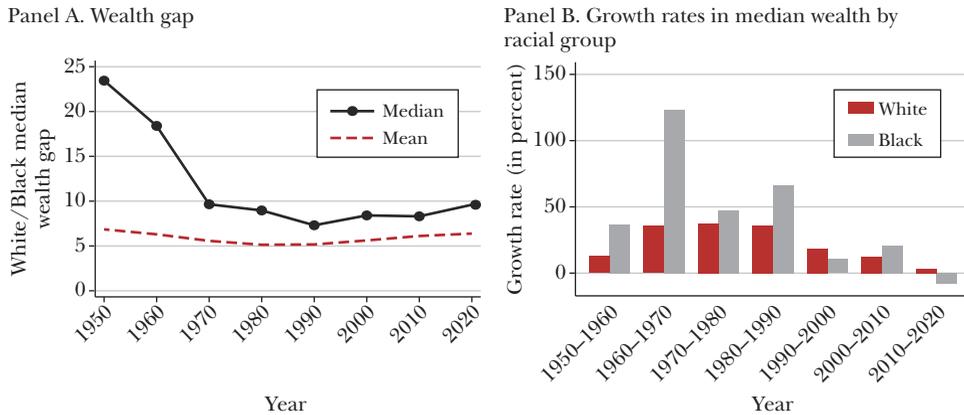
Source: SCF+ (Derenoncourt et al. 2023b).
Note: Panel A presents the household-level white-to-Black wealth gaps at the mean and median. Panel B presents growth rates in Black and white wealth at the median for each decade from 1950 to 2020.

documented in the previous section, the majority of Black Americans remained propertyless. In the next section, we concentrate on the post-1950 period using data from the augmented SCF+ to investigate the development of the wealth gap at the median after World War II.

Post–World War II: The Evolution of the Median White-Black Wealth Gap

For the post–World War II period, we track the evolution of the racial wealth gap for the median household using the microdata from the augmented SCF+ (Kuhn, Schularick, and Steins 2020). In panel A of Figure 6, we present the evolution of the white-to-Black wealth gap at the median of the household wealth distribution— while also showing the mean for comparison. Throughout the post–World War II period, the Black-white wealth gap at the median was substantially larger than the wealth gap at the mean. In 1950, the median wealth gap was nearly 25:1. By 1970, this number has fallen substantially, reaching a level of 10:1—however, the gap has remained at this level for the last five decades. In contrast to the wealth gap at the median, the gap at the mean followed a u-shape over the last 70 years, ranging from just under 5:1 to around 7:1. Hence, a focus on the average wealth gap alone hides important trends in terms of wealth-building and financial inclusion for the typical Black household. Financial inclusion and the support for broad-based wealth building are regularly on policymakers' agendas, and analysis of the average alone can only inform these debates to a limited extent.

What can explain the sharp drop in the median wealth gap, particularly between 1960 and 1970? Panel B of Figure 6 shows the growth rates in median wealth by racial group for each decade between 1950 and 2020. Recall that it was only during

the decade between 1940 and 1950 that our estimation points to positive wealth levels for the median Black household. Black wealth at the median grew strongly from low wealth levels between 1960 and 1970, precisely when the wealth gap at the median fell by more than half. This stark increase in median Black wealth during this decade suggests that civil rights era policies and improvements in labor standards that disproportionately benefited Black workers in the late 1960s may have also translated into absolute and relative improvements in the wealth position of median Black households. Although the average wealth gap shows a declining pattern over this time period, too, it misses the contemporaneous and much stronger trends at the median that, as argued before, might be of particular policy relevance.

In short, the post–World War II period and, in particular, the first two postwar decades witnessed improvement in a substantial widening of wealth holdings among Black Americans. The trend continued until the 1980s, but stagnated in recent decades.

The Evolution of Black Wealth at the Top and Bottom of the Distribution

In a final step in our investigation, we study the available information for trends in Black wealth in the tails of the distribution. As a starting point, Figure 7 presents the Black population shares in the bottom 50 percent, 50–90 percent, and top 10 percent of the US wealth distribution. As a comparison, we also plot the total Black population share in the United States as a red dashed line.

While the Black population share shows a secular upward trend from roughly 8 to 13 percent of the total US population during 1950–2020, the Black population share in the bottom 50 percent is significantly higher. Whereas one in eight Americans is Black overall, almost one in four Americans is Black if we look at households below the median wealth level. Interestingly, the share in the bottom 50 percent has steadily increased since 1950 and closely follows the trend of the overall population share. Such trends emphasize once again the large representation of Black Americans in the poorest group in the United States. Conversely, Black Americans are underrepresented in the upper half of the wealth distribution. Black Americans represent less than 10 percent of the 50–90 percent slice of the wealth distribution in the United States, and less than 2 percent of the top 10 percent. In particular, Black representation at the very top is persistent with barely any change since 1950. Looking along the entire wealth distribution, we find that the Black wealth distribution is shifted towards lower levels throughout the entire distribution.

The concept of a racial "rank gap" offers an alternative way to conceptualize the evolution of wealth differences between groups at specific parts of the distribution (Bayer and Charles 2018; Kuhn, Schularick, and Steins 2020). The rank gap is the percentage point difference between the rank of a given percentile in the Black and white wealth distribution. A rank gap of –20 for the median Black household wealth means that the place of that household in the white distribution would be

Figure 7
Black Population Shares along the Wealth Distribution
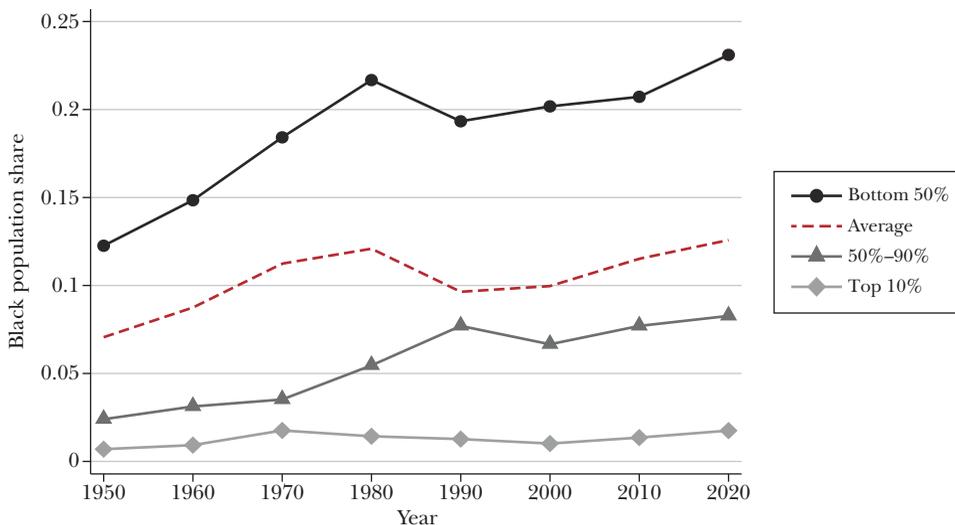
Notes: The red dashed line presents the Black population shares in the United States. The solid black line with dots plots the Black population share among the bottom 50 percent, the dark grey line with triangles the share among the 50 percent–90 percent, and the light grey line with squares the share among the top 10 percent of the total US wealth distribution.

20 percentile points lower—that is, only at the 30th percentile. Figure 8 shows the wealth rank gap at the median (red solid line) and at the 90th percentile (black dashed line). We expand earlier results for the wealth rank gap at the median and the 90th percentile from Kuhn, Schularick, and Steins (2020).

For the median, the average rank gap since 1950 is close to –30 percentage points, implying that the median Black household finds itself at the 20th percentile of the wealth distribution of white households. Put differently, the typical black household is poorer than 80 percent of white households. The rank gap at the median has improved consistently until 2010, but has increased again since then. Today, the median Black household would find itself in bottom third of the white distribution. The rank gap is equally large and persistent at the top. The 90th percentile of the Black wealth distribution yields a rank gap of close to –30 on average. In other words, the Black household at the 90th percentile is fairly close to the median white household. Although some progress is visible here, too, the gap has stalled at –25 since the 1990s. Being at the threshold for the top 10 percent of Black households puts such households only at the 60th percentile of the distribution of white households.

In summary, our inspection of the tails shows that Black Americans are disadvantaged at both ends. Black Americans are overrepresented in the bottom 50 percent

Figure 8
Racial Wealth Rank Gaps: Median and 90th Percentile
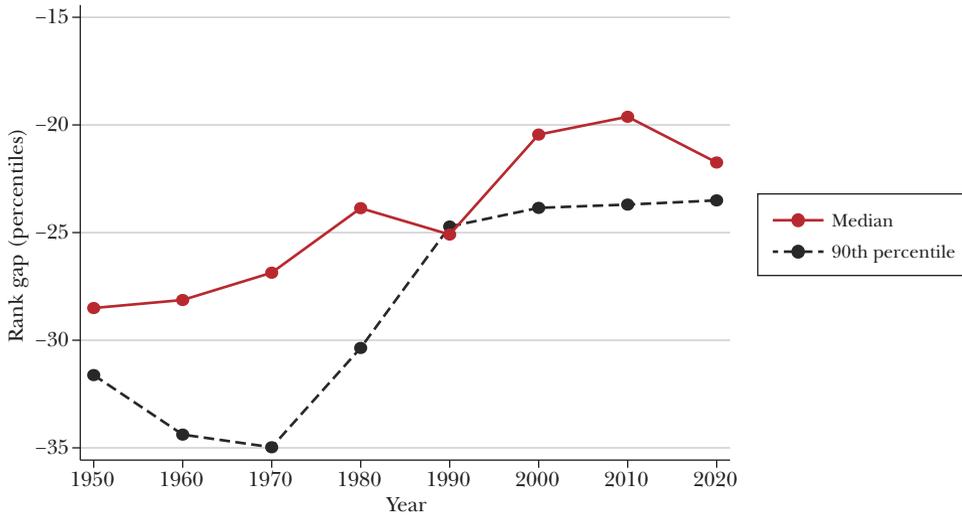
of the overall US wealth distribution and underrepresented at the top. Also, given the same rank, Black Americans are consistently worse off than their white counterparts. In particular, this gap seems to be highly persistent at the top. Given that the recent divergence in the racial wealth gap can be explained through asset price surges and rising wealth concentration at the very top, such trends set a sobering perspective for the future development of the average racial wealth gap.

Conclusion

The exclusion of the enslaved Black population from wealth accumulation until 1865 continues to leave its mark on the wealth distribution today. While there has been some convergence in wealth levels over the past 150 years, convergence processes can be inherently slow when initial gaps are so large. Even under hypothetical conditions where Black Americans had slightly higher income growth than white Americans, but equal savings rates and capital gains, full wealth convergence remains a distant goal over the next two centuries. Reality is less optimistic, given that we live in an era of high wealth-to-income ratios in which differences in capital gains on existing assets play a dominant role for wealth growth. In recent decades,

capital gains differentials between Black and white Americans were much more than before, mainly due to the greater exposure of the wealth of white households to equity markets and the strong performance of those markets.

There have been some relative improvements for the median Black household. The wealth gap at the median remains wide, but has shrunk substantially in the post–World War II era, especially early in this period. In fact, the median Black household only started to dispose of meaningful wealth around the time of World War II. In other parts of the wealth distribution, Black progress is visible over the past century, but often at a glacial pace and with a tendency towards stagnation or even outright reversal during the past three decades of asset price surges and rising wealth concentration at the very top in the hands of mostly white, rich households.

References

Ager, Philipp, Leah Boustan, and Katherine Eriksson. 2021. "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War." *American Economic Review* 111 (11): 3767–94.

Albright, Alex, Jeremy A. Cook, James J. Feigenbaum, Laura Kincaide, Jason Long, and Nathan Nunn. 2021. "After the Burning: The Economic Effects of the 1921 Tulsa Race Massacre." NBER Working Paper 28985.

Bayer, Patrick, and Kerwin Kofi Charles. 2018. "Divergent Paths: A New Perspective on Earnings Differences between Black and White Men since 1940." *Quarterly Journal of Economics* 133 (3): 1459–1501.

Berlin, Ira. 1975. *Slaves without Masters: The Free Negro in the Antebellum South.* New York: Pantheon Books.

Collins, William J., and Robert A. Margo. 2011. "Race and Home Ownership from the End of the Civil War to the Present". *American Economic Review* 101 (3): 355–59.

Cook, Lisa D. 2014. "Violence and Economic Activity: Evidence from African American Patents, 1870–1940." *Journal of Economic Growth* 19 (2): 221–57.

Department of Commerce. 1924. *Wealth, Public Debt, and Taxation: Estimated National Wealth.* Washington, DC: Bureau of the Census.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. 2023a. "Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860–2020." *Quarterly Journal of Economics.*

Derenoncourt, Ellora, Chi Hyn Kim, Moritz Kuhn, and Moritz Schularick. 2023b. "Replication data for: Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860–2020." Harvard Database. https://doi.org/10.7910/DVN/H6NXUH.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. 2023c. "Replication data for: Changes in the Distribution of Black and White Wealth since the US Civil War." American

Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193729V1.

Dray, Sacha, Camille Landais, and Stefanie Stantcheva. 2023. "Wealth and Property Taxation in the United States." NBER Working Paper 31080.

Du Bois, William Edward Burghardt. 1901. "The Negro Landholder of Georgia." In *Bulletin of the United States Department of Labor*, 35, 647–777. Washington, DC: Government Printing Office.

Fetter, Daniel K. 2013. "How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-Century GI Bills." *American Economic Journal: Economic Policy* 5 (2): 111–47.

Higgs, Robert. 1982. "Accumulation of Property by Southern Blacks before World War I." *American Economic Review* 72 (4): 725–37.

Jordà, Òscar , Moritz Schularick, and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." In *NBER Macroeconomics Annual 2016*, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

Katznelson, Ira. 2005. *When Affirmative Action Was White: An Untold History of Racial Inequality in Twentieth-Century America.* New York: WW Norton and Company.

Kuhn, Moritz, Moritz Schularick, and Ulrike I. Steins. 2020. "Income and Wealth Inequality in America, 1949–2016." *Journal of Political Economy* 128 (9): 3469–3519.

Margo, Robert A. 1984. "Accumulation of Property by Southern Blacks before World War I: Comment and Further Evidence." *American Economic Review* 74 (4): 768–76.

Messer, Chris M., Thomas E. Shriver, and Alison E. Adams. 2018. "The Destruction of Black Wall Street: Tulsa's 1921 Riot and the Eradication of Accumulated Wealth." *American Journal of Economics and Sociology* 77 (3–4): 789–819.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 133 (2): 553–609.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2022. "Distributional National Accounts (Tables II: distributional series, 2022 vintage)." https://gabriel-zucman.eu/files/PSZ2022AppendixTables II(Distrib).xlsx (December 2022).

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "IPUMS USA: Version 11.0." Minneapolis, Minnesota: IPUMS. https://doi.org/10.18128/D010.V11.0. (accessed November 2021).

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131 (2): 519–78.

Schweninger, Loren. 1989. "Black-Owned Businesses in the South, 1790–1880." *Business History Review* 63 (1): 22–60.

Schweninger, Loren. 1990. *Black Property Owners in the South, 1790–1915.* Urbana, IL: University of Illinois Press.

Sutch, Richard. 1988. "Replication Data for: Appendix: The Value of the Slave Population, 1805–1860 (Table A.4 and A.1)." https://hsus.cambridge.org/HSUSWeb/toc/showTable.do?id=Bb209-218 (accessed June 2023).

Walker, Juliet E. K. 1983. "Black Entrepreneurship: An Historical Inquiry." *Business and Economic History* 12: 37–55.

Walker, Juliet E. K. 1986. "Racism, Slavery, and Free Enterprise: Black Entrepreneurship in the United States before the Civil War." *Business History Review* 60 (3): 343–82.

Work, Monroe N., ed. 1932. *Negro Year Book: An Annual Encyclopedia of the Negro, 1931–1932.* Tuskegee, AL: Negro Year Book Publishing Co.

Work, Monroe N., ed. 1938. *Negro Year Book: An Annual Encyclopedia of the Negro, 1937–1938.* Tuskegee, AL: Negro Year Book Publishing Co.

Why Do Retired Households Draw Down Their Wealth So Slowly?

Eric French, John Bailey Jones, and Rory McGee

The life-cycle framework of Modigliani and Brumberg (1954, 1980) postulates that households will smooth consumption by accumulating wealth during their prime earning years and spending it once they retire. The simplest version of the model, with no bequest motives or uncertainty about length of life, further predicts that households will begin decumulating their wealth as soon as they retire and will die with no wealth. This prediction stands in sharp contrast with the data, which show that retired households, especially those with high lifetime income, decumulate their assets very slowly. Indeed, many die leaving large estates. In the past two decades, a growing literature has sought to explain this "retirement savings puzzle."

In this article we review and evaluate the three leading explanations for why older households seem reluctant to draw down their wealth. Although Modigliani and Brumberg (1954) did not formalize these explanations, they described all three. First, the precautionary motive arises because retired households face the risk of

■ *Eric French is the Montague Burton Professor of Industrial Relations and Labour Economics, University of Cambridge, Cambridge, United Kingdom. He is also a Research Fellow, Centre for Economic Policy Research, and Co-Director at the ESRC Centre for the Microeconomic Analysis of Public Policy, Institute for Fiscal Studies, both in London, United Kingdom. John Bailey Jones is Vice President of Microeconomic Analysis in the Research Department, Federal Reserve Bank of Richmond, Richmond, Virginia. Rory McGee is an Assistant Professor of Economics, University of Western Ontario, London, Ontario, Canada. He is also an International Research Associate, Institute for Fiscal Studies, London, United Kingdom. Their email addresses are ebf26@cam.ac.uk, john.jones@rich.frb.org, and rmcgee4@uwo.ca.*

living long and incurring catastrophic medical and long-term care expenses, and so households may hold onto their wealth to cover such expenses. Second, the bequest motive can arise either because individuals enjoy leaving bequests, or because they use bequests to reward their caregivers and elicit care. Third, households may need "to have an equity in certain kinds of assets before [they] can receive services from them." In particular, retirees may be reluctant to decumulate their housing wealth, as many enjoy living in their homes, find it costly to move, and face an underdeveloped market for "reverse mortgages" that would allow them to decumulate housing wealth without leaving their homes. Indeed, older households decumulate housing more slowly than other forms of wealth.

These three explanations are neither original to the life-cycle model nor mutually exclusive. For example, well before the introduction of the life-cycle model, Keynes (1936) discussed at length the desires to "build up a reserve against unforeseen contingencies" (the precautionary motive) and "bequeath a fortune" (the bequest motive). What is new in the last couple of decades is the availability of better data, which allow researchers to measure risks and outcomes more accurately, and greater computing power, which facilitates estimation of models containing multiple motivations for saving. Simultaneously accounting for all motivations is important because, as Modigliani and Brumberg (1954) pointed out, assets can serve multiple purposes: "For example, the ownership of a house is a source of current services; it may be used to satisfy part of the consumption planned for after retirement; it may be bequeathed; and, finally, it is a source of funds in emergencies."

In this article, we discuss what researchers in this area have learned about the retirement savings puzzle, in effect updating and extending the more technical survey in De Nardi, French, and Jones (2016b). We first describe how postretirement wealth changes with age and income. We then introduce the competing explanations for these savings patterns, along with supporting evidence. The fungibility of wealth makes it difficult to disentangle the explanations, and thus we explain the leading approaches for doing so. In our opinion, there is considerable evidence that the precautionary motive and the bequest motive are both important. Although the evidence on its role is less developed than for the other two motives, housing deserves further study as well, if only for its prominence in most household portfolios.

The relative importance of the precautionary, bequest, and housing motives in explaining the slow decumulation of wealth is not only of academic interest. It is of great policy relevance in an aging society. We therefore conclude by discussing longer-term savings trends and the importance of understanding retirement saving motives when contemplating the welfare implications of reforms to Social Security, Medicare, or Medicaid. We also discuss the limited use of financial products such as long-term care insurance and annuities, which in principle should insure retirees against their risks more effectively than wealth. Once again, understanding the different saving motivations is key. If the precautionary motive is strong, the low take-up of these products may reflect market failures, and government intervention in these markets may be warranted. If on the other hand the precautionary motive is weak and retirees save mostly for bequests or homeownership, then the value of

these financial products may be modest and government intervention may provide little if any benefit.

Wealth Profiles after Retirement

An important factor determining the welfare of retirees is their consumption, which is funded primarily by net worth, Social Security benefits, and defined benefit private pensions. With the notable exception of households in the bottom lifetime income decile, who rely almost completely on Social Security, net worth is a major source of funds. For households with above-median lifetime income, it is the most important source of funds (Scholz, Seshadri, and Khitatrakun 2006).

As is standard in this literature, our measure of wealth is net worth excluding annuitized wealth, the (discounted) value of the Social Security and other defined benefit pension income that households expect to receive over the remainders of their lives. Although annuitized wealth is an important source of retirement funding, it behaves very differently from other forms of wealth: it cannot be bequeathed, and its value is largely a mechanical function of how long individuals expect to live. In particular, annuitized wealth declines mechanically as individuals age and expected lifespans shorten, and falls to zero at death. This means that, in contrast to net worth, annuitized wealth falls rapidly after retirement (Love, Palumbo, and Smith 2009).

In this section we establish three facts about net worth: (1) the wealth of older households declines slowly with age; (2) the decline is slower among the rich; and (3) those with low income have little wealth.

To document the wealth of the elderly, we use data on older US households from the Assets and Health Dynamics of the Oldest Old (AHEAD) cohort of the Health and Retirement Survey (HRS). The HRS has several features that make it well-suited for studying the wealth dynamics of older households. It is a nationally representative longitudinal dataset that follows households to the ends of their lives and beyond, using "exit" interviews with survivors to measure end-of-life expenses and bequests. It combines detailed financial information with a battery of health measures, allowing researchers to quantify the longevity and medical spending risks that older households face and to observe households as they respond to major life events like the death of a spouse. Moreover, it is linked to several administrative data sources. These include the National Death Index, which provides an accurate measure of mortality.

We measure wealth in terms of net worth, which is the sum of the value of housing and real estate, automobiles, liquid assets (money market accounts, savings accounts, Treasury bills, and so on), individual retirement accounts (IRA) and Keogh accounts (and other defined contribution plans), stocks, the value of any farms or businesses, mutual funds, bonds, "other" assets, and investment trusts—minus mortgages and other debts. We use data starting in 1996 and every two years thereafter through 2014. Our sample selection restrictions follow De Nardi et al. (2023a) and are discussed there in more detail.

Older households differ along a variety of dimensions that potentially affect their saving decisions. Many of these differences (like education level) are correlated with the households' lifetime earnings or permanent income. Households with different permanent income ranks receive different flows of retirement income and face different processes for health, mortality, and medical expenses. Because permanent income is determined prior to retirement, it provides a useful basis for stratifying retired households. Our proxy for permanent income is based on post-retirement annuitized income, which is the sum of Social Security benefits, defined benefit pension benefits, veteran's benefits, and annuities. Because households with higher lifetime earnings tend to have higher annuity incomes—for example, Social Security payments are higher for people with a history of higher earnings—this measure is a good indicator of the income people received when they worked. We use annuitized income to construct a permanent income measure comparable across households of different ages and sizes.[1]

Figure 1 presents median wealth conditional on age and permanent income tercile for the cohort aged 71–76 (which we index as 75) in 1996, an age by which the great majority of households have completely retired. These profiles come from De Nardi et al. (2023a), who show that the facts we highlight here hold for other cohorts as well. Figure 1 presents wealth profiles for the unbalanced panel; each point represents the median for all the members of an age-income tercile cell who are alive at a particular date.[2]

The left panel shows wealth profiles for households who are single (most of whom are widowed or divorced) throughout the entire sample period. The median 75-year-old in the top income tercile enters our sample with about $200,000 in wealth (in 2014 dollars), while the one at the bottom holds essentially no wealth at all. Over time, those in the top income tercile tend to hold substantial wealth well into their 90s, those in the middle tercile display some asset decumulation as they age, and those at the bottom hold little wealth at any age. Thus, even at older ages, richer people save more, a finding first documented by Dynan, Skinner, and Zeldes (2004) for the whole life cycle.

The right panel of Figure 1 reports median wealth for households who are couples in 1996. In later years, many of these households lose a member and become singles, in which case we report the wealth of the surviving spouses. Couples are richer than singles. Couples in the highest income tercile hold around $300,000, and even the couples in the lowest income tercile hold over $70,000 in the early stages of their retirement. As with the singles, couples in the highest income tercile hold large amounts of wealth well into their 90s, while those in the lowest tercile hold little wealth. Many couples experience a significant decline in wealth when

---

[1]More specifically, to construct our income measure, we regress annuity income on a household fixed effect and controls for household composition and age. The rank order of each household's estimated fixed effect provides our measure of its income. This is a time-invariant measure that follows the household even after one of its members dies.

[2]We also include the value of bequests left after the final household member dies.

Evolution of Median Wealth for Retirees
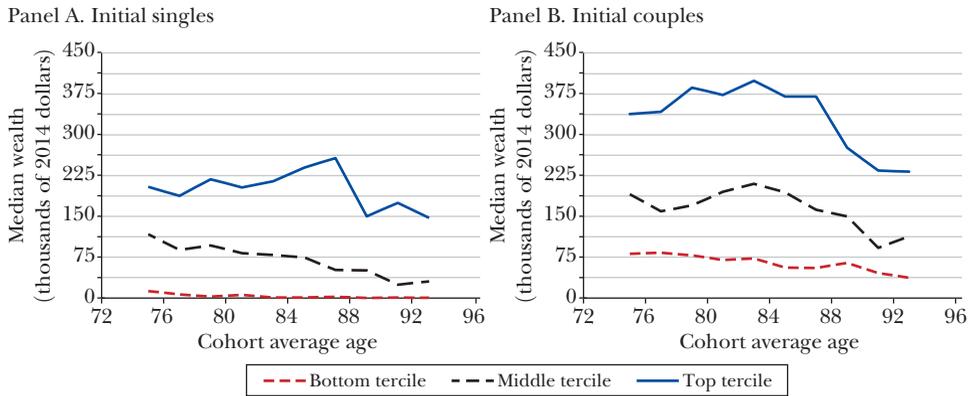
Source: Data from De Nardi et al. (2023b), based on the Assets and Health Dynamics of the Oldest Old (AHEAD) cohort of the Health and Retirement Survey (HRS).
Note: Each line shows medians wealth over the period 1996–2014 for a subset of the AHEAD households aged 71–76 in 1996.

one spouse dies (French et al. 2006; Poterba, Venti, and Wise 2011; De Nardi et al. 2023a). Some of this decline is attributable to end-of-life medical and other expenses, but most is due to bequests to nonspousal heirs. As married households become single, this drop imparts a downward slope on their asset profiles. While both spouses are still alive, couples run down their assets at least as slowly as singles.

It is well-documented that health and wealth are positively correlated (for instance, Smith 1999). As a result, poor people die more quickly, and as a cohort ages, its surviving members are increasingly likely to be rich. Failing to account for this mortality bias will lead a researcher to understate asset decumulation late in life (Shorrocks 1975). De Nardi, French, and Jones (2010) show that mortality bias is quantitatively important, although conditioning on income, as we do in the above graphs, reduces its effects. But regardless of how mortality bias is addressed, the puzzle remains: the asset decumulation of older households is significantly slower than that implied by a simple life-cycle model where individuals face no uncertainly and receive no utility from leaving bequests. In the next section we discuss extensions to the life-cycle model that encourage older households to save more and to die with positive wealth.

Drivers of Savings

Precautionary Saving Motives

One explanation for why retirees appear reluctant to spend down their wealth is that, by saving, they insure themselves against the risk and associated costs of living long and having high medical spending.

In an important early study, Davies (1981) showed that when lifespans are uncertain and there is no annuity income, individuals with reasonable levels of risk aversion will never fully deplete their wealth. The risk of living long may be especially strong for rich people, women, and people in good health, who tend to live longer than their poor, male, and sick counterparts. Using mortality rates estimated from the AHEAD data, De Nardi, French, and Jones (2009) find that an unhealthy 70-year-old male at the bottom quintile of the income distribution expects to live only six more years, while a healthy woman at the top quintile of the permanent income distribution expects to live 16 more years. Similar gradients of longevity with respect to income are found in administrative data and in other countries (for example, see Waldron 2007; Chetty et al. 2016; Banks et al. 2021). The greater longevity of those with high income can partially explain their higher rates of saving, as they have longer lifespans to finance.

Older households also face the risk of high medical spending. Although almost all Americans aged 65 and older receive public health insurance through the Medicare program, Medicare does not cover all health care costs. For example, Medicare only pays for the first 20 days of a nursing home stay (and part of the cost for the next 80 days). Some households have these and other expenses covered by Medicaid, another public health insurance program, but Medicaid is available only to those with limited financial resources. This leaves many retirees having to make significant payments out of pocket.

To give a sense of the medical spending risk for the elderly, Figure 2 shows average medical expenses conditional on age and income for singles, taken from De Nardi et al. (2023a). We focus here on the medical spending of singles, but the spending of couples exhibits similar patterns: for any level of age and income, the medical spending of married households is roughly double that of singles. Because all households receive support from Medicare, we do not include Medicare expenditures in the figure. The left panel of the Figure includes Medicaid payments, however, because Medicaid is means-tested. This means that medical expenses that are covered by Medicaid among poorer households are paid out of pocket by richer ones.

The left panel of Figure 2 shows the sum of out-of-pocket expenses and the payments made by Medicaid. Out-of-pocket expenses are the sum of what an individual spends on drugs, hospital stays, nursing home care, home health care, doctor visits, dental visits, and outpatient care, along with premia for private and Medicare insurance. The public component of the Health and Retirement Survey lacks Medicaid spending data, but we can impute it by combining the HRS with the administrative data contained in the Medicare Current Beneficiary Survey (for a description, see De Nardi et al. 2023a).

The left panel of Figure 2 shows that medical expenses rise rapidly with age. For individuals in the middle income tercile, mean spending rises from roughly $6,000 at age 76 to $26,000 at age 100. Medical expenses rise with age because older individuals are more likely to incur costly end-of-life expenses (French et al. 2006; Marshall, McGarry, and Skinner 2011) and because older individuals incur higher out-of-pocket expenses, such as nursing home care, while alive.

Mean Medical Spending of Retired Singles

Panel A. Out-of-pocket plus Medicaid

Panel B. Out-of-pocket
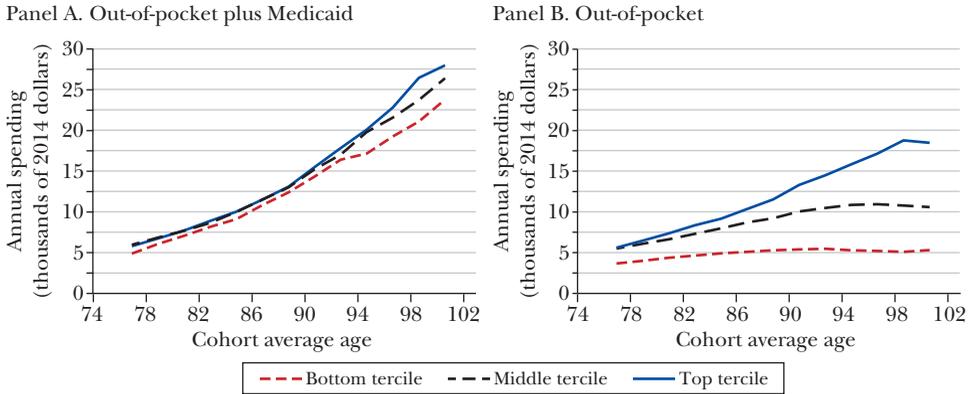
Source: Data from De Nardi et al. (2023b).
Note: Each line shows simulated mean spending of single households by income tercile.

The right panel of Figure 2 shows out-of-pocket medical expenses in isolation; comparing this panel to the one on the left reveals the extent to which Medicaid reduces out-of-pocket expenditures. Because people with low wealth on average receive more assistance from Medicaid, the income gradient for out-of-pocket spending is far steeper than the gradient for total spending. Given that out-of-pocket medical expenditures rise with permanent income, the saving motives they generate should be stronger for those with higher income, causing them to decumulate wealth more slowly.

Medical spending among retirees is not only high, but its distribution is very concentrated (De Nardi et al. 2016), with the top 5 percent of spenders accounting for 49.1 percent of out-of-pocket expenditures in any year. The risk does not average out over time. Calculating the present value of remaining lifetime medical spending, Arapakis et al. (2021) find that the 90th percentile of discounted medical spending at age 65 is twice the size of the mean.

Medical expense uncertainty reinforces the risks associated with lifespan uncertainty (De Nardi, French, and Jones 2009) and increases the impact of medical expenses on saving. The total effect of medical spending is potentially large: modelling the entire life cycle, Kopecky and Koreshkova (2014) calculate that 13.5 percent of aggregate US wealth is attributable to saving for old-age medical expenditures.

Because poor health raises medical spending and shortens lifespans, it affects a household's lifetime financial resources and spending horizon. This in turn affects the household's nonmedical consumption and its marginal utility. Poor health may also affect the marginal utility of nonmedical consumption more directly. For example, functional limitations likely reduce the marginal utility of recreational goods like ski equipment, while raising the marginal utility of home services like

housecleaning and lawn care. If the marginal utility of nonmedical consumption generally rises at older ages because of declining health, retirees would have another reason to hold onto wealth. Laitner, Silverman, and Stolyarov (2018) show that the risk of an increase in the marginal utility of consumption is in many respects equivalent to the risk of higher medical expenses. The literature has yet to reach a consensus, however, about whether bad health raises or lowers the marginal utility of consumption.[3]

In addition to changes in health, events such as the need to acquire a new car or support a grandchild's education can impact the marginal utility of consumption. Inferring fluctuations in marginal utility directly from observed consumption, Christensen, Kallestrup-Lamb, and Kennan (2022) find these fluctuations to be an important driver of retiree savings.

It is not obvious whether marriage increases or reduces household risk. Couples may be able to pool their risks and wealth, and they may be able to partially self-insure by having the healthier partner care for the sicker one. Conversely, two-person households face the risk of having one person die. While single households likely have lower needs, the death of the husband often leads to a large reduction in the wife's income: widows are much more likely to be impoverished than wives (Braun, Kopecky, and Koreshkova 2017). Saving is an important mechanism for insuring against this risk.

Programs that provide social insurance to poorer households should weaken precautionary saving motives. In the United States, the two most important programs in which the elderly receive income- and asset-tested transfers from the government are Medicaid (for medical expenses) and Supplemental Security Income (cash payments). Such means-tested programs discourage saving. They reduce both the average level of medical spending (as shown in Figure 2) and the risk of catastrophic expenses. Moreover, they impose a steep implicit tax: when a low-income household receives means-tested insurance, increases in its wealth lead to lower benefits, with little if any change in the resources available for consumption. Means-tested insurance thus has the potential to crowd out private saving, especially among the poor (Hubbard, Skinner, and Zeldes 1995). As a result, social insurance programs could help to explain why low-income retirees hold such modest amounts of wealth.

The quasi-experimental evidence on the impact of Medicaid and other transfer programs on savings is mixed, with some studies finding evidence that asset-tested transfer programs reduce private savings (Greenhalgh-Stanley 2012), and others finding they do not (Hurst and Ziliak 2006; Gardner and Gilleskie 2012). One potential reason why these results are mixed is that policy reforms affect the rate of saving, which in turn alters the level of wealth only slowly. A reform that has a relatively large effect on wealth in the long run may have very modest short-run effects. Consequently, a popular approach for evaluating the impact of policy reforms is to calibrate or estimate structural models with realistic risks and means-tested

---

[3] To give two recent examples, Blundell et al. (2020) find that declines in health reduce the marginal utility of consumption, while Ameriks et al. (2020) find that requiring long-term care raises it.

insurance. These models are usually required to match additional features of the data, such as wealth levels for different households at different ages. Once estimated, the models can be used to evaluate policy reforms, including their long-term effects. Models of this sort will feature prominently in our discussion below.

Bequest Motives

Efforts to quantify the role of bequests in generating aggregate wealth date back at least to the debate between Kotlikoff and Summers (1981) and Modigliani (1988). Many studies find bequests to be important; for example, in this journal Gale and Scholz (1994) estimate that bequests account for around 30 percent of US wealth holdings. At the same time, most bequests are very modest. As Figure 3 shows, 41 percent of households leave no bequests, and many other bequests are small. Nonetheless, some estates are large—the 95th percentile is over $1,000,000—and the mean nonzero bequest is $335,000. Most estates, but not all, go to children.

The presence of bequests need not imply that households possess bequest motives, because households that die prematurely or incur unusually low medical expenses may find themselves leaving *accidental bequests* as a byproduct of their precautionary saving. In such a case, the skewed distribution of bequests observed in the data may reflect the skewed distribution of the precautionary motives behind the accidental bequests—for example, the tendency of low-income households to rely more heavily on means-tested insurance.

Alternatively, households may enjoy conferring wealth on their heirs and would, even in the absence of risk, choose to make *intentional bequests*. In this case, the concentrated distribution of bequests may indicate that bequests are luxury goods, giving bequests the potential to explain why high-income households decumulate their wealth more slowly. One reason why bequests may be luxuries is that high-income parents are relatively likely to have higher incomes than their children. This gives altruistic high-income parents an incentive to transfer resources to their children that low-income parents lack. Using calibrated overlapping generations models, Castañeda, Díaz-Giménez, and Ríos-Rull (2003) and De Nardi (2004) show that these intergenerational incentives can explain both why high-income parents hold so much wealth and why the distribution of bequests is skewed.

Finally, households may save primarily for precautionary reasons or to maintain their home, but also receive utility from any *incidental bequests* that they might leave. In this case, bequest motives lower the opportunity cost of saving for other reasons (Dynan, Skinner, and Zeldes 2002; Lockwood 2018).

Accidental and incidental bequests are best understood in the context of "terminal" bequests, which are the bequests left when the final member of the household dies. (Figure 3 shows terminal bequests.) In addition to terminal bequests, many couples who lose a spouse leave significant bequests to nonspousal heirs. De Nardi et al. (2023a) show that 31 percent of couples transfer wealth to nonspousal heirs when the first spouse dies, with an average value (when nonzero) of $248,000. Bequests left to nonspousal heirs after the death of the first spouse

Figure 3
Bequests and Their Recipients

Panel A. Distribution

Panel B. Recipients
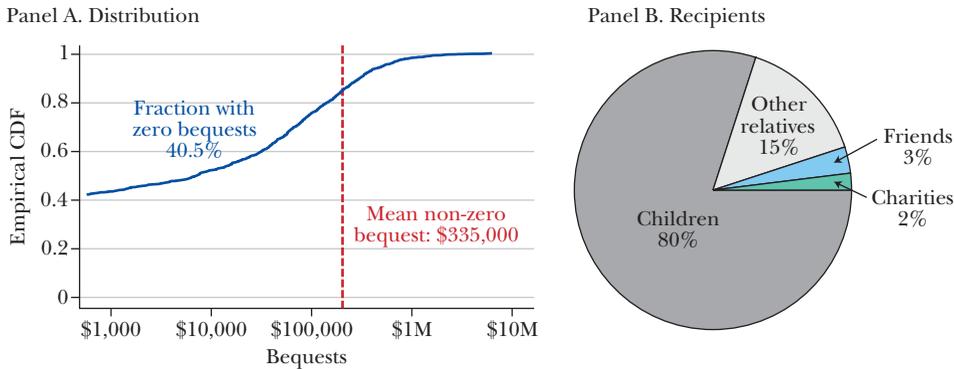
Source: Authors' calculations from the AHEAD cohort of the HRS.
Note: Panel A: Distribution of nonzero bequests at the death of the final household member and their mean value (red dashed line). "CDF" stands for cumulative distribution function. Distribution censored at 99th percentile. Panel B: Average share disbursed to each recipient group.

are almost surely intentional, because the wealth could have been directed to the surviving spouse instead.

Housing

The most important asset for most households in most countries is their primary home. According to data from the Health and Retirement Survey, US retirees on average hold 46 percent of their wealth in housing; the fraction rises to 69 percent among homeowners. Housing differs from other assets by providing consumption services as well as financial returns. Many older individuals seem to prefer living in owner-occupied housing to living in rental properties, perhaps for sentimental reasons or because they can more easily modify their own property to fit their needs (Nakajima and Telyukova 2020). In most countries, the elderly run down their nonhousing wealth more quickly than their housing wealth (Nakajima and Telyukova 2020; Blundell et al. 2016).

There are other reasons why older individuals might liquidate their financial wealth before they liquidate their housing wealth. Most of these explanations center on the costs associated with selling a home or with tax-related issues (for example, Engen, Gale, and Uccello 1999). Liquidating a house entails substantial transaction costs. Most buyers and sellers use real estate agents, who typically charge 5–6 percent of the selling price of the house. These charges are in addition to the taxes and other fees associated with selling a house and the time and effort spent moving. Several papers suggest that households are sensitive to these transaction costs (Yang 2009; McGee 2022).

Moreover, housing is typically tax-advantaged relative to other assets. In the United States, housing can often be bequeathed to one's heirs tax-free, whereas selling a house will often force the seller to pay capital gains taxes. Furthermore, housing assets are often exempt from the asset tests associated with the Medicaid and Supplemental Security Insurance programs (De Nardi et al. 2012; Chang and Ko 2022). As a result, households that sell their home and convert the proceeds to financial assets become ineligible for these government transfers until the financial assets are depleted. Finally, income from financial assets is usually taxable, but the implicit "rent" homeowners pay themselves is untaxed.

Regardless of its cause, the desire to remain in one's own home will slow down the decumulation of total wealth only if there are impediments to extracting home equity while remaining in the house. With a "reverse mortgage," a homeowner can receive a stream of payments for as long as they live in the home, to be repaid from the later sale of the home. But as of 2011, only 2.1 percent of age-65+ homeowners had reverse mortgages (Nakajima and Telyukova 2017). The low take-up of reverse mortgages may reflect market frictions, such as difficulties by consumers in understanding these products. Alternatively, retirees may wish to hold on to their wealth for precautionary reasons or to leave bequests, reducing their willingness to borrow, in effect, against their homes.

To the extent that homeownership explains the slow rundown of wealth, its effects will be strongest among high-income households, who are more likely to own their home (Achou 2023). The homeownership motive is therefore consistent with the observation that those with high income are less likely to decumulate their assets.

Disentangling the Different Motivations

The three sets of saving motivations, precautionary, bequest, and housing, have similar implications for saving at older ages, making it difficult to disentangle their relative importance. All three motivations encourage saving, and all three motivations are strongest for the rich. Although we can estimate many of the risks facing households from the data, studies that attempt to quantify the competing hypotheses depend on preferences that are not observed. In particular, we need measures of risk aversion, patience, the strength of the bequest motive, the extent to which bequests are a luxury good, and the desire to remain in one's own home.

Numerical simulations of life-cycle models show that different values of these parameters can fit the observed asset data more or less equally well. For example, De Nardi, French, and Jones (2010) show that a model without bequests and with reasonable preference parameters and risks can match observed median wealth holdings by age, income quintile, and cohort, while still generating a realistic distribution of unintended bequests. But when the model is augmented to allow for intentional bequests, they estimate strong bequest motives, especially for the richest, with only modest changes in other parameters. The ability of such models

to fit wealth data almost equally well with or without bequest motives embodies the fundamental identification problem in this literature.

Disaggregating the data more finely, or considering savings over the entire life cycle, yields some evidence that bequest motives are at times important (Kaji, Manresa, and Pouliot 2020; Pashchenko and Porapakkarm 2023). Nonetheless, the broad message of the literature is that precautionary and bequest motives explain retiree wealth data equally well.

A number of papers attempt to resolve this problem by going beyond savings and considering additional features of the data. Here, we describe these approaches.

Insurance Choices

The life-cycle model with longevity and health risk, but without bequest motives, implies a high demand for insurance products such as annuities (insurance against a long life) and long-term care insurance (insurance against poor health at the end of life). These products, if fairly priced, can insure against lifespan or medical expense risk much more efficiently than standard assets. For example, using a simple version of the life-cycle model with only lifespan uncertainty, Yaari (1965) shows that people should immediately annuitize all their wealth. Nonetheless, US households hold only small amounts of annuities and long-term care insurance (Fang 2016).[4] This suggests that precautionary motives cannot be the only explanation for high savings at old ages.

Purchases of annuities and long-term care insurance reduce wealth left to heirs, but insure against medical and longevity risks. The fact that most households do not purchase these products is sometimes taken as evidence that people have a bequest motive (Lockwood 2018).

However, there may be other reasons why risk-averse households rarely purchase annuities or long-term care insurance. Many studies of the under-annuitization puzzle focus on adverse selection: long-lived people are more likely to purchase annuities, driving annuity prices up and pricing out those who do not expect to live so long (for a well-known paper in this vein, see Mitchell et al. 1999). But at observed levels of adverse selection, when the only risk facing households is lifespan uncertainty, most reasonably calibrated life-cycle models will imply that people should completely annuitize (for example, Lockwood 2012). On the other hand, many individuals seem to underestimate their expected lifespans and thus undervalue the returns to annuitization, substantially reducing annuity demand (O'Dea and Sturrock 2023).

Annuity demand may also be low because of medical expense risk. Annuities offer high returns to surviving individuals, but are very illiquid. This makes annuities more desirable to households who expect to live long and incur high medical expenses at very old ages, but less desirable to households who need liquid assets

---

[4] In an earlier symposium in this journal, Benartzi, Previtero, and Thaler (2011) and Brown and Finkelstein (2011) discuss in detail the low take-up of annuities and long-term care insurance, respectively.

to cover medical expenses in the near future (Davidoff, Brown, and Diamond 2005; Reichling and Smetters 2015; Peijnenburg, Nijman, and Werker 2017). Some studies that model health and medical spending risk carefully, however, still find that bequest motives are necessary to explain low annuity demand (Lockwood 2012; Pashchenko 2013).

Unlike annuities, which pay out benefits as long as the individual remains alive, long-term care insurance pays out only when the individual needs expensive long-term care services. In principle, the demand for long-term care insurance should be large, because long-term care needs often occur very late in life when other financial resources have been exhausted. In practice, access to comprehensive long-term insurance is often limited. The typical long-term care insurance contract caps both the maximum number of days covered over the life of the policy and the maximum daily payment for a nursing home stay, with the maximum often fixed in nominal terms (Fang 2016). In addition, many individuals have health conditions that preclude them from buying coverage: Hendren (2013) estimates that 23 percent of 65-year-olds fall into this category.

Suppliers of long-term care insurance face the risk that the insurance will lead households to switch from informal long-term care provided by family members to formal long-term care paid for by the insurer. This moral hazard problem not only drives up the cost of long-term care insurance, but it makes coverage unappealing to individuals who prefer to be cared for by their relatives (Pauly 1990; Mommaerts 2023). Ko (2022) finds that purchasers of long-term care insurance would be, even in the absence of insurance, more likely to utilize long-term formal care.

Middle- and low- income households may view themselves as reasonably insured against long-term care expenses by Medicaid. Because Medicaid is the "payer of last resort," and it covers only expenses not reimbursed by other insurers, among Medicaid recipients private long-term care insurance mostly displaces Medicaid payments—and thus Medicaid should crowd out private insurance. Brown and Finkelstein (2008) calculate that Medicaid imposes an implicit tax on private insurance of about 65 percent for the median-wealth individual. Braun, Kopecky, and Koreshkova (2019) likewise find that Medicaid crowd-out explains low holdings of long-term care insurance among poorer households, although adverse selection and administrative costs are more important in explaining low take-up among the rich.

The extent to which retirees run down their wealth to qualify for Medicaid and other means-tested benefits should tell us something about the value retirees place on these benefits. If people view Medicaid-funded care as being of low quality (they have "public care aversion" in the language of Ameriks et al. 2011; 2018), they will maintain high asset levels to avoid it, even though Medicaid care would be close to free. Thus, public care aversion strengthens precautionary saving motives. However, most low-income individuals receive Medicaid, suggesting that they are not overly averse to this insurance (De Nardi, French, and Jones 2016a). To match observed wealth holdings and Medicaid recipiency jointly, their model must attribute a significant part of savings to bequest motives.

Home equity may also substitute for long-term care insurance (and for annuities as well). Indeed, it has been shown that health shocks and loss of a spouse are associated with housing wealth decumulation (Poterba, Venti, and Wise 2011; Chang and Ko 2022). The extent to which home equity performs this function is unresolved (for conflicting results, see Davidoff 2010; Achou 2021). But the broader recurrent theme worth emphasizing is that postretirement assets can simultaneously serve many purposes and can be used for many contingencies.

In contrast to annuities and long-term care insurance, life insurance is widely held. Because (term) life insurance pays out only when its holder dies, its popularity has been taken as evidence for the existence of bequest motives (Inkmann and Michaelides 2012; Hong and Ríos-Rull 2012).

Finally, the limited use of reverse mortgages may suggest the presence of other saving motives. Estimating a structural model of saving and housing decisions, Nakajima and Telyukova (2017) find that bequest motives, nursing-home risk, house price risk, and loan costs all contribute to the low take-up of reverse mortgages. An alternative explanation is market frictions. Reverse mortgages often contain a requirement that homes be maintained, which may discourage their use by preventing home equity decumulation through foregone repairs (Cocco and Lopes 2020). Many reverse mortgages also impose a debt-to-income requirements, which older homeowners often fail (Caplin 2002). Information frictions and low levels of financial literacy may play a role as well (Davidoff, Gerhard, and Post 2017; Boyer et al. 2020).

It is worth stressing that the low use of these financial products does not imply a complete absence of precautionary motives, but only that there are other considerations leading households to self-insure through savings rather than insurance products. For example, households with modest bequest motives may prefer to insure against medical or longevity risks by holding assets that, when not spent, can be left to their heirs as incidental bequests.

Strategic Surveys

One way to find out why households are saving is to ask them. Ameriks et al. (2011, 2020) consider the responses to "strategic survey questions" that present the respondents with hypothetical, explicit trade-offs between consuming long-term care and leaving bequests. For example, Ameriks et al. (2011) ask survey respondents how they would divide a $100,000 (or $250,000) prize between a "bequest locked box" that would be given to the respondents' heirs when they died and an "long-term care locked box" that could be accessed only to pay for long-term care. Requiring the life-cycle model to match respondents' choices provides additional identifying variation that helps pin down the competing motivations. Their results suggest that for many older individuals, precautionary motives are at least as important as bequest motives.

Variation across Countries and Time

Although most countries have universal public health insurance programs for the elderly, considerable cross-country variation exists in the coverage of medical

and long-term care. If self-insuring against these expenses is a pressing concern, households should save more in countries with less public funding, all else equal.

Cross-country evidence supports the view that the precautionary saving motive is important. Institutional differences in health insurance generosity can explain one-third of the difference in retiree wealth decumulation between Sweden and the United States (Nakajima and Telyukova 2023) as well as differences in retiree spending patterns between England and the United States (Banks et al. 2019).

A related approach is to study the effects of policy changes over time. For example, Lee and Tan (2019) examine the effects of a calculation error in the Social Security benefit formula (made in the 1970s) that left retirees born between 1911 and 1916 with higher benefits than those born immediately before or after. They find that the benefit increase led to significantly higher bequests, which they interpret as evidence in favor of bequest motives.

Bequests and Children

If bequest motives are mostly due to parents' desire to leave resources to their offspring—as opposed to other relatives, friends, or charity—then households without children should have weaker bequest motives. The evidence on this question remains unsettled. The discrepancies stem from differences in approach.

One approach for identifying bequest motives is to compare wealth accumulation with and without children. Empirically, there is little evidence that retirees with living children decumulate their wealth at a slower rate than those without (Hurd 1987, 1989; De Nardi et al. 2023a; but see Kopczuk and Lupton 2007 for an alternative perspective). A second approach is to ask individuals about their bequest motives, using either stated preference information (Laitner and Juster 1996) or responses to strategic survey questions (Ameriks et al. 2011). These studies find that those with children tend to answer questions in a way consistent with stronger bequest motives.

In addition to the debate about whether those with children have stronger bequest motives, there is a debate about whether these bequest motives represent altruism or strategic considerations. Long-term care is often provided informally by children, especially in countries with limited public long-term care insurance (Barczyk and Kredler 2019). Retirees may accumulate funds for bequests (or inter vivos transfers) that will encourage their children to provide care. Bequests driven by the need to reward informal caregivers are known as *strategic bequests* (Bernheim, Shleifer, and Summers 1985). Strategic bequests share many similarities with precautionary saving. In both cases, households hold wealth late in life to insure against the risk of living long and having high medical needs.

Empirical evidence on the magnitude of the strategic bequest motive is mixed. Although many retirees receive care from their children, few pay for that care formally (Brown 2006). While written wills may reward caregivers with bequests (Groneck 2017), the additional transfers are typically modest and financial transfers from living parents do not favor caregivers (McGarry and Schoeni 1997). Studies estimating models that include altruistic and strategic motives find that strategic

motives alone cannot explain transfer behavior (Barczyk and Kredler 2018; Ko 2022; Barczyk, Fahle, and Kredler forthcoming; Mommaerts 2023). In summary, the evidence for an operative strategic bequest motive is modest. To the extent bequests are intentional, rather than accidental outcomes of the precautionary motive, they appear to be largely altruistic.

Taking Stock

A number of recent studies, exploiting different features of the data, suggest that both precautionary and bequest motives are present. However, the relative importance of these motives remains an open question. Research based on demand for annuities and long-term care insurance tends to find stronger bequest motives. Papers utilizing strategic survey questions tend to find a larger role for precautionary motives. The slower decumulation rates of homeowners imply that the desire to remain in one's own home is also important, but the limited use of reverse mortgages suggests that it cannot be the only motive present.

Different motives likely dominate at different points of the income distribution, reflecting differences in the extent to which each motive behaves as a luxury good. For example, De Nardi et al. (2023a) find that precautionary motives are dominant in the middle tercile of the income distribution, while bequest motives play a larger role at the top.

In addition to disentangling saving motives, these studies raise pointed questions about the effectiveness of several financial products and the scope for government intervention. If households have strong precautionary motives but are deterred by market imperfections such as adverse selection, there may be a role for policy to improve these products. On the other hand, if strong bequest motives are limiting the demand for these products, their low utilization may be efficient.

Savings Trends and Policy Implications

Nearly 40 percent of total nonpension wealth in the United States is held by households whose heads are 65 or older (Bhutta et al. 2020). As the population continues to age, the importance of retiree savings will only increase. Concerns about low and declining savings rates earlier in life have led some to believe that younger cohorts may be unprepared for retirement (Skinner 2007). Although it is difficult to know how future generations will accumulate wealth during their working years or how they will draw down this wealth once they retire, comparing savings patterns across cohorts may provide some clues. Using the same data from the Health and Retirement Survey and wealth measure as Figure 1, Figure 4 plots the median wealth of four cohorts, each born in a different decade, against the average age of the cohorts' members, for the years 1998–2018.

To fix ideas, note that in 1998, the War Babies cohort has an average age of 55 and median wealth holdings of roughly $150,000, while the Late–Greatest

Figure 4
Median Wealth of Successive Cohorts
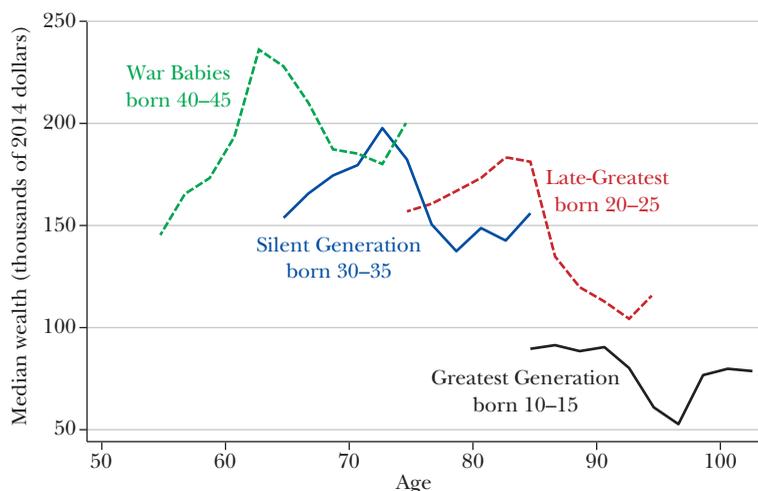
Source: Authors' calculations from the Health and Retirement Survey.
Note: Each line shows median wealth for a particular 5-year birth cohort over the years 1998–2018.

Generation cohort has an average age of 75 and holds a similar amount of wealth. In 2000, the War Babies and Late-Greatest cohorts are 57 and 77 years old, respectively, and both hold more wealth. Except for the oldest old (the "Greatest Generation"), the profiles have similar shapes across the cohorts. Wealth rises between 1998 and 2006 before falling, reflecting the rise and fall of asset prices around the Great Recession. For those in their 50s and 60s, there is some evidence of wealth accumulation, while there is evidence of decumulation after age 70. These dynamics aside, Figure 4 shows that cohorts born more recently hold more wealth.

On the other hand, among cohorts younger than those shown in Figure 4, wealth accumulation has stalled (Gale et al. 2021). Sturrock (2023) reports a similar stalling in UK data, attributing much of the slowdown to lower earnings growth. Such findings raise the possibility that younger generations could enter retirement with less wealth than older ones.

The well-being of retired households depends not only on their income and wealth, but also on their exposure to the risks of outliving their wealth or incurring expensive medical conditions. Even as wealth accumulation has halted, longevity and medical expenses have continued to grow, raising questions about how future generations will fund their retirements. Between 1950 and 2019, the United States saw a steady increase in life expectancy at age 65, rising from 13.9 to 19.6 years.[5]

[5]Compiled from National Center for Health Statistics (2012, Table 22; 2018, Table 15), Arias and Xu (2022a, 2022b, Table A), and Arias et al. (2022). Estimates for 2021 are provisional.

Since 2020, COVID-19 has reduced life expectancy by over a year, although perhaps only temporarily. It remains unclear how the pandemic and other health trends, such the rapid growth of obesity or the rise in "deaths of despair" (Case and Deaton 2021), will impact retirees' lifespans or the ages at which they stop working. Changes along either dimension will affect the number of years that retirees need to fund.

Although exact growth rates are hard to predict, lifespans and medical expenditures will most likely continue to rise (for example, see forecasts by the Social Security Administration 2022; Centers for Medicare and Medicaid Services 2022a) as new medical technologies are adopted (Chandra, Holmes, and Skinner 2013). Given that households care most about what they pay out-of-pocket, changes in the generosity of health insurance will also be important. In recent decades, expansions in coverage have held down out-of-pocket spending despite substantial increases in total spending.[6]

If current trends continue, however, the cost of government programs such as Social Security, Medicare, and Medicaid will not be sustainable. Reforms are needed, but their impact will depend on the relative strengths of the precautionary, bequest, and homeownership motives. To give a prominent example, it has been long understood that if households do not face risk and do not possess altruistic bequest motives, an unfunded Social Security system crowds out private savings, reduces the aggregate capital stock, and likely reduces welfare (Diamond 1965). On the other hand, altruistic bequest motives can undo many of these distortions (Barro 1974), and precautionary motives may allow Social Security to have insurance value and perhaps even improve welfare (Harenberg and Ludwig 2019).

Understanding saving motivations is also important when considering how to insure retirees against risks such as long-term care. Given the low take-up of private insurance, many retirees face the possibility of catastrophic long-term care expenses. Whether and how to reform the long-term care insurance market (through promoting private insurance or expanding Medicaid) has been a topic of recent policy debate (Commission on Long-Term Care 2013). When considering such reforms, policymakers need to know the extent to which the limited use of private insurance reflects market frictions, rather than bequest and/or homeownership motives that lower the insurance's value (Arapakis et al. 2022).

In short, we expect the retirement savings puzzle to only increase in salience. We hope that future cohorts of economists will continue to make it a research priority.

---

[6]Between 1990 and 2020, even as per capita medical spending rose by 113 percent, out-of-pocket spending rose by only 13 percent. Among those 65 and older, between 2002 and 2014 (when data are available), per capita out-of-pocket spending fell by 5 percent, even as total spending increased by 7 percent. Figures calculated from the National Health Expenditure Accounts. Data for all ages come from the main tables (Centers for Medicare and Medicaid Services 2022c, Tables 1 and 6), and data for older individuals come from the Age and Gender tables (Centers for Medicare and Medicaid Services 2022b, Table 7). All values deflated by the Consumer Price Index.

References

Achou, Bertrand. 2021. "Housing Liquidity and Long-Term Care Insurance Demand: A Quantitative Evaluation." *Journal of Public Economics* 194: 104353.

Achou, Bertrand. 2023. "Housing in Medicaid: Should it Really Change?" *American Economic Journal: Economic Policy* 15 (1): 1–36.

Ameriks, John, Andrew Caplin, Steven Laufer, and Stijn Van Nieuwerburgh. 2011. "The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives." *Journal of Finance* 66 (2): 519–61.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti. 2018. "The Long-Term-Care Insurance Puzzle: Modeling and Measurement." NBER Working Paper 22726.

Ameriks, John, Joseph Briggs, Andrew Caplin, Matthew D. Shapiro, and Christopher Tonetti. 2020. "Long-Term Care Utility and Late-in-Life Saving." *Journal of Political Economy* 128 (6): 2375–2451.

Arapakis, Karolos, Eric French, John Bailey Jones, and Jeremy McCauley. 2021. "On the Distribution and Dynamics of Medical Expenditure among the Elderly." Michigan Retirement and Disability Research Center Working Paper 2021–436.

Arapakis, Karolos, Eric French, John Bailey Jones, and Jeremy McCauley. 2022. "How Should We Fund End-of-Life Care in the US?" *The Lancet Regional Health—Americas* 15: 100359.

Arias, Elizabeth, Betzaida Tejada-Vera, Kenneth D. Kochanek, and Farida B. Ahmad. 2022. *Provisional Life Expectancy Estimates for 2021.* Vital Statistics Rapid Release Report 23. Hyattsville, MD: National Center for Health Statistics.

Arias, Elizabeth, and Jiaquan Xu. 2022a. *United States Life Tables, 2019.* National Vital Statistics Report 19. Vol. 70. Hyattsville, MD: National Center for Health Statistics.

Arias, Elizabeth, and Jiaquan Xu. 2022b. *United States Life Tables, 2020.* National Vital Statistics Report 2. Vol 71. Hyattsville, MD: National Center for Health Statistics.

Banks, James, Richard Blundell, Peter Levell, and James P. Smith. 2019. "Life-Cycle Consumption Patterns at Older Ages in the United States and the United Kingdom: Can Medical Expenditures Explain the Difference?" *American Economic Journal: Economic Policy* 11 (3): 27–54.

Banks, James, Janet Currie, Sonya Krutikova, Kjell G. Salvanes, and Hannes Schwandt. 2021. "The Evolution of Mortality Inequality in 11 OECD Countries: Introduction." *Fiscal Studies* 42 (1): 9–23.

Barczyk, Daniel, Sean Fahle, and Matthias Kredler. Forthcoming. "Save, Spend or Give? A Model of Housing, Family Insurance, and Savings in Old Age." *Review of Economic Studies.* https://doi.org/10.1093/restud/rdac081.

Barczyk, Daniel, and Matthias Kredler. 2018. "Evaluating Long-Term-Care Policy Options, Taking the Family Seriously." *Review of Economic Studies* 85 (2): 766–809.

Barczyk, Daniel, and Matthias Kredler. 2019. "Long-Term Care across Europe and the United States: The Role of Informal and Formal Care." *Fiscal Studies* 40 (3): 329–73.

Barro, Robert J. 1974. "Are Government Bonds Net Wealth?" *Journal of Political Economy* 82 (6): 1095–1117.

Benartzi, Shlomo, Alessandro Previtero, and Richard H. Thaler. 2011. "Annuitization Puzzles." *Journal of Economic Perspectives* 25 (4): 143–64.

Bernheim, B. Douglas, Andrei Shleifer, and Lawrence H. Summers. 1985. "The Strategic Bequest Motive." *Journal of Political Economy* 93 (6): 1045–76.

Bhutta, Neil, Jesse Bricker, Andrew C. Chang, Lisa J. Dettling, Sarena Goodman, Joanne W. Hsu, Kevin B. Moore, Sarah Reber, Alice Henriques Volz, and Richard A. Windle. 2020. "Changes in U.S. Family Finances from 2016 to 2019: Evidence from the Survey of Consumer Finances." *Federal Reserve Bulletin* 106 (5): 1–42.

Blundell, Richard, Margherita Borella, Jeanne Commault, and Mariacristina De Nardi. 2020. "Why Does Consumption Fluctuate in Old Age and How Should the Government Insure It?" NBER Working Paper 27348.

Blundell, Richard, Rowena Crawford, Eric French, and Gemma Tetlow. 2016. "Comparing Retirement Wealth Trajectories on Both Sides of the Pond." *Fiscal Studies* 37 (1): 105–30.

Boyer, M. Martin, Philippe De Donder, Claude Fluet, Marie-Louise Leroux, and Pierre-Carl Michaud. 2020. "Long-Term Care Insurance: Information Frictions and Selection." *American Economic Journal: Economic Policy* 12 (3): 134–69.

Braun, R. Anton, Karen A. Kopecky, and Tatyana Koreshkova. 2017. "Old, Sick, Alone, and Poor: A Welfare Analysis of Old-Age Social Insurance Programmes." *Review of Economic Studies* 84 (2): 580–612.

Braun, R. Anton, Karen A. Kopecky, and Tatyana Koreshkova. 2019. "Old, Frail, and Uninsured: Accounting for Features of the U.S. Long-Term Care Insurance Market." *Econometrica* 87 (3): 981–1019.

Brown, Jeffrey R., and Amy Finkelstein. 2008. "The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market." *American Economic Review* 98 (3): 1083–1102.

Brown, Jeffrey R., and Amy Finkelstein. 2011. "Insuring Long-Term Care in the United States." *Journal of Economic Perspectives* 25 (4): 119–42.

Brown, Meta. 2006. "Informal Care and the Division of End-of-Life Transfers." *Journal of Human Resources* 41 (1): 191–219.

Caplin, Andrew. 2002. "Turning Assets into Cash: Problems and Prospects in the Reverse Mortgage Market." In *Innovations in Retirement Financing*, edited by Olivia S. Mitchell, Zvi Bodie, P. Brett Hammond, and Stephen Zeldes, 234–53. Philadelphia, PA: University of Pennsylvania Press.

Case, Anne, and Angus Deaton. 2021. *Deaths of Despair and the Future of Capitalism*. Princeton, NJ: Princeton University Press.

Castañeda, Ana, Javier Díaz-Giménez, and José-Víctor Ríos-Rull. 2003. "Accounting for the U.S. Earnings and Wealth Inequality." *Journal of Political Economy* 111 (4): 818–57.

Centers for Medicare and Medicaid Services. 2022a. *2022 Annual Report of the Boards of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds*. Washington, DC: Centers for Medicare and Medicaid Services.

Centers for Medicare and Medicaid Services. 2022b. "National Health Expenditure Data: Age and Gender." Washington, DC: Centers for Medicare and Medicaid Services.

Centers for Medicare and Medicaid Services. 2022c. "National Health Expenditure Data: Historical." Washington, DC: Centers for Medicare and Medicaid Services.

Chandra, Amitabh, Jonathan Holmes, and Jonathan Skinner. 2013 "Is This Time Different? The Slowdown in Health Care Spending." *Brookings Papers on Economic Activity* 44 (2): 261–302.

Chang, Minsu, and Ami Ko. 2022. "Marital Transitions, Housing, and Long-Term Care in Old Age." Unpublished.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. "The Association between Income and Life Expectancy in the United States, 2001–2014." *JAMA* 315 (16): 1750–66.

Christensen, Bent Jesper, Malene Kallestrup-Lamb, and John Kennan. 2022. "Consumption and Saving after Retirement." NBER Working Paper 29826.

Cocco, João F., and Paula Lopes. 2020. "Aging in Place, Housing Maintenance, and Reverse Mortgages." *Review of Economic Studies* 87 (4): 1799–1836.

Commission on Long-Term Care. 2013. *Commission on Long-Term Care Report to the Congress*. Washington,

DC: Government Printing Office.

Davidoff, Thomas. 2010. "Home Equity Commitment and Long-Term Care Insurance Demand." *Journal of Public Economics* 94 (1–2): 44–49.

Davidoff, Thomas, Jeffrey R. Brown, and Peter A. Diamond. 2005. "Annuities and Individual Welfare." *American Economic Review* 95 (5): 1573–90.

Davidoff, Thomas, Patrick Gerhard, and Thomas Post. 2017. "Reverse Mortgages: What Homeowners (Don't) Know and How It Matters." *Journal of Economic Behavior and Organization* 133: 151–71.

Davies, James B. 1981. "Uncertain Lifetime, Consumption, and Dissaving in Retirement." *Journal of Political Economy* 89 (3): 561–77.

Diamond, Peter A. 1965. "National Debt in a Neoclassical Growth Model." *American Economic Review* 55 (5): 1126–50.

De Nardi, Mariacristina. 2004. "Wealth Inequality and Intergenerational Links." *Review of Economic Studies* 71 (3): 743–68.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2009. "Life Expectancy and Old Age Savings." *American Economic Review* 99 (2): 110–15.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2010. "Why Do the Elderly Save? The Role of Medical Expenses." *Journal of Political Economy* 118 (1): 39–75.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2016a. "Medicaid Insurance in Old Age." *American Economic Review* 106 (11): 3480–3520.

De Nardi, Mariacristina, Eric French, and John Bailey Jones. 2016b. "Savings after Retirement: A Survey." *Annual Review of Economics* 8: 177–204.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Angshuman Gooptu. 2012. "Medicaid and the Elderly." *Federal Reserve Bank of Chicago Economic Perspectives* 36 (1): 17–34.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Jeremy McCauley. 2016. "Medical Spending of the US Elderly." *Fiscal Studies* 37 (3–4): 717–47.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Rory McGee. 2023a. "Why Do Couples and Singles Save during Retirement? Household Heterogeneity and Its Aggregate Implications." NBER Working Paper 28828.

De Nardi, Mariacristina, Eric French, John Bailey Jones, and Rory McGee. 2023b. "Why Do Couples and Singles Save during Retirement? Household Differences and their Aggregate Implications: Data Underlying Figures 3 and 5." NBER Working Paper 28828 (accessed October 28, 2022).

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes. 2002. "The Importance of Bequests and Life-Cycle Saving in Capital Accumulation: A New Answer." *American Economic Review* 92 (2): 274–78.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes. 2004. "Do the Rich Save More?" *Journal of Political Economy* 112 (2): 397–444.

Engen, Eric M., William G. Gale, and Cori E. Uccello. 1999. "The Adequacy of Household Saving." *Brookings Papers on Economic Activity* 30 (2): 65–165.

Fang, Hanming. 2016. "Insurance Markets for the Elderly." In *Handbook of the Economics of Population Aging*, Vol. 1, edited by John Piggott and Alan Woodland, 237–309. Amsterdam: North-Holland.

French, Eric, John Bailey Jones, and Rory McGee. 2023. "Replication data for: Why Do Retired Households Draw Down Their Wealth So Slowly?" American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193226V1.

French, Eric, Mariacristina De Nardi, John Bailey Jones, Olesya Baker, and Phil Doctor. 2006. "Right before the End: Asset Decumulation at the End of Life." *Federal Reserve Bank of Chicago Economic Perspectives* 30 (3): 2–13.

Gale, William G., and John Karl Scholz. 1994. "Intergenerational Transfers and the Accumulation of Wealth." *Journal of Economic Perspectives* 8 (4): 145–60.

Gale, William G., Hilary Gelfond, Jason J. Fichtner, and Benjamin H. Harris. 2021. "The Wealth of Generations, with Special Attention to the Millennials." In *Measuring Distribution and Mobility of Income and Wealth*, edited by Raj Chetty, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell, 145–74. Chicago: University of Chicago Press.

Gardner, Lara, and Donna B. Gilleskie. 2012. "The Effects of State Medicaid Policies on the Dynamic Savings Patterns and Medicaid Enrollment of the Elderly." *Journal of Human Resources* 47 (4): 1082–1127.

Greenhalgh-Stanley, Nadia. 2012. "Medicaid and the Housing and Asset Decisions of the Elderly: Evidence from Estate Recovery Programs." *Journal of Urban Economics* 72 (2–3): 210–24.

Groneck, Max. 2017. "Bequests and Informal Long-Term Care Evidence from HRS Exit Interviews."

Journal of Human Resources 52 (2): 531–72.

Harenberg, Daniel, and Alexander Ludwig. 2019. "Idiosyncratic Risk, Aggregate Risk, and the Welfare Effects of Social Security." *International Economic Review* 60 (2): 661–92.

Hendren, Nathaniel. 2013. "Private Information and Insurance Rejections." *Econometrica* 81 (5): 1713–62.

Hong, Jay H., and José-Víctor Ríos-Rull. 2012. "Life Insurance and Household Consumption." *American Economic Review* 102 (7): 3701–30.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *Journal of Political Economy* 103 (2): 360–99.

Hurd, Michael D. 1987. "Savings of the Elderly and Desired Bequests." *American Economic Review* 77 (3): 298–312.

Hurd, Michael D. 1989. "Mortality Risk and Bequests." *Econometrica* 57 (4): 779–813.

Hurst, Erik, and James P. Ziliak. 2006. "Do Welfare Asset Limits Affect Household Saving? Evidence from Welfare Reform." *Journal of Human Resources* 41 (1): 46–71.

Inkmann, Joachim, and Alexander Michaelides. 2012. "Can the Life Insurance Market Provide Evidence for a Bequest Motive?" *Journal of Risk and Insurance* 79 (3): 671–95.

Kaji, Tetsuya, Elena Manresa, and Guillaume Pouliot. 2020. "An Adversarial Approach to Structural Estimation." arXiv: 2007.06169.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money.* New York: MacMillan.

Ko, Ami. 2022. "An Equilibrium Analysis of the Long-Term Care Insurance Market." *Review of Economic Studies* 89 (4): 1993–2025.

Kopczuk, Wojciech, and Joseph P. Lupton. 2007. "To Leave or Not to Leave: The Distribution of Bequest Motives." *Review of Economic Studies* 74 (1): 207–35.

Kopecky, Karen A., and Tatyana Koreshkova. 2014. "The Impact of Medical and Nursing Home Expenses on Savings." *American Economic Journal: Macroeconomics* 6 (3): 29–72.

Kotlikoff, Laurence J., and Lawrence H. Summers. 1981. "The Role of Intergenerational Transfers in Aggregate Capital Accumulation." *Journal of Political Economy* 89 (4): 706–32.

Laitner, John, and F. Thomas Juster. 1996. "New Evidence on Altruism: A Study of TIAA-CREF Retirees." *American Economic Review* 86 (4): 893–908.

Laitner, John, Dan Silverman, and Dmitriy Stolyarov. 2018. "The Role of Annuitized Wealth in Post-retirement Behavior." *American Economic Journal: Macroeconomics* 10 (3): 71–117.

Lee, Siha, and Kegon T. K. Tan. 2019. "Bequest Motives and the Social Security Notch." HCEO Working Paper 2019-061.

Lockwood, Lee M. 2012. "Bequest Motives and the Annuity Puzzle." *Review of Economic Dynamics* 15 (2): 226–43.

Lockwood, Lee M. 2018. "Incidental Bequests and the Choice to Self-Insure Late-Life Risks." *American Economic Review* 108 (9): 2513–50.

Love, David A., Michael G. Palumbo, and Paul A. Smith. 2009. "The Trajectory of Wealth in Retirement." *Journal of Public Economics* 93 (1–2): 191–208.

Marshall, Samuel, Kathleen M. McGarry, and Jonathan S. Skinner. 2011. "The Risk of Out-of-Pocket Health Care Expenditure at the End of Life." In *Explorations in the Economics of Aging*, edited by David Wise, 101–28. Chicago: University of Chicago Press.

McGarry, Kathleen, and Robert F. Schoeni. 1997. "Transfer Behavior within the Family: Results from the Asset and Health Dynamics Study." *Journals of Gerontology Series B* 52B (S): S82–92.

McGee, Rory. 2022. "Old Age Savings and House Price Shocks." University of Western Ontario Research Report 2021-4.

Mitchell, Olivia S., James M. Poterba, Mark J. Warshawsky, and Jeffrey R. Brown. 1999. "New Evidence on the Money's Worth of Individual Annuities." *American Economic Review* 89 (5): 1299–1318.

Modigliani, Franco. 1988. "Measuring the Contribution of Intergenerational Transfers to Total Wealth: Conceptual Issues and Empirical Findings," In *Modeling the Accumulation and Distribution of Wealth*, edited by Denis Kessler and André Masson, 21–52. Oxford, UK: Clarendon Press, Oxford.

Modigliani, Franco, and Richard Brumberg. 1980. "Utility Analysis and Aggregate Consumption Functions: An Attempt at Integration," In *The Collected Papers of Franco Modigliani*, Vol. 2, *The Life Cycle Hypothesis of Saving*, edited by Andrew Abel, 128–97. Cambridge, MA: MIT Press.

Modigliani, Franco, and Richard Brumberg. 1954. "Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data." In *The Collected Papers of Franco Modigliani*, Vol. 6, edited by Francesco Franco, 3–45. Cambridge, MA: MIT Press, 2005.

Mommaerts, Corina. 2023. "Long-Term Care Insurance and the Family." Unpublished. https://sites.

google.com/site/corinamommaerts/research.

Nakajima, Makoto, and Irina A. Telyukova. 2017. "Reverse Mortgage Loans: A Quantitative Analysis." *Journal of Finance* 72 (2): 911–50.

Nakajima, Makoto, and Irina A. Telyukova. 2020. "Home Equity in Retirement." *International Economic Review* 61 (2): 573–616.

Nakajima, Makoto, and Irina A. Telyukova. 2023. "Medical Expenses and Saving in Retirement: The Case of U.S. and Sweden." Unpublished.

National Center for Health Statistics. 2012. *Health, United States, 2011: With Special Feature on Socioeconomic Status and Health.* Hyattsville, MD: National Center for Health Statistics.

National Center for Health Statistics. 2018. *Health, United States, 2017: With Special Feature on Mortality.* Hyattsville, MD: National Center for Health Statistics.

O'Dea, Cormac, and David Sturrock. 2023. "Survival Pessimism and the Demand for Annuities." *Review of Economics and Statistics* 105 (2): 442–57.

Pashchenko, Svetlana. 2013. "Accounting for Non-annuitization." *Journal of Public Economics* 98: 53–67.

Pashchenko, Svetlana, and Ponpoje Porapakkarm. 2023. "Saving Motives over the Life-Cycle." Unpublished.

Pauly, Mark V. 1990. "The Rational Nonpurchase of Long-Term-Care Insurance." *Journal of Political Economy* 98 (1): 153–68.

Peijnenburg, Kim, Theo Nijman, and Bas J. M. Werker. 2017. "Health Cost Risk: A Potential Solution to the Annuity Puzzle." *Economic Journal* 127 (603): 1598–1625.

Poterba, James, Steven Venti, and David Wise. 2011. "The Composition and Drawdown of Wealth in Retirement." *Journal of Economic Perspectives* 25 (4): 95–117.

Reichling, Felix, and Kent Smetters. 2015. "Optimal Annuitization with Stochastic Mortality and Correlated Medical Costs." *American Economic Review* 105 (11): 3273–3320.

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving 'Optimally' for Retirement?" *Journal of Political Economy* 114 (4): 607–43.

Shorrocks, Antony F. 1975. "The Age-Wealth Relationship: A Cross-Section and Cohort Analysis." *Review of Economics and Statistics* 57 (2): 155–63.

Skinner, Jonathan. 2007. "Are You Sure You're Saving Enough for Retirement?" *Journal of Economic Perspectives* 21 (3): 59–80.

Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation between Health and Economic Status." *Journal of Economic Perspectives* 13 (2): 145–66.

Social Security Administration. 2022. *The 2022 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Federal Disability Insurance Trust Funds.* Washington, DC: Social Security Administration.

Sturrock, David. 2023. "Wealth and Welfare across Generations." Institute for Fiscal Studies Working Paper 23/15.

Waldron, Hilary. 2007. "Trends in Mortality Differentials and Life Expectancy for Male Social Security-Covered Workers, by Socioeconomic Status." *Social Security Bulletin* 67 (3): 1–28.

Yaari, Menahem E. 1965. "Uncertain Lifetime, Life Insurance, and Theory of the Consumer." *Review of Economic Studies* 32 (2): 137–50.

Yang, Fang. 2009. "Consumption over the Life Cycle: How Different Is Housing?" *Review of Economic Dynamics* 12 (3): 423–43.

Where Does Wealth Come From? Measuring Lifetime Resources in Norway

Sandra E. Black, Paul J. Devereux, Fanny Landaud, and Kjell G. Salvanes

The United States and many other countries have experienced dramatic increases in wealth inequality in recent decades. As panel A of Figure 1 shows, the share of wealth held by the top 10 percent of the population ranges from 0.48 (Netherlands) to 0.80 (Chile) across OECD countries, highlighting dramatic disparities in wealth within countries. In addition, this is part of a growing trend: when we look at the change of the top 10 percent share between 1995 and 2021 (Figure 1, panel B), we see that, for most OECD countries, inequality has

■ *Sandra E. Black is Professor of Economics and International and Public Affairs at Columbia University, New York City, New York. She is also an Affiliated Professor, Norwegian School of Economics, Bergen, Norway; Research Fellow, IZA (Institute of Labor Economics), Berlin, Germany; and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. Paul J. Devereux is Professor of Economics at University College Dublin, Dublin, Ireland, NHH, IZA, and CEPR. He is also an Affiliated Professor, Norwegian School of Economics, Bergen, Norway; Research Fellow, IZA (Institute of Labor Economics), Berlin, Germany; and Fellow, Centre for Economic Policy Research, London, United Kingdom. Fanny Landaud is a CNRS (Centre national de la recherche scientifique) Researcher at THEMA (Théorie Économique, Modélisation, et Applications), CY Cergy Paris University, Cergy-Pontoise, France. Kjell G. Salvanes is Professor of Economics, Norwegian School of Economics, Bergen, Norway. He is also Research Associate, IZA (Institute of Labor Economics), Berlin, Germany; Research Associate, Center for Economic Policy Research, London, United Kingdom; and Research Associate, Center for the Economics of Human Development, University of Chicago, Chicago, Illinois. Their email addresses are sblack@columbia.edu, devereux@ucd.ie, fanny.landaud@cyu.fr, kjell.salvanes@nhh.no.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.37.4.115.
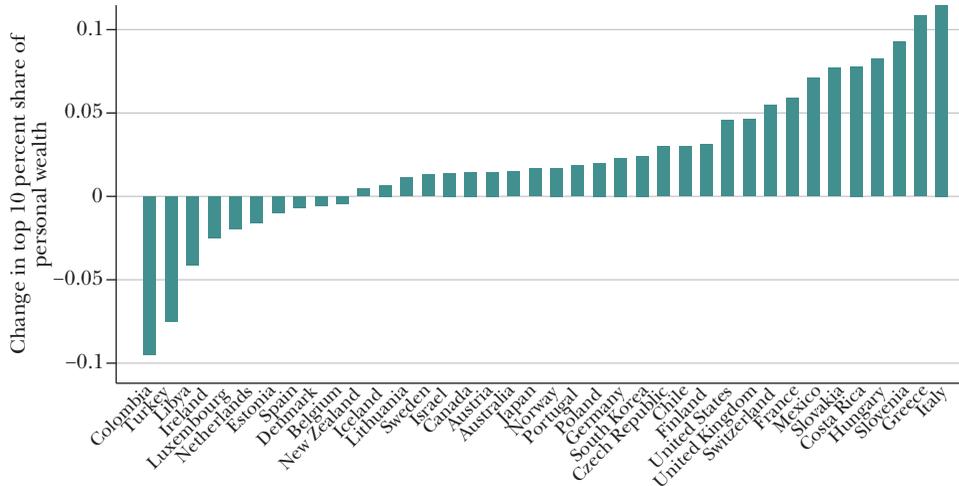
Figure 1

Top 10 Percent Share and Change in Top 10 Percent Share of Net Personal Wealth for OECD Member Countries

Panel A. Top 10 percent share in 2021
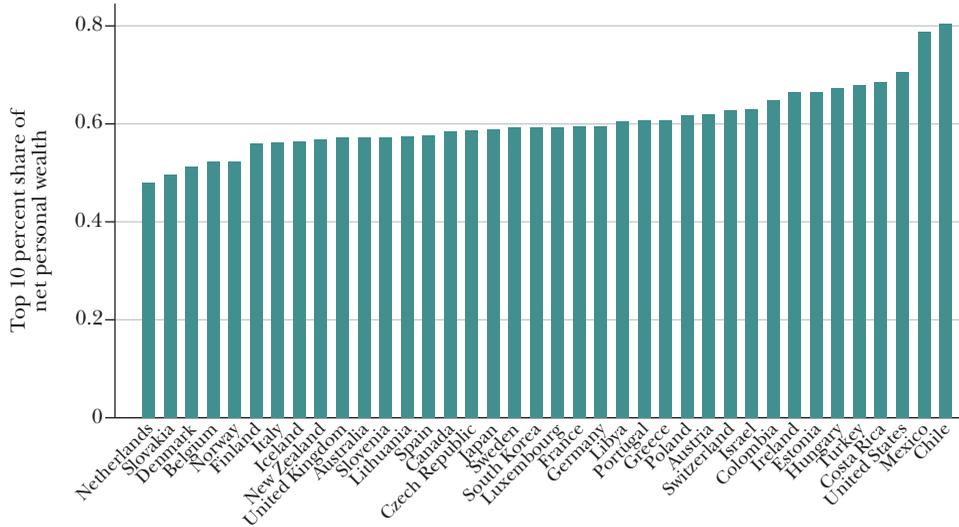
Panel B. Change in top 10 percent share from 1995 to 2021

Source: The figure uses information from the World Inequality Database (WID.world 2023).
Note: Panel A shows the share of net personal wealth held by the top 10 percent in 2021 for each OECD country, and panel B shows the change in the share of net personal wealth held by the top 10 percent between 1995 and 2021.

increased; the increase in the top 10 percent share ranges from –0.10 (Columbia) to 0.11 (Italy). Wealth inequality is particularly salient, given that it is both substantially larger than income inequality and is much more persistent across generations

(for US-centered discussions, see, for example, Kopczuk and Saez 2004; Saez and Zucman 2016, 2020a, 2020b; Smith et al. 2019).

However, wealth is a particularly problematic outcome to interpret—more so than education or even income. Unlike income, wealth is a stock. It results both from actions to increase the flow of money to a household (through, for example, income from labor) as well as from decisions by the household on how to use the money that comes in (for example, choices about consumption, saving, and assets in which to invest). As a result, individuals may have similar wealth but from very different sources; for instance, some may have inherited their wealth, while others earned it through working and saving. Moreover, individuals with the same income may ultimately have different wealth because of differing consumption or investment patterns. If two individuals had the same wealth at age 65 but one had a large inheritance from parents and resulting capital gains from investments, while the second received no inheritance but worked hard and saved their labor income, would we say that this is an equal society? Similarly, if two individuals had the same level and sources of income but one consumed much more, so at age 65 they had very different net wealth, would we say this is an unequal society?

To better understand inequality of wealth, we might ideally like to abstract from differences in consumption and spending behavior and compare the opportunity to accumulate wealth across individuals. To carry out such a task, we would need a large sample of households for which we have information on levels of household wealth as well as income and capital gains received from all sources over a sustained period of time. Few countries have the data available to carry out such a calculation, but the comprehensive Norwegian administrative data allow us to carry out an accounting exercise in which we sum all the income flows into a household over an extended period.

More specifically, we use data on the population of Norway to create a measure of *lifetime resources*, calculated as net wealth in 1994 plus the sum of income and capital gains received from all sources from 1995 to 2013. Once we have this measure of lifetime resources for each individual, we can answer several questions: (1) How correlated is lifetime resources with net wealth? (2) What are the most important sources of lifetime resources, and how does this vary by age and across the distribution of lifetime resources? (3) How do measures of inequality change when we use lifetime resources relative to wealth or labor income? The concept of measuring lifetime resources is not new; for example, Boserup, Kopczuk, and Kreiner (2017) also define a measure of lifetime resources as a more appropriate metric than net wealth when studying intergenerational transmission, but they do not calculate lifetime resources.

Our goal is to think broadly about the determinants of lifetime resources and to examine the relative importance of various sources of income as people move from youth to middle age. While analyses of wealth have generally either used aggregate data from national accounts or cross-sectional data from wealth surveys, our data allow us to observe wealth for a balanced population panel over a 19-year period. Our analysis is also potentially related to ongoing policy debates around the best

approaches to reducing wealth inequality. For example, if inheritances constitute a large proportion of the income of the wealthy, this may be an argument for higher inheritance taxes to counter inequality. If, on the other hand, most of the income of the wealthy comes from capital gains, these may instead be a better target for tax policy.

Norwegian data are particularly well suited to our analysis. Norway has a wealth tax that is assessed annually and is based on net wealth including financial assets, housing wealth, cars, and bank deposits. Until it was eliminated in January 2014, Norway also had a unified inheritance and gift tax. Therefore, until 2013, there are accurate administrative data on wealth and inheritances, data that are unavailable in most countries.

While the Norwegian data are uniquely suited to this analysis, how generalizable are conclusions from the Norwegian data likely to be? As shown in Figure 1, Norway displays a similar top 10 percent share of wealth to that in many other countries but is on the more equal side of the scale; also, like most countries, Norway experienced a modest rise in wealth inequality from 1995 to 2013. Moving beyond the top 10 percent share, when we compare the distribution of wealth in our data to that for other countries in the Household Finance and Consumption Survey (HFCS), while Norway looks somewhat different at the very bottom and very top of the distribution—which is consistent with the fact that our data are likely better than the survey data at the extremes of the distribution—the overall distributions look quite similar across countries.[1] Using the Gini coefficient as an overall measure of wealth inequality, the Gini coefficient for Norway at 0.68 is very similar to that for the Euro area as a whole (0.69). In the conclusion, we will discuss further how generalizable our findings are likely to be to other countries such as the United States.

Different Measure of Wealth: Lifetime Resources

As noted earlier, we would ideally like to abstract from differences in consumption and spending behavior and compare the opportunity to accumulate wealth across individuals. While we cannot observe this counterfactual, we proxy for it with a simple accounting exercise summing all the income flows into a household over an extended period of time (1995 to 2013).

Net wealth equals gross wealth (the total value of real assets such as housing, financial assets including cash deposits and stocks, and any other wealth held) minus total debts (which include mortgage debts, student loans, and other debts). By definition, the net wealth at the end of any period equals the net wealth at the beginning plus all income inflows during the period minus any outflows during the

---

[1]Appendix Table A1 shows deciles of the distribution of household net wealth in 2014, calculated for Norway using the Norwegian administrative data and for other countries using survey data (Household Finance and Consumption Network 2020). Unfortunately, Norway is not included in the Household Finance and Consumption Survey.

period. The income inflows include net labor income, net government transfers, net capital income, net inheritances and gifts received, lottery gains, net capital gains on real assets, and net capital gains on financial assets.[2] Income outflows are the wealth tax paid and consumption (while income taxes are an outflow they are already accounted for in our measure of "net" labor income). Because capital gains add to wealth, they are included even if they are not realized. We consider the 20 years from 1994 up through 2013. As a result, net wealth as measured in 2013 can be considered as the total income inflows and gains (net of taxes) from all sources between 1995 and 2013 minus the cumulated wealth tax paid and cumulated consumption, added to the baseline measure of net wealth in 1994. Our measure of lifetime resources abstracts from consumption decisions and so equals net wealth in 1994 combined with the sum of the income inflows over the following 19 years (and subtracting wealth tax payments over the period). Specifically,

Lifetime resources = net wealth in 1994 + net labor income (1995–2013)

+ net government transfers (1995–2013)

+ net capital income (1995–2013)

+ net gifts and inheritances (1995–2013)

+ lottery gains (1995-2013)

+ net capital gains on real assets (1995–2013)

+ net capital gains on financial assets (1995–2013)

− wealth tax payments (1995–2013).

This measure of lifetime resources and its components allows us to see how differences in wealth arise. For example, two households with the same wealth in 1994 and the same inflow of income and capital gains over the previous 19 years would have the same lifetime resources. However, if one household consumed much more or saved and invested differently than the other, then the two households could have different wealth in 2013.

While lifetime resources provides an accounting of the various sources of income that can be used to accumulate wealth, it is not wholly unaffected by consumption decisions of households, as income that is used for consumption is not invested, so we will miss the returns that would have been gained on those investments. This slippage will be more important for individuals who consume more relative to their income inflows.

We use several Norwegian administrative registers to construct our dataset (Statistisk Sentralbyrå 2020a, 2020b, 2020c, 2020d, 2020e). We begin with the

---

[2] For all these inflows, "net" refers to after taxes, so, for example, net labor income equals total earned income minus all relevant taxes on that income.

Norwegian population register, which includes demographics and detailed family information. To construct measures of the sources of income and wealth, we combine the population register with information from Norwegian tax records, which are available with detailed information from 1993 until 2014. We restrict our sample to include a balanced panel of all adults registered as Norwegian residents in each year between 1994 and 2013. The sample ranges in age from 21 to 46 in 1994 and from 40 to 65 in 2013, and includes over 1.5 million individuals. We focus on this age range as we believe that wealth from youth into middle age (compared to wealth acquired at older ages) is particularly important for quality of life, residential choice, and human capital investment in children.

Because of the existence of a wealth tax in Norway, the government collects detailed information on all forms of wealth; as a result, we have detailed information on financial wealth, including physical capital held in private businesses, asset values, and debts. Our wealth data include all types of nonpension wealth of individuals and households, but we have no information on wealth held in pension funds. The exclusion of pension wealth is consistent with other studies, such as Charles and Hurst (2003) for the United States and Boserup, Kopczuk, and Kreiner (2017) for Denmark. The inclusion of pension wealth, which comes largely from labor income, would magnify the importance of labor income in our conclusions.

In the Norwegian data, most data reporting is by third parties—employers report employee earnings to the tax authorities, and bank and financial intermediaries report assets such as savings, stock values, and bonds—so usual measurement issues in household-reported survey data are greatly reduced.[3] Notably, the dataset encompasses the entire population of Norway, including all taxpayers. This allows us to avoid the typical "top-coding" problem that arises in studies of wealth. "Top-coding" arises where wealth (or income) at the very top of the distribution in a certain dataset is combined into a broad category like "$1 billion or more." Because wealth is highly concentrated at the top, top-coding will hide details of the wealth distribution, and the lack of top-coding—with the use of actual data instead—is an important feature of our data. Also, as further confirmation that measurement error is not a significant problem, we have verified that our main conclusions are robust to re-weighting our data to match aggregates from Norway's National Accounts. Another advantage of the Norwegian data is that our wealth measure includes private business wealth that entrepreneurs report to the tax authorities as the assessed value of their shares in a private business. Throughout the paper, we measure wealth and income variables in nominal Norwegian krona.

Traditionally, wealth is measured at either the individual or household level. Because we are tracking individuals over time—during periods of family formation

---

[3]One potential problem is wealth held in tax havens abroad and unreported to the Norwegian tax authorities (Alstadsæther, Johannesen, and Zucman 2019). Very few people hold wealth abroad, and a Norwegian tax amnesty in the early 2000s for holdings of assets abroad revealed that such behavior was concentrated among the very wealthy. Of those who disclosed holdings, half were among the 400 wealthiest (Zucman 2015).

and possibly dissolution—we use a hybrid measure to calculate our wealth and income variables. For each year, we consider the wealth and income of each individual: if they have a spouse or cohabitating partner, we add the wealth and income of both and allocate half to each individual. This metric enables us to allocate individual income to both members of a couple and to do the analysis at the individual level.

Further details about how exactly each variable is constructed are relegated to the appendix, but here is an overview.[4] We define gross wealth at the start of the period as financial wealth plus the value of real assets, and net wealth is defined as gross wealth minus debts.

We measure "net labor income" as gross labor income (from employment and self-employment) minus taxes and deductions. Similarly, "net transfer income" is income from government transfers—including pensions, unemployment benefits, sickness benefits, social assistance, and child support benefits—net of any taxes.

To measure "net capital income," we first define gross capital income as the sum of interest received from banks, plus dividends, plus other capital income, plus imputed rents from owner occupancy, minus interest paid on debts.[5] We then subtract taxes paid on capital income, excluding taxes paid on the sale of real or financial assets (which are included in the calculation of the returns to real and financial assets), plus the tax deduction on interest paid (interest paid on debts is deductible from capital income for tax purposes).

To calculate capital gains on real assets in each year, we multiply the value of housing assets—which includes primary residence and secondary residence—at the beginning of the year by the estimated percentage capital gain on these. To calculate the percentage capital gain in housing, we assume that primary and secondary houses appreciate in value at a rate equal to the percentage increase in house prices in the region of residence in that year. We assume there is no capital appreciation for nonhousing real wealth.

The capital gains on risky financial assets are calculated for each year as the total value of risky financial assets (total financial assets minus bank deposits) at the beginning of the year multiplied by the percentage annual return of the OBX index (the main stock market index in Norway). Capital gains on financial assets are subject to taxes, so we subtract an estimate of accumulated tax liabilities on capital gains from estimated capital gains.

Information on inheritances and gifts is taken from Norway's administrative registers and is available from 1995 to 2013. In Norway, inheritances and gifts are reported even if they are below the tax thresholds. Black et al. (2022) provide much evidence to suggest that misreporting is not a major issue and that our data on gifts and inheritances are unlikely to suffer from serious error. Whenever we mention

---

[4]For descriptive statistics with means, medians, and standard deviations on all the key measures in our study, see Appendix Table A2.

[5]We include imputed rental income in capital income as it provides a flow return to home ownership by obviating the need to pay rent. In essence, the imputed rent is an estimate of the rent that a homeowner who lives in their own home would have received if they instead rented out their house.

inheritances in this paper, we are referring to the sum of both inheritances and inter vivos gifts received.

Lottery gains do not matter much to the overall wealth picture because they are so rare, but we do have data from an administrative register with information on all large lottery winnings. Small sums—amounts below 10,000 or 100,000 Norwegian kroner depending on the tax period (that is, roughly $1,000 or $10,000 in US dollars)—are not reported in the register. These winnings are not subject to taxation.

In creating lifetime resources, we subtract only one item—the total wealth tax paid in the 1995–2013 period. The wealth tax paid each year is available in the data.

## Relationship between Net Wealth and Lifetime Resources

If net wealth and lifetime resources are highly correlated, this suggests that differences in spending and consumption patterns are not a large component of differences in wealth across individuals and that having a measure of net wealth at a point in time provides useful insight into the underlying distribution of lifetime resources. In contrast, if they are not highly correlated, this suggests that the observed wealth distribution is heavily influenced by differences in consumption patterns, and any inferences about the wealth distribution should be viewed through this lens. We investigate the relationship between net wealth and lifetime resources in this section and find that net wealth is highly correlated with lifetime resources; these correlations are strongest at the tails of the distribution. Overall, this suggests that net wealth provides a useful but imperfect proxy for lifetime resources.

Fact 1: Net Wealth is More Unequally Distributed than Lifetime Resources

Net wealth is more unequally distributed than lifetime resources. Figure 2 plots the distribution of actual net wealth in 2013, as well as our measure of lifetime resources. Given the presence of some extreme values, we have trimmed the top 1 percent and the bottom 0.1 percent of each distribution on the horizontal axis for this visualization. Because lifetime resources are partly used for consumption and spending, the mean of this measure is much larger than the mean of actual net wealth. We can see that the net wealth distribution is very skewed. Net wealth is negative on average in the bottom 20 percent as many people have debts that exceed the value of their gross financial or real assets. Additionally, net wealth remains low throughout the bottom half of the distribution, with median net wealth in 2013 being only about NOK1,000,000; 45 percent of net wealth belongs to the top 10 percent, 17 percent is held by the top 1 percent, and 7 percent is held by the top 0.1 percent. Therefore, given net wealth is low in the bottom half of the net wealth distribution, we can infer that the lifetime resources accrued by these households are primarily consumed rather than used to accumulate wealth.

While it is not obvious from Figure 2, the distribution of lifetime wealth is much less unequal than that of net wealth—the Gini coefficient is lower (0.27 compared to 0.59), the share of the total going to the lowest 50 percent is higher (33 percent

Distributions of Net Wealth and Lifetime Resources
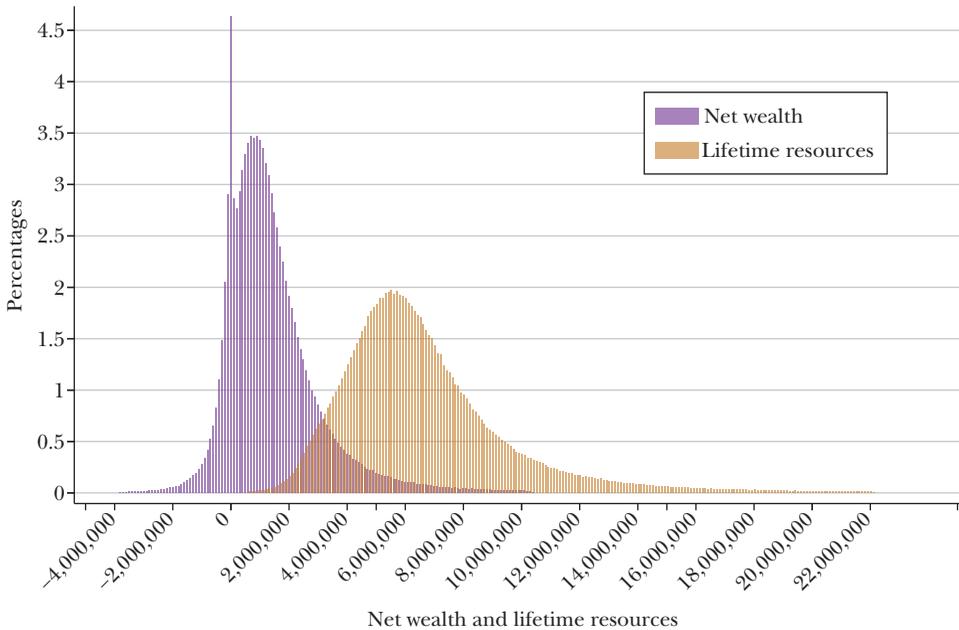
Source: Our calculations, based on several administrative registries maintained by Statistics Norway that we link through unique identifiers for individuals.
Notes: The figure refers to a balanced panel of all adults born between 1948 and 1973 and registered as Norwegian residents in each year between 1994 and 2013. The figure shows the distributions of lifetime resources and net wealth in 2013. All amounts in Norwegian Kroner. The top 1 percent and bottom 0.1 percent of lifetime resources and net wealth are trimmed.

versus 8 percent), and the share going to the top 1 percent is lower (7 percent versus 17 percent). A common lesson from the literature on inequality applies here as well: conclusions about inequality may be quite sensitive to the outcome considered.

Fact 2: Lifetime Resources and Net Wealth are Highly Correlated

In Figure 3, we divide individuals by decile of lifetime resources (horizontal axis) and see which deciles they occupy in the 2013 net wealth distribution (vertical axis), both calculated by age. We include the bottom nine deciles and, because of the importance of the top 1 percent of wealth-holders, we split the top decile into three groups—the 90th to 99th percentiles, the top percentile excluding the top 0.1 percent, and the top 0.1 percent. There is a strong correlation between the two variables, particularly at the bottom and top end; for example, about 90 percent of people in the top 1 percent of lifetime resources are in the top 10 percent of actual net wealth in 2013, and about 60 percent of this group are in the top 1 percent of 2013 net wealth. However, as Figure 3 suggests, the relationship is weaker in the middle of the two distributions.

Figure 3

Proportion of Individuals Belonging to the Different Deciles of Net Wealth by Deciles of Lifetime Resources
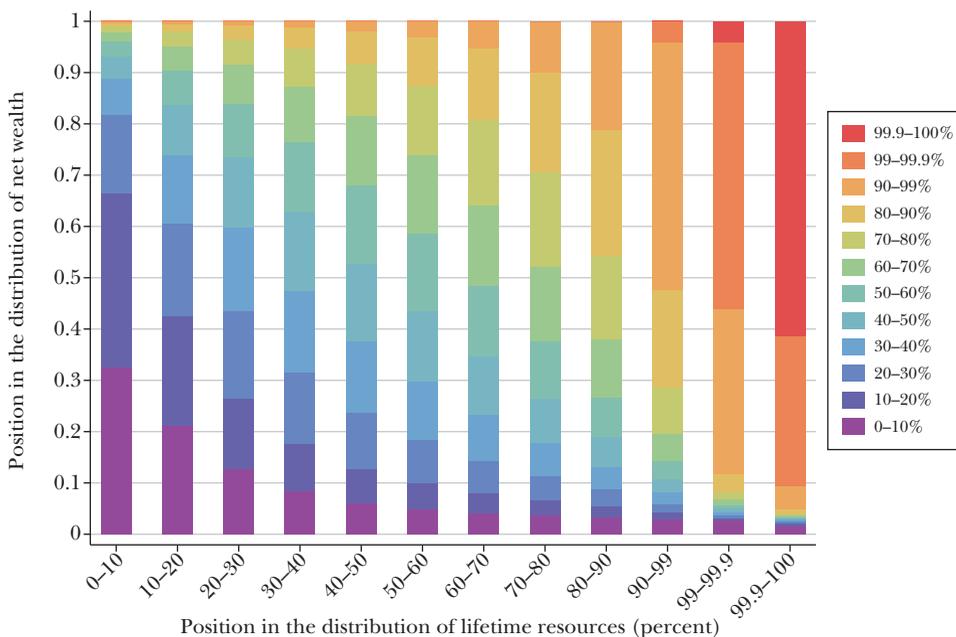
Source: Our calculations, based on several administrative registries maintained by Statistics Norway that we link through unique identifiers for individuals.

Notes: The figure refers to the same sample as Figure 2, and shows the proportion of individuals belonging to the different deciles of net worth in 2013 by deciles of lifetime resources, with the top 1 percent (excluding the top 0.1 percent) and the top 0.1 percent as separate categories. For each decile of lifetime resources, the corresponding colored bars represent the average share of indivduals belonging to this decile of net wealth. The shares sum to 1, and the deciles are age-specific.

We find little evidence that the correlation between lifetime resources and net worth in 2013 varies based on individual characteristics.[6] When we break the sample along a variety of dimensions, such as family size, age, and education, the correlation is quite similar across the range of characteristics. For example, the correlation varies only between 0.71 for those who are aged 40 in 2013 and 0.80 for those who are aged 65 in 2013; when we break by two broader age categories (45–54 in 2013 and 55–65 in 2013), we see, again, that the correlations are quite similar, 0.75 and 0.79, respectively. When we break the sample by family size (number of children), we see correlations of 0.77 for individuals with no children, 0.76 for individuals with 1 child, 0.77 with 2 children, and 0.78 with three or more children. (This is perhaps

---

[6]The results discussed here are obtained when winsorizing the top and bottom 0.1 percent of these distributions. Perhaps unsurprisingly, given the large outliers in wealth, the unwinsorized correlations are not very stable across characteristics.

surprising if we consider family size as a proxy for consumption.) When we break by educational attainment, the coefficients are again almost identical, with those who are more educated (defined as having more than a high school education) having a correlation of 0.78 while less educated individuals have a correlation of 0.75.

Fact 3: Lifetime Resources Predicts the Educational Outcomes of Children Better than Net Wealth Does

Considerable attention has been paid to understanding the effect of family wealth on the educational outcomes of children. Because parental wealth is correlated with many other variables that also affect children's outcomes, some researchers have sought natural experiments that generate plausibly random variation in family wealth. For example, Lovenheim and Reynolds (2013) show that increases in household wealth arising from increases in US housing prices make children of low- and middle-income families more likely to attend a four-year state university than a community college. Other research studies the effect of parental wealth on children by studying adoptees, which enables researchers to distinguish between the effects of nature (biological components) and nurture (environmental components). As one example, Black et al. (2020) study adopted children in Sweden and find that nurture is more important than nature in the intergenerational transmission of wealth. Trying to summarize the findings of such a varied group of studies is of course subject to caution, but it is fair to say that wealthier parents tend to pass along benefits to their children in the form of higher education levels, as well as through other human capital investments and financial transfers.

However, there is a more basic question: Is net wealth a more relevant measure for children's outcomes than lifetime resources? As a very simple first check, we test which of these two measures best predicts the education outcomes of children. Note that we are not trying to estimate causal effects; rather, we seek only to understand which variable provides a better prediction of the human capital outcomes of children. Perhaps surprisingly, we find that lifetime resources is a better predictor than net wealth.

To conduct this analysis, we restrict our sample to parents whose first child is between 16 and 22 in 2013 (or between 20 and 22 when studying higher education outcomes). We focus on four outcomes: middle school grade point average; whether the child enrolled in a more rigorous academic high school track; whether the child enrolled in college; and the number of completed years of schooling. Most Norwegians complete middle school around age 16, and a student's grade point average, ranging between 10 and 60, is a measure of achievement during core schooling. After middle school, students choose between two different tracks for high school, the academic track and the vocational track. The academic track is a preparation for university and other higher educational studies. Restricting our sample to first-born children who are aged between 16 and 22 in 2013 ensures that they have been in the household for most or all of the 1995–2013 period over which we measure lifetime resources. We have information on educational outcomes for all years up to and including 2018, when children are aged between 21 and 27.

Because many individuals do not complete education before age 25, we restrict the sample to children aged between 20 and 22 in 2013 (between 25 and 27 in 2018) when studying completed years of education.

To compare predictive power, we run separate regressions for lifetime resources and for net wealth; in the regressions, we regress children's outcomes on either the family's net wealth in 2013 or their lifetime resources in 2013. To do so, we include 99 indicators for the parents' percentile of the distribution of lifetime resources or net wealth: the predictive power of the regression (as measured by R-squared) is about twice as high for lifetime resources as it is for net wealth.[7] This is possibly because, for much of the population, lifetime resources is primarily comprised of labor income, which, in turn, is more highly correlated with their parents' human capital than with their wealth.[8] As a result, it may be more useful to focus on lifetime resources rather than net wealth when studying the effects of inequality of opportunity on the human capital accumulation of the next generation.

Sources of Lifetime Resources across the Distribution

We next examine which particular income sources, aggregated over the 1995–2013 period, are the largest components of lifetime resources calculated over the same period. We find that lifetime resources is disproportionately composed of labor income for all except the top 1 percent of the distribution, who get most of their resources from capital income and capital gains on financial assets. The relative contributions of the components are almost identical across deciles of lifetime resources and actual net wealth in 2013.
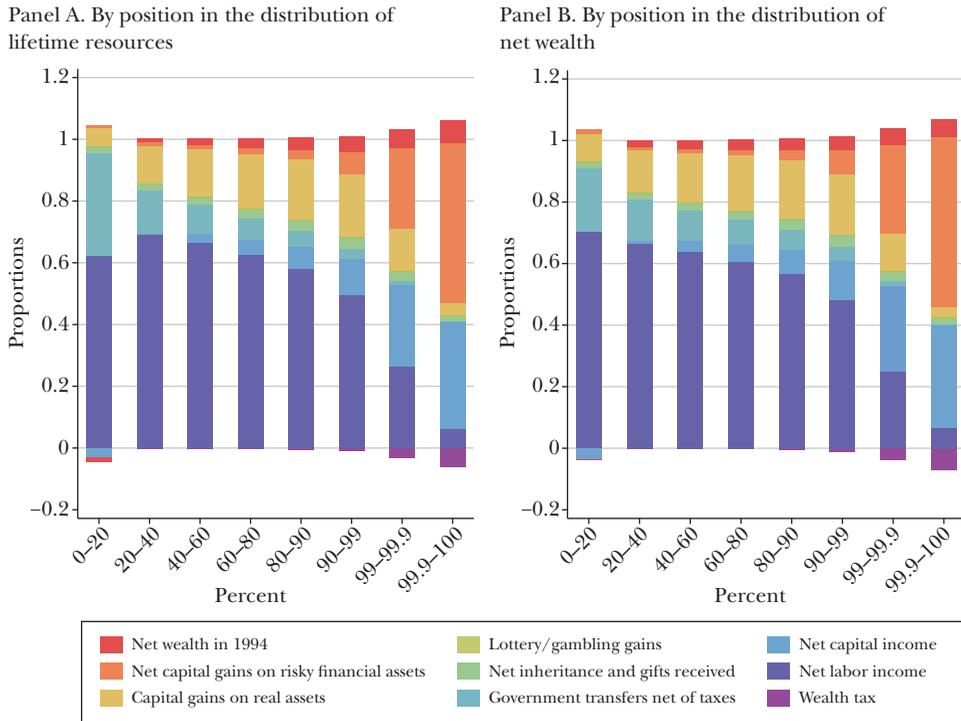
Before discussing the findings, it is important to be clear that this analysis is purely descriptive. For example, when we show the contribution of inheritances to lifetime resources, this does not account for higher investment returns that accrue as a result of having higher inheritances. Any such returns will show up as capital gains, capital income, or returns on financial assets. So, while inheritances may show up as a small component of lifetime resources, this does not imply that lifetime resources are not heavily influenced by inheritances. Similar arguments can be made for other components of lifetime resources, so the analysis should be considered as simply an accounting of how the sources of income vary over the lifetime resources and net wealth distributions.

---

[7] Again, for details of these regressions, see the online Appendix Table A3.
[8] We find that the correlation between parents' years of education and their position in the distribution of lifetime resources is higher (0.33) than the correlation between parents' years of education and their position in the distribution of net wealth in 2013 (0.19).

Figure 4

Components of Lifetime Resources by Position in the Distribution of Lifetime Resources or Net Wealth

Panel A. By position in the distribution of lifetime resources

Panel B. By position in the distribution of net wealth

Legend:
- Net wealth in 1994
- Net capital gains on risky financial assets
- Capital gains on real assets
- Lottery/gambling gains
- Net inheritance and gifts received
- Government transfers net of taxes
- Net capital income
- Net labor income
- Wealth tax

Source: Our calculations, based on several administrative registries maintained by Statistics Norway that we link through unique identifiers for individuals.

Note: The figure refers to the same sample as Figure 2, and shows the components of lifetime resources by position in the distribution of lifetime resources (panel A), or by position in the distribution of net wealth in 2013 (panel B). For each quantile and each component, the corresponding colored bars represent the average share of that component in lifetime resources. The shares sum to 1, and the quantiles are age-specific.

Fact 4: Lifetime Resources is Disproportionately Composed of Labor Income for All Except the Top 1 Percent of the Distribution

We document our findings in Figure 4. Here, blocks below zero represent negative income (outflows), while blocks above zero represent positive sources of income. The size of each block reflects the proportion of lifetime resources attributable to that source and, by construction, the totals add up to 1. At the bottom of the lifetime resources distribution, for those with the lowest lifetime resources, we see a larger role for government transfers, at 33 percent of lifetime resources for the bottom 20 percent. This share declines as we move up the distribution, with a low of 0.3 percent for the top 0.1 percent. At the same time, capital income is very low or even negative (due to interest payments on debt) at the bottom of the

lifetime resources distribution, but increases steadily across the distribution, with a high of 35 percent for the top 0.1 percent. When considering capital gains on real assets (largely due to homeownership), the share is relatively low at the bottom of the lifetime resources distribution (6 percent for the bottom 20 percent), then increases to 18 percent for the 60–80th percentile and 20 percent for the 80–90th percentile and the 90–99th percentile, before falling at the very top of the distribution (with 4 percent for the top 0.1 percent). Net capital gains on risky financial assets only become significant at the top of the distribution, at 26 percent for the top 1 percent and 52 percent for the top 0.1 percent. Importantly, we find only a small role for inheritances, even at the top of the distribution, with a range of 2 percent for the bottom 20 percent to 4 percent for the 90–99th percentile.[9]

A major takeaway from Figure 4 is that lifetime resources is disproportionately composed of labor income for all except the top 1 percent of the distribution, who get most of their resources from capital income and capital gains on financial assets. In addition, perhaps surprisingly, inheritances are not an important source of lifetime resources at any point in the distribution even at the very top.

Fact 5: Components of Lifetime Resources across the 2013 Net Wealth Distribution Are Similar to Those across the Distribution of Lifetime Resources

While we have focused on the distribution of lifetime resources so far, one question is how different does this figure look if we consider the distribution of net wealth—a variable we are much more likely to have in data. As we saw earlier, while there is a high correlation between net wealth and lifetime resources, particularly at the top of the distribution, this correlation is far from perfect.

To answer this, we compare panel A and panel B of Figure 4; panel B shows how the constituents of lifetime resources vary across the distribution of net wealth. With the exception of the bottom decile—government transfers are a much larger component of lifetime resources for the bottom decile of the lifetime resources distribution than the bottom decile of the net wealth distribution—the role of the components is almost identical across deciles of lifetime resources and actual net wealth in 2013, suggesting that, on this dimension, net wealth might be a reasonable proxy for lifetime resources.

The households that have large amounts of net wealth (approximately those in the top decile) have lifetime resources that comprise less of labor income and more

---

[9]As mentioned earlier, this analysis may understate the role of inheritances if people with higher inheritances have higher capital income and capital gains as a result. To check this, we have created an alternative measure of lifetime resources that only includes the primary sources of income—net labor income, government transfers, gifts and inheritances, and initial wealth—and cumulate these over time, allowing them to grow at a specified interest rate that varies based on an individual's decile of net wealth in each year. Using this measure, we still find a very small role for inheritance relative to labor income: for the richest 1 percent, capitalized inheritance and gifts represent 5 percent of their cumulated primary incomes, while capitalized labor income accounts for 41 percent of this sum. The role of inheritances is explored in more detail in Black et al. (2022).

of capital gains and financial returns. A key takeaway here is that, while labor earnings are very important throughout the distribution of net wealth, at the very top of the wealth distribution, individuals get most of their resources from capital gains and returns on financial assets.

Fact 6: Children of Wealthy Parents Accumulate Wealth Differently

While inheritance does not appear to be an important factor in wealth accumulation on average, there is a well-established correlation in net wealth across generations. However, while we know that children with wealthier parents are more likely to be wealthy themselves, very little is known about how parental wealth affects the sources of wealth of their children. Do children of wealthy parents accumulate wealth differently than children from less wealthy families?

To answer this, we examine the components of lifetime resources for children whose parents are at various points of the net wealth distribution in 1994. We measure parental wealth as the sum of both parents' wealth in 1994 (whether or not both parents are in the same household)—if only one parent is alive in 1994, we measure parental wealth as the wealth of that parent. Note that the sample we use for this analysis is somewhat smaller than in our previous analysis, as we drop individuals for whom we do not know parental wealth in 1994. We do not have information on parental net wealth for about 11 percent of the sample, either because both parents died before 1994 or because they were not living in Norway in 1994.

Given that our sample includes only persons with at least one living parent in 1994, the inheritances received over the 1995–2013 period are an underestimate of the inheritances that will eventually be received. For about 30 percent of the sample, the last surviving parent dies between 1994 and 2013 so, in the majority of cases, children have not received inheritances from their parents by 2013. However, they may have received inter-vivos gifts from their parents or an inheritance from their spouse's parents during our sample period, and these are included in our measure of inheritances.

In Figure 5, we split the sample by decile of parental net wealth in 1994 to determine how the relative importance of the components of lifetime resources differs by parental wealth. We see that individuals with rich parents get their wealth disproportionately from capital income and returns on financial wealth. Also, wealth in 1994 and inheritances and gifts are a more important component of lifetime resources for the children of the rich. Taking our results together, children of parents in the top 1 percent of the wealth distribution have higher initial wealth in 1994 and receive greater inheritances than others, and they invest these resources, thereby receiving a disproportionate amount of their wealth from returns on financial investments and capital income. Overall, we conclude that, not only do children with wealthy parents end up with greater wealth (the average net wealth in 2013 of children in the top 0.1 percent of 1994 parental wealth is five times that of children of the 99–99.9 percent group and twelve times that of the

Figure 5

Components of Lifetime Resources by Position in the Distribution of Parental Wealth in 1994
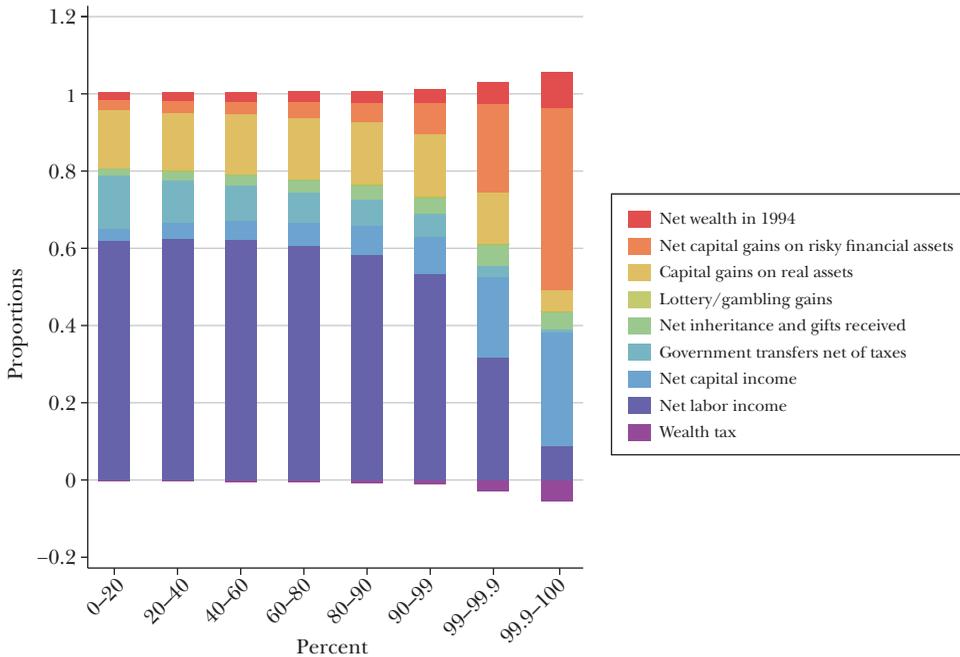
Source: Our calculations, based on several administrative registries maintained by Statistics Norway that we link through unique identifiers for individuals.

Note: The figure refers to the same sample as Figure 2, restricted to individuals with at least one parent alive and resident of Norway in 1994. The Figure shows the components of lifetime resources by position in the distribution of parental wealth measured in 1994. For each quantile and each component, the corresponding colored bars represent the average share of that component in lifetime resources. The shares sum to 1, and the quantiles are age-specific.

90–99 percent group), but they also accumulate their wealth in very different ways compared to most members of society.[10]

Fact 7: Variation in Labor Income Increases Variation in Lifetime Resources throughout the Distribution, While Other Components Predominantly Affect Only the Bottom or the Top of the Distribution

Another way to assess the importance of different sources of lifetime resources in terms of overall inequality is to consider what the lifetime resources distribution

---

[10] In addition, the average net wealth in 2013 of children in the top 0.1 percent of 1994 parental wealth is 19 times that of children of the 50–90 percent group, and 29 times that of children in the bottom 50 percent of 1994 parental wealth.
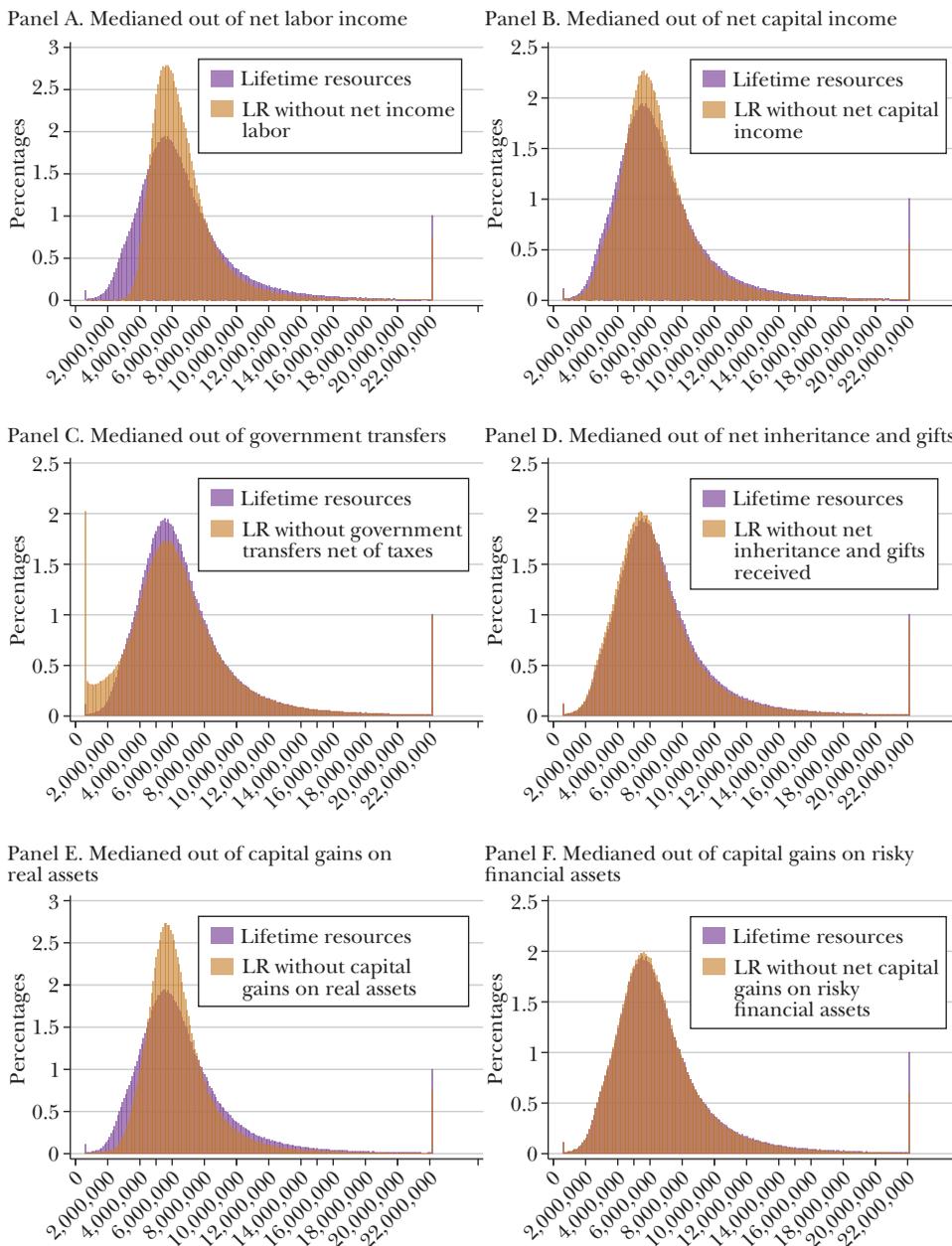
would look like if variation in any one component is removed. For example, what would the lifetime resources distribution look like if everyone received the same labor income? Because lifetime resources is just the sum of the components over the entire time period, it is easy to equalize any component and see how the lifetime resources distribution changes. Note that a key limitation with this type of analysis is that it ignores spillover effects—changing one component could lead to changes in other components (for example, changing labor income may change investment behavior and hence returns from assets). Because of these omitted spillover effects, we are likely underestimating the effect of a change in any one component on the overall distribution, and the analysis should be considered as purely descriptive.

Figure 6 does these analyses, showing the full distribution of actual lifetime resources (in purple) overlaid with what the distribution would look like under different scenarios (in yellow). To enable us to focus on the variation (as distinct from the location) of the lifetime resources distribution, our adjusted distributions replace the actual value of the component for each person with the median value of that component, so it is the distribution if everyone had the same amount (the median value) for a particular component—we refer to this process as "medianing out" the component.[11] (The median values of each component are reported in Appendix Table A2.)

A visual inspection of Figure 6 yields several main findings. First, equalizing labor income substantially compresses the lifetime resources distribution, reducing the Gini coefficient from 0.27 to 0.20. The large effects of labor income on the lifetime resources distribution are particularly noteworthy given how compressed the income distribution is in Norway compared to other countries (OECD 2020). Second, equalizing net capital income also reduces the variance of lifetime resources, but to a much smaller extent, reducing the Gini coefficient from 0.27 to 0.23. However, it has substantially larger effects at the very top of the distribution: the fraction of lifetime resources held by the top 1 percent of the lifetime resources distribution goes down from 7 percent to 5 percent when we give everyone the median net capital income. While this may appear small, it represents a 30 percent decrease in the share of lifetime resources held by the top 1 percent. Net capital gains on real assets increase the variance of the lifetime resources distribution— both at the top and the bottom—so setting them equal to the median value reduces the variance of the distribution, and the Gini coefficient declines from 0.27 to 0.22. Net capital gains on risky financial assets increase variation in lifetime resources only towards the top of the distribution and have a particularly large effect on the top 1 percent. Equalizing net capital gains on risky financial assets reduces the Gini coefficient from 0.27 to 0.24, with the share going to the top 1 percent falling from 7 percent to 5 percent.

---

[11] Additionally, to make the pictures tractable, we winsorize the top 1 percent and bottom 0.1 percent, where these quantiles are defined based on lifetime resources rather than on lifetime resources with the relevant component medianed out.

Figure 6

Distribution of Lifetime Resources and Lifetime Resources Medianed Out of Its Components

Panel A. Medianed out of net labor income

Panel B. Medianed out of net capital income

Panel C. Medianed out of government transfers

Panel D. Medianed out of net inheritance and gifts

Panel E. Medianed out of capital gains on real assets

Panel F. Medianed out of capital gains on risky financial assets

Source: Our calculations, based on several administrative registries maintained by Statistics Norway that we link through unique identifiers for individuals.
Note: The figure refers to the same sample as Figure 2. The figure shows the distribution of lifetime resources, and its adjusted distributions, replacing individuals' net labor income, net capital income, government transfers, net inheritance or gifts, capital gains on real assets or capital gains on risky financial assets by its median value in the sample. "LR" is lifetime resources. The top and bottom values of lifetime resources and its adjusted distributions are winsorized, using as caps the top 1 percent and bottom 0.1 percent of lifetime resources.

In contrast, government transfers compress the lifetime resources distribution, and their effect is concentrated at the bottom—the Gini coefficient increases from 0.27 to 0.31 and the share going to the bottom 40 percent decreases from 24 percent to 22 percent. A very small proportion of individuals (0.2 percent) have less than NOK1,000,000 of lifetime resources; however, this proportion increases to 3 percent when we equalize government transfers.

The adjusted distribution for lifetime resources when we give everyone the median value of inheritances and gifts (only about NOK9,000, or $1,200 in 2013) is very similar to the actual lifetime resources distribution (the Gini coefficient only changes from 0.267 to 0.265) implying that this component has little effect on the overall variation in lifetime resources.

Unsurprisingly, given their rarity, lottery gains have no discernable effect on the lifetime resources distribution, and we do not show this in Figure 6.

Overall, government transfers, net labor income, and net capital gains on real assets tend to affect the bottom of the distribution, while net labor income, net capital income, net capital gains on real assets, and net capital gains on financial assets affect the spread at the top of the lifetime resources distribution.

Conclusion

How well might the Norwegian patterns of lifetime resource accumulation apply to other countries such as the United States? There are several reasons to be cautious about extrapolating our findings to the United States, as the United States is an outlier in many ways. First, the US distribution of wealth is considerably more unequal than the norm for Europe and other high-income countries, as illustrated in Figure 1: among the countries featured, the top 10 percent share of the wealth distribution in the United States is exceeded only by Mexico and Chile. In Norway, the Gini coefficient for inequality of household net wealth was 0.68 in 2014; for a rough comparison, Wolff (2016) calculates the Gini coefficient for the United States to be 0.87 in 2013. Second, Norway ranks among the countries with the most compressed distribution of labor income. Based on OECD data, the Gini coefficient of income was 0.26 in 2018—one of the lowest within the OECD but similar to that in other Nordic countries and some central European countries like the Czech Republic (OECD 2020). In contrast the US Gini coefficient for income is quite high at about 0.4.

Third, individuals in the United States need to self-insure against negative health or income shocks in a way that is not necessary in other countries with public provision of health insurance and strong social safety nets, so we might expect differential savings behavior that affects the link between net wealth and lifetime resources. Fourth, differential tax treatment of investments in the United States and different forms of retirement savings (through defined contribution instead of defined benefit pension plans) might also lead to different resource allocation across assets and hence influence the relative importance of various sources of lifetime resources. Fifth, patterns of returns on assets may differ in the US economy.

Despite such differences, there is reason to believe that some of the more general lessons from the study of Norway's data continue to hold. For example, studies based on US data have also emphasized the role of labor earnings for wealth accumulation, even at the top of the distribution. Using cross-sectional US data from the 2010 and 2016 Survey of Consumer Finances, Kaymak, Leung, and Poschke (2020) show that even the top 1 percent of the wealth distribution have a large proportion of their income coming from labor earnings. The World Inequality Database uses national accounts to provide comparable estimates of private wealth to household income ratios across countries; for Norway, the ratio was 5.38 in 2013, similar to the United States (5.66). Because inheritances come from (parental) wealth, these ratios imply that the importance of inheritances relative to other income sources may be similar in Norway to that in the United States. Indeed, Bauluz and Meyer (2021) use US data from the Survey of Consumer Finances to show that inheritances account for a small proportion of annual income, consistent with our findings here and with our detailed study on inheritances in Norway (Black et al. 2022).

Moving beyond the United States, it is more likely that the findings from Norway also apply to other European countries. As we discussed earlier, along many dimensions, the wealth distribution of Norway emulates that of many other European countries, so we might expect similar conclusions. In Norway, house price appreciation and returns to risky financial assets were broadly similar over the 1994–2013 period to those in many other European countries (Jordà et al. 2019). In terms of home ownership rates, Norway is right at the mean of the euro area. In our data, 64.1 percent of households owned their main residence in 2014. The comparable number in the euro area in the Household Finance and Consumption Survey was 61.0 percent. Much of what we know about differences in savings behavior and returns to investments between poorer and wealthier households comes from studies using Norwegian data (Fagereng et al. 2019, 2020). It is not obvious that these patterns differ in a substantial way between Norway and other economies of western Europe. On the other hand, the World Inequality Database suggests that the level of private wealth is relatively low in Norway compared to other many other countries, but similar to that in Sweden and Finland.[12]

One could speculate about the likely effects of differences between Norway and other countries for the accumulation of lifetime resources and net wealth. However, we simply emphasize that these variations highlight the importance of having a more comprehensive understanding of wealth accumulation over the life cycle, in the United States and in other countries as well, as more detailed wealth and income data become available internationally.

---

[12] From the World Inequality Database, average individual private net wealth in 2013 (in €2019) was €101,026 for Norway, €127,357 for Germany, €155,754 for France, €125,871 for Sweden, €91,573 for Finland, and €177,394 for Italy. These data include pension wealth and are thus not directly comparable to our data. Using our administrative data, we find that average individual private net wealth in 2013 (in €2019) was €108,493 for Norway.

References

Alstadsæther, Annette, Niels Johannesen, and Gabriel Zucman. 2019. "Tax Evasion and Inequality." *American Economic Review* 109 (6): 2073–2103.

Bauluz, Luis, and Timothy Meyer. 2021. "The Great Divergence: Intergenerational Wealth Inequality in the U.S. and France." Unpublished.

Black, Sandra E., Paul J. Devereux, Fanny Landaud, and Kjell J. Salvanes. 2022. "The (Un)Importance of Inheritance." NBER Working Paper 29693.

Black, Sandra E., Paul J. Devereux, Fanny Landaud, and Kjell G. Salvanes. 2023. "Replication data for: Where Does Wealth Come From? Measuring Lifetime Resources in Norway." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193581V1.

Black, Sandra E., Paul J. Devereux, Petter Lundborg, and Kaveh Majlesi. 2020. "Poor Little Rich Kids? The Role of Nature versus Nurture in Wealth and Other Economic Outcomes and Behaviours." *Review of Economic Studies* 87 (4): 1683–1725.

Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner. 2017. "Intergenerational Wealth Formation over the Life Cycle: Evidence from Danish Wealth Records 1984–2013." Unpublished.

Charles, Kerwin Kofi, and Erik Hurst. 2003. "The Correlation of Wealth across Generations." *Journal of Political Economy* 111 (6): 1155–82.

Eika, Lasse, Magne Mogstad, and Ola L. Vestad. 2020. "What Can We Learn about Household Consumption Expenditure from Data on Income and Assets?" *Journal of Public Economics* 189: 104163.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and Persistence in Returns to Wealth." *Econometrica* 88 (1): 115–70.

Fagereng, Andreas, Matin Blomhoff Holm, Benjamin Moll, and Gisle Natvik. 2019. "Saving Behavior across the Wealth Distribution: The Importance of Capital Gains." NBER Working Paper 26588.

Household Finance and Consumption Network. 2020. "The Household Finance and Consumption Survey, Wave 2014, Statistical tables." ECB Statistics Paper, Technical Report.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor. 2019. "The Rate of Return on Everything, 1870–2015." *Quarterly Journal of Economics* 134 (3): 1225–98.

Kaymak, Baris, David Leung, and Markus Poschke. 2020. "Accounting for Wealth Concentration in the US." IZA Discussion Paper 13082.

Kopczuk, Wojciech, and Emmanuel Saez. 2004. "Top Wealth Shares in the United States, 1916–2000: Evidence from Estate Tax Returns." *National Tax Journal*: 57(2): 445–87.

Kuhn, Mortiz, Moritz Schularick, and Ulrike I. Steins. 2020. "Income and Wealth Inequality in America, 1949–2016." *Journal of Political Economy* 128 (9): 3469–3519.

Lovenheim, Michael F., and C. Lockwood Reynolds. 2013. "The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom." *Journal of Human Resources* 48 (1): 1–35.

OECD. 2020. "Income Inequality." https://data.oecd.org/inequality/income-inequality.htm (accessed

on July 21, 2023).

Saez, Emmanuel, and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131 (2): 519–78.

Saez, Emmanuel, and Gabriel Zucman. 2020a. "The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts." *Journal of Economic Perspectives* 34 (4): 3–26.

Saez, Emmanuel, and Gabriel Zucman. 2020b. "Trends in US Income and Wealth Inequality: Revising after the Revisionists." NBER Working Paper 27921.

Statistisk Sentralbyrå. 2020a. "Arveavgift (opphørt), 2013." https://www.ssb.no/inntekt-og-forbruk/statistikker/arv/aar/2015-06-18?fane=om#content (accessed on September 29, 2023).

Statistisk Sentralbyrå. 2020b. "Data om folkemengden eller familier og husholdninger, innvandring, og andre demografiske opplysninger." https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/befolkning (accessed on September 29, 2023).

Statistisk Sentralbyrå. 2020c. "Data om inntekter, formue, barnetrygd eller pensjoner." https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/inntekt (accessed on September 29, 2023).

Statistisk Sentralbyrå. 2020d. "Data om personers utdanningsnivå og utdanningsretning, karakterer og resultater." https://www.ssb.no/data-til-forskning/utlan-av-data-til-forskere/variabellister/utdanning (accessed on September 29, 2023).

Statistisk Sentralbyrå. 2020e. "Dødelighetsutvikling og dødsårsaksmønster 1961–2018."

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick. 2019. "Capitalists in the Twenty-First Century." *Quarterly Journal of Economics* 134 (4): 1675–1745.

WID.world. 2023. "The World Inequality Database." https://wid.world/news-article/new-get-wid-world-inequality-data-stata-ssc-package (accessed on August 23rd, 2023).

Wolff, Edward N. 2016. "Household Wealth Trends in the United States, 1962 to 2013: What Happened over the Great Recession? *RSF: Russell Sage Foundation Journal of the Social Sciences* 2 (6): 24–43.

Zucman, Gabriel. 2015. *The Hidden Wealth of Nations: The Scourge of Tax Havens.* Chicago: University of Chicago Press.

The Importance of Financial Literacy: Opening a New Field

Annamaria Lusardi and Olivia S. Mitchell

People face complex financial decisions with potentially long-lasting consequences at all stages of life. As young people grow into adulthood, they make decisions about loans for college tuition, cars, and houses, along with how to manage credit cards, health and other kinds of insurance, and living within a budget. The shift from defined benefit to defined contribution retirement plans implies that ordinary people must now shoulder decisions about saving, investing, and more. Older people face decisions about how to manage risks and costs of aging, as well as drawing down their retirement assets. These decisions have only become more complex with the advent of new financial products (which, with the help of technology, one can access with a click), novel ways to make payments ("buy now, pay later"), risky instruments such as crypto assets, and most recently the rise of inflation. According to Google Trends, searches for how to budget or save for retirement have increased fourfold since 2004.

For these reasons and others, *financial literacy*, by which we mean people's knowledge of and ability to use fundamental financial concepts in their economic decision-making, matters and is more important than ever. The fact that so many people lack financial knowledge not only limits their ability to utilize their resources to the fullest, but also contributes to macroeconomic problems. Recent economic crises related to the subprime mortgage debacle and the COVID-19 pandemic

■ *Annamaria Lusardi is Senior Fellow at the Stanford Institute for Economic Policy Research (SIEPR) and the Director of the Financial Freedom Initiative, a collaboration between the Graduate School of Business (GSB), SIEPR, and the Economics Department, Stanford University, Stanford, California. Olivia Mitchell is Professor of Insurance/Risk Management and Business Economics/Policy, and Executive Director of the Pension Research Council, at The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania. Their email addresses are alusardi@stanford.edu and mitchelo@wharton.upenn.edu.*

further illuminate how the confluence of debt and peoples' lack of financial cushions can affect the economy at large. In the wake of the pandemic and its consequent economic dislocations, it has become even clearer that people must hold precautionary savings and manage money properly to secure their economic wellbeing (Demertzis, Domínguez-Jiménez, and Lusardi 2020; Clark, Lusardi, and Mitchell 2021). In the longer term, differences in financial literacy also contribute to wealth inequality (Lusardi, Michaud, and Mitchell 2017).

In the early 2000s, we designed and fielded a survey in the United States to measure financial literacy and to understand how it affects financial decision-making. These questions have now been adopted in numerous surveys around the world. In what follows, we illustrate what we have learned about financial literacy in the United States and elsewhere, as well as how this knowledge differs across population subgroups. In turn, quantitative measurement of financial literacy permits researchers to measure the impact of financial literacy on important economic outcomes. In fact, financial literacy has become its own field of study and many countries have mandated financial literacy in school, beginning with elementary education. We conclude our discussion with thoughts on how our research findings can be used in research, teaching, and policy.

Quantifying Financial Literacy

Our early work on financial literacy was prompted by concerns about inadequate saving, financial vulnerability, and retirement insecurity among Americans. At that time, no nationally representative datasets existed to measure what people knew (or did not know) about the fundamentals of economics and finance. In 2004, we created and fielded an experimental module on financial literacy for the Health and Retirement Study (HRS). Four principles informed the design of those questions: simplicity, relevance, brevity, and capacity to differentiate. Specifically, our questions sought to measure knowledge of the key building blocks for financial decision-making in an intertemporal setting (simplicity). We also required that the questions relate to concepts pertinent to peoples' day-to-day financial decisions over the life cycle (relevance) and captured general rather than context-specific ideas. Finally, the number of questions had to be parsimonious to ensure widespread adoption (brevity), while still differentiating across people in terms of their financial knowledge (capacity to differentiate).

This effort produced what is now known as the "Big Three," a short set of questions that over the years has proven to be an extremely good measure of peoples' understanding of basic financial concepts (Lusardi and Mitchell 2011a, b, c). Table 1 lists these questions.

In practice, the fact that we were limited to only a handful of questions in the Health and Retirement Study module proved to be a blessing in disguise, as it has been easy for other surveys to add the Big Three questions too. Among US surveys, the Big Three have been included in the National Longitudinal Survey of Youth (NLSY), the Rand American Life Panel (ALP), the Understanding America Study

Table 1

The "Big Three" Financial Literacy Questions and Answers

(correct answers indicated with two asterisks)

1) Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?
>   More than $102**
>   Exactly $102
>   Less than $102
>   Do not know
>   Refuse to answer

2) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?
>   More than today
>   Exactly the same
>   Less than today**
>   Do not know
>   Refuse to answer

3) Please tell me whether this statement is true or false. "Buying a single company's stock usually provides a safer return than a stock mutual fund."
>   True
>   False**
>   Do not know
>   Refuse to answer

Source: The Big Three originated with Lusardi and Mitchell (2011a).

(UAS), the National Financial Capability Study (NFCS), and the Survey of Consumer Finances (SCF), just to mention some of the most prominent. The Big Three have also been included in more than 40 surveys fielded in Europe, Latin America, and Asia (for a review of non-US findings, see Lusardi and Mitchell 2011c, 2014). The widespread use of these questions has enabled us to develop international comparisons of financial literacy, and financial literacy questions akin to the Big Three were also included in the S&P Global Financial Literacy Survey covering more than 140 countries (Klapper and Lusardi 2020).[1]

The Big Three questions are simple, yet they test for fundamental knowledge at the basis of most economic decisions. If people are unfamiliar with these topics, they are much less likely to know about more complex concepts such as the relationship between risk and return, the term structure of interest rates, and how interest compounds over long periods. In addition, answering these questions does not require difficult calculations, as we do not test for mathematical skills but rather for an understanding of how interest rates and inflation work. The questions also test knowledge of the language of finance, because, for example, we do not explain what interest rates, inflation, or a stock mutual fund are. Making the questions multiple choice and giving respondents the option to say "do not know" (or refuse to answer) avoids forcing respondents to pick an answer at random. Indeed, the prevalence

---

[1] These questions were designed by Leora Klapper and Annamaria Lusardi in collaboration with Gallup.

Table 2
Financial Literacy in the US Population: Big Three Questions

	Correct	*Incorrect*	*Do not know/ refuse*
Interest	80.6%	16.4%	3.0%
Inflation	75.5%	20.7%	3.8%
Risk	60.7%	17.1%	22.2%
All Big Three correct	43.3%		

Source: Authors' tabulations, 2019 Survey of Consumer Finance.
Note: All data weighted using sampling weights.

of these choices offers insights that go beyond correct versus incorrect responses, and they help us evaluate when respondents are unsure of their knowledge, as we explain in more detail below.

This question set offers rich information about peoples' financial literacy. To illustrate findings, we next provide evidence from the 2019 Survey of Consumer Finances, the most recent data available from one of the best surveys on wealth (Bhutta et al. 2020). We then describe our more detailed measures of financial literacy, which show similar patterns.

As reported in many of our previous studies, and as shown in Table 2, financial literacy is strikingly low in the United States. For instance, while 81 percent of Americans understand simple interest rates, about three-quarters get the inflation question correct, and only 61 percent of the population knows that a single stock is riskier than a stock mutual fund. Overall, only 43 percent of Americans answer all three questions correctly. Thus, knowledge of basic financial concepts cannot be taken for granted, even in a country with well-developed financial markets and where these topics have been important for decades. Knowledge is particularly low about risk diversification, a relevant and fundamental concept, and where the percentage of "do not know" answers is strikingly high with respect to the first two questions.

Financial illiteracy is not only widespread in the general population, but it also differs markedly across population subgroups. Panel A of Figure 1 reports financial literacy levels for women and men, where we see a sizeable gender gap for each of the financial literacy questions, as well as for all of the Big Three questions combined. Overall, only 29 percent of women answer all three questions correctly, versus 48 percent of men. This gender difference is remarkably stable across topics (Yakoboski, Lusardi, and Hasler 2022). It is also strikingly stable across the 140 countries examined by Klapper and Lusardi (2020). The percentage of those who refuse to answer is normally very small, but as Panel B of Figure 1 shows, women are much more likely than men to respond that they do not know the answer to at least one financial literacy question, especially the one about risk diversification. Such gender differences are likely to be the result of lack of self-confidence, in addition to lack of knowledge (Bucher-Koenen et al. 2021).
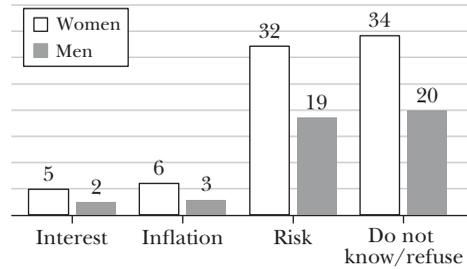
Figure 1
Financial Literacy Differences by Sex

Panel A. Percent of respondents answering
the Big Three questions correctly, by sex
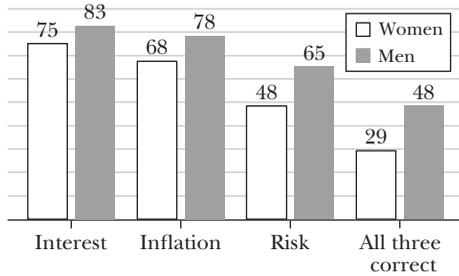
Panel B. Percent of respondents answering the
Big Three questions with at least one
"do not know/refuse," response, by sex

Source: Authors' tabulations, 2019 Survey of Consumer Finance.
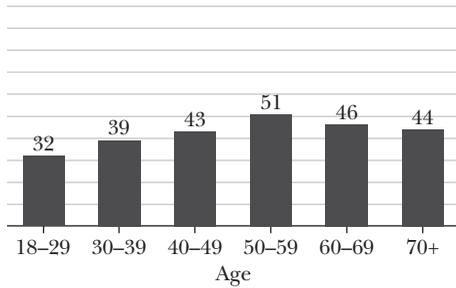*Note:* All data weighted using sample weights.

Financial literacy follows a hump-shaped pattern with age. As Figure 2 shows, young adults display very low financial literacy, with only one-third being able to answer all three questions correctly. Of course, this age group is also making important financial decisions with long-lasting consequences, including taking out student loans, contributing (or not) to a pension, managing credit card debt, buying a home, and raising a family. The hump-shaped financial literacy pattern peaks with only about half answering all questions correctly in the 50–59 age bin. Thereafter, the percentage getting all three answers correct falls. Because these data are cross sectional, one cannot distinguish whether the lower level of financial literacy among older respondents is an age or a cohort effect. For example, cognition may decline with age, or people may not have needed high financial knowledge in the past. Of course, the pattern raises concerns either way, because apparently many elderly Americans are managing their money in retirement with only a limited grasp of basic finance. A very recent longitudinal study (Boyle et al. under review), shows that literacy declines among older adults by 1 percent per year over time, confirming the importance of the age effect.

Given the recent rise in inflation, in panel B of Figure 2 we report the proportion of people by age who know about the loss of purchasing power in the presence of inflation. Clearly, it is the younger respondents who lack knowledge of this topic, as only 65 percent of those under age 40 understand inflation, most likely because they had not experienced it during their economic lifetimes until now. Older generations who lived through the 1970s period of double-digit inflation were more knowledgeable about inflation.

There are also sharp differences by educational levels. While 65 percent of those with college degrees or more earned a perfect score on the Big Three, only 18 percent of the high school dropouts attained this level, as shown in panel A of Figure 3. Nevertheless, even within the college-plus group, more than one-third

Figure 2
Financial Literacy Differences by Age

Panel A. Percent of respondents answering
all Big Three questions correctly, by age

Panel B. Percent of respondents answering
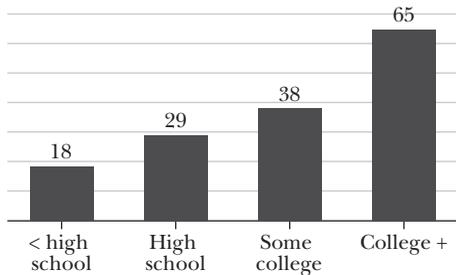the inflation question correctly, by age
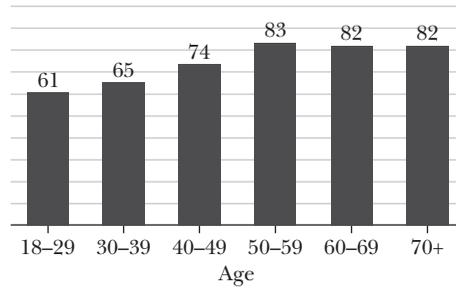
Source: Authors' tabulations, 2019 Survey of Consumer Finance.
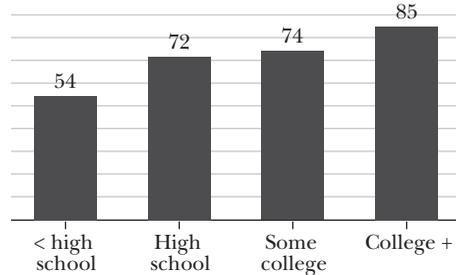*Note:* All data weighted using sampling weights.

Figure 3
Financial Literacy Differences by Education

Panel A. Percent of respondents answering all
Big Three questions correctly, by education

Panel B. Percent of respondents answering the
inflation question correctly, by education

Source: Authors' tabulations, 2019 Survey of Consumer Finance.
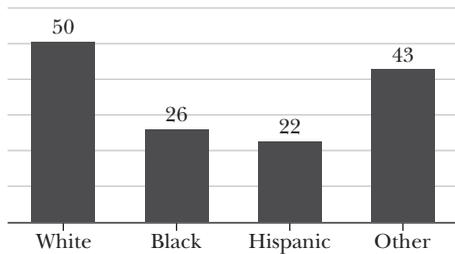*Note:* All data weighted using sampling weights.

of respondents did not know one or more of the Big Three questions. In other words, higher education *per se* is insufficient to instill financial literacy in consumers. This underscores the fact that acquisition of financial know-how requires additional investment not currently part of a general education.

Knowledge of inflation also differs markedly across education groups. Only about half of those without a high school degree know about the eroding power of inflation, versus 85 percent of those with a college or more education, as shown in panel B of Figure 3. This speaks of the potential negative impact of inflation on the finances of the poor.
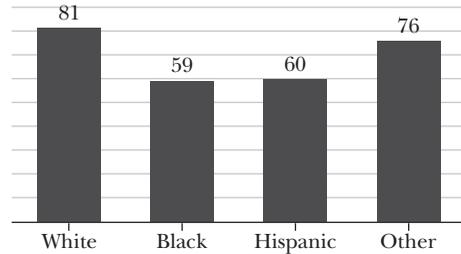
Figure 4
Financial Literacy Differences by Race/Ethnicity

Panel A. Percent of respondents answering all
Big Three questions correctly, by race/ethnicity

Panel B. Percent of respondents answering the
inflation question correctly, by race/ethnicity

*Source:* Authors' tabulations, 2019 Survey of Consumer Finance.
Note: All data weighted using sampling weights.

Sharp differences in financial literacy are also observed by race/ethnicity. Panel A of Figure 4 indicates that half of Whites could correctly answer all three questions, versus only 26 percent of Blacks and 22 percent of Hispanics. This low level of basic financial knowledge among Blacks and Hispanics may help account for the persistence of inequality along financial dimensions such as wealth. Similarly, there are sharp differences in the knowledge of inflation by race/ethnicity, with Blacks and Hispanics being less knowledgeable than Whites, as shown in panel B of Figure 4. Again, lack of knowledge about the loss of purchasing power in the presence of inflation can aggravate financial vulnerability.

These findings are relatively similar across developed nations, providing a robust set of evidence regarding the degrees of difficulty in answering these questions and the quality of the measurement as well (Lusardi and Mitchell 2011c, 2014). A major finding from other studies, shown in Figure 5, is that the level of financial literacy is low even in high-income countries (such as Italy and Japan), as well as those with a strong education system (such as Sweden). Our project on "Financial Literacy around the World" (FLAT World) showed that the world is *flat* indeed in terms of financial literacy, based on our comparisons of knowledge of fundamental concepts across countries (Lusardi and Mitchell 2011c, 2014). In developing countries, financial literacy tends to decline rather than increase with age (Klapper and Lusardi 2020), suggesting that younger cohorts may be acquiring financial literacy over time compared to their older counterparts, whose knowledge is low or has depreciated over time. Nevertheless, the young in emerging economies are still relatively poorly informed when compared to young people in developed economies.

While the Big Three is a strikingly effective indicator, we have also developed more complex measures of financial literacy which offer useful insights (Lusardi and Mitchell 2017; Clark, Lusardi, and Mitchell 2017). One index, now known as

Figure 5
Financial Literacy around the World
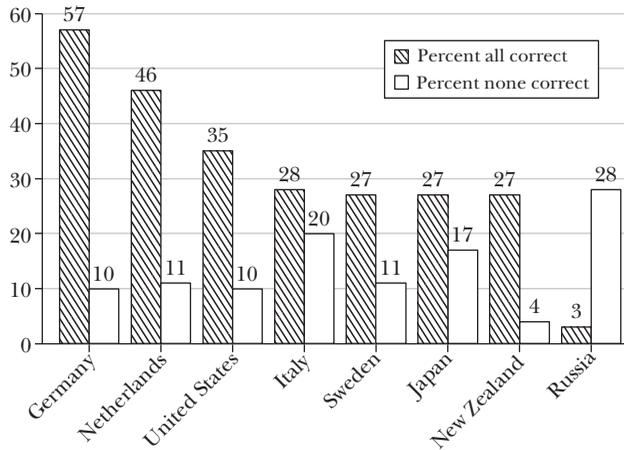
Source: Authors' tabulations, based on data in Lusardi and Mitchell, eds. (2011d).

the "Big Five," added two questions to the original Big Three.[2] These additional questions were originally designed for the 2009 National Financial Capability Study to evaluate financial literacy more broadly, particularly in the wake of the global financial crisis led by subprime financial mortgage market, and to further assess knowledge related to investing. The Personal Finance Index (or P-Fin Index), which started in 2016 and is done at an annual frequency, includes 28 questions (Yakoboski, Lusardi, and Hasler 2022). This survey covers a representative sample of the US population, and one or more subgroups are normally oversampled to

---

[2] See the Big Five in a self-quiz format at https://gflec.org/education/questions-that-indicate-financial-literacy/.
The two additional questions are as follows (correct answers identified with asterisks):
If interest rates rise, what will typically happen to bond prices?
  A) They will rise
  B) They will fall**
  C) They will stay the same
  D) There is no relationship between bond prices and the interest rate
  E) Don't know
  F) Prefer not to say
A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.
  A) True**
  B) False
  C) Don't know
  D) Prefer not to say
See the review of our questions in Hastings, Madrian, and Skimmyhorn (2013).

provide additional information. The questions in the P-Fin Index are divided into eight broad topics, and risk management is once again the category where people score the lowest (and there is a high proportion of "do not know" answers).

The P-Fin Index also shows that, overall, financial literacy has not changed much in the past six years (2017–2023), yet financial literacy has grown more unequal over time, in that the more financially literate subgroups in 2022 were also the most financially literate in 2017 (for example, men and the better-educated). The somewhat good news is that there was also an improvement in financial literacy scores among the young, possibly due to more US states having recently mandated financial education in school (Urban et al. 2015). These more extensive measures of financial literacy confirm the broad conclusions provided by the Big Three.

We have focused considerable attention on measurement of financial literacy, in the belief that good measurement is critically important to understand a new topic and identify potential problems. Indeed, this effort has propelled the opening of the new field of financial literacy.

How Financial Literacy Shapes Financial Decision-Making

In this section, we investigate the effects of financial literacy on financial decision-making. Our early work using the Health and Retirement Survey showed that financial literacy is a strong predictor of retirement planning and wealth (Lusardi and Mitchell 2007, 2011a). We have replicated this finding in many other US datasets, as well as in several other countries (for example, Behrman et al. 2012; Hastings and Mitchell 2020; Lusardi and Mitchell 2014).

Our research has also shown that financial literacy produces better investment outcomes. For example, the more financially literate are more likely to invest in the stock market, and hence, earn higher (risk-adjusted) returns on their investments (van Rooij, Lusardi, and Alessi 2011; Clark, Lusardi, and Mitchell 2017). Stock market participation can be a conduit to higher wealth and, potentially for society as a whole, to greater wealth inequality, as we discuss below. Financial literacy also shapes the liability side of the balance sheet. Debt has risen across generations in the United States, and people are increasingly carrying debt well into retirement (Mitchell and Lusardi 2020). People who are more financially literate are also better able to manage this debt (Lusardi, Mitchell, and Oggero 2020).

Recent analysis has also sought to explain some of the observed demographic differences in financial literacy noted above. For example, several studies have sought to explain why women tend to be less financially literate than men, including from stereotypes and household specialization of labor (Hsu 2016; Cupák et al. 2018; Driva, Lührmann, and Winter 2016). Other work has focused on differences by race/ethnicity, with candidate explanations including differential maternal education and lack of financial exposure in the home (see, for instance, Angrisani et al. 2021). Other studies, including Okamoto and Komamura (2021) and Finke, Howe, and Huston (2017), evaluate the links between financial literacy and age.

In Table 3, we report coefficient estimates from median regressions of three wealth measures, as well as the wealth/income ratio, using our measure of financial literacy as the explanatory variable. Our goal in this analysis is simply to show the associations between financial literacy and wealth. Additional control variables include age, gender, education, marital status, employment status, race, and income, which are all factors that can proxy for other determinants of wealth. The first column of panel A indicates that one additional correct answer on the financial literacy (FinLit) measure is associated with 13 percent higher median net wealth, 24 percent higher financial wealth, and 7 percent more nonfinancial wealth. Moreover, the median wealth/income ratio is higher by 15 percent, holding other factors constant. Estimates are even larger when using the Big Three in Panel B. Both sets of results underscore the powerful positive association between financial literacy and wealth.

The observation that financial literacy is associated with wealth led us to develop a theoretical model to examine the impact of financial literacy on wealth and wealth inequality. In particular, our life-cycle model embeds several types of uncertainty regarding labor income, out-of-pocket medical expenses, and asset returns, as well as borrowing constraints and other features of the economy (Lusardi, Michaud, and Mitchell 2017). Of central interest is the role of financial literacy, where we posit that becoming financially knowledgeable requires people to expend time and money. In turn, this investment permits them to reap the benefits of having access to a better saving technology. Depreciation of knowledge is also a factor, so consumers must decide whether to keep investing in financial literacy or to let their knowledge decline with time. In this setup, financial literacy becomes an endogenous decision variable: people choose their financial knowledge optimally by comparing the costs and benefits of doing so.

Our theoretical framework is useful for several reasons. First, it rationalizes some of the financial literacy facts reported above. For example, financial literacy is predicted to be low among the young, but it should rise with age as people start investing in financial literacy. At some point, it can be optimal to let knowledge depreciate, generating the observed hump-shaped pattern.[3] Second, our model predicts that financial literacy will be higher for the better-educated, as this group is more likely to need to save more for retirement, compared to the lower-paid who receive relatively higher Social Security replacement rates. Third, this economic model makes it clear that people can be perfectly rational and yet choose not to be particularly financially sophisticated. Fourth, understanding the causal impact of financial literacy on wealth must take into account the fact that financial literacy is an endogenous variable. While the results in Table 2 are reported for descriptive reasons, different estimation strategies are required to assess the effect of financial literacy on wealth, as summarized in Lusardi and Mitchell (2014).

To this end, we use our theoretical model to assess the impact of financial literacy, not just on wealth, but also on wealth inequality. We document that

---

[3]In addition, of course, people may also elect to delegate money management decisions to financial advisers in lieu of devoting their own time and effort to acquiring the knowledge (Kim, Maurer, and Mitchell 2021).

Table 3

Estimated Coefficients from Median Regressions of Financial Literacy on Various Wealth Measures

	Net wealth ($100k)	Financial wealth ($100k)	Non-financial wealth ($100k)	Wealth/income ratio
A. Using FinLit index				
FinLit index	0.161***	0.059***	0.121***	0.305***
	(0.036)	(0.010)	(0.022)	(0.046)
Age	0.040***	0.008***	0.019***	0.082***
	(0.003)	(0.001)	(0.002)	(0.003)
Female	0.195**	0.078***	0.069	−0.145
	(0.068)	(0.024)	(0.047)	(0.099)
Black	−0.383***	−0.121***	−0.304***	−1.081***
	(0.092)	(0.034)	(0.052)	(0.080)
Hispanic	0.197*	−0.043	0.011	−0.171
	(0.091)	(0.030)	(0.076)	(0.119)
Race, others	−0.052	−0.076	−0.056	−0.469***
	(0.105)	(0.043)	(0.061)	(0.099)
<High school	0.078	0.024	−0.526***	−1.370***
	(0.136)	(0.042)	(0.071)	(0.186)
High school	0.009	−0.024	−0.467***	−0.823***
	(0.119)	(0.044)	(0.076)	(0.114)
Some college	−0.073	−0.096*	−0.378***	−0.797***
	(0.086)	(0.037)	(0.055)	(0.098)
Pseudo R^2	0.241	0.167	0.240	0.138
Median of dep. var.	1.215	0.250	1.643	1.982
B. Using all Big Three correct				
All Big Three correct	0.546***	0.206***	0.256***	0.725***
	(0.065)	(0.024)	(0.048)	(0.066)
Age	0.038***	0.008***	0.019***	0.084***
	(0.002)	(0.001)	(0.002)	(0.003)
Female	0.204**	0.099***	0.076	−0.104
	(0.066)	(0.027)	(0.044)	(0.087)
Black	−0.355***	−0.128***	−0.297***	−1.072***
	(0.093)	(0.036)	(0.054)	(0.093)
Hispanic	0.228**	−0.019	0.027	−0.102
	(0.087)	(0.030)	(0.079)	(0.103)
Race, other	−0.023	−0.074	−0.065	−0.449***
	(0.091)	(0.041)	(0.068)	(0.081)
<High school	0.176	0.045	−0.517***	−1.426***
	(0.130)	(0.045)	(0.067)	(0.169)
High school	0.053	−0.006	−0.437***	−0.753***
	(0.114)	(0.044)	(0.074)	(0.093)
Some college	−0.061	−0.083*	−0.354***	−0.741***
	(0.092)	(0.039)	(0.059)	(0.090)
Pseudo R^2	0.242	0.168	0.240	0.140
Median of dep. var.	1.215	0.250	1.643	1.982

Source: Authors' tabulations, 2019 Survey of Consumer Finance. All data weighted using sampling weights.
Notes: In both panels, median regression models also control on marital status, number of children at home, employment status, and income; reference groups are White, college+, never married, and working. N = 5,777. *p<0.10; **p<0.05; ***p<0.01.

financial literacy is indeed not a sideshow; instead, in the United States, we show that 30–40 percent of wealth inequality near retirement can be accounted for by financial literacy (Lusardi, Michaud, and Mitchell 2017). In other words, financial literacy matters, and it matters a great deal. Accordingly, it is time to incorporate financial literacy into standard intertemporal models of saving behavior, because leaving it out omits a key explanation for why wealth holdings are so heterogeneous.

Adding financial literacy into calibrated intertemporal models of saving is one important way to measure its effects. Another is to seek sources of exogenous variation in access or opportunities to become financially literate. This includes, for example, financial education mandates in school or initiatives in the workplace, and studying the effects of specific programs.[4]

Empirical evaluations are most credible when they result from randomized control trials, as these can avoid issues resulting from more able people self-selecting into the programs of interest. A recent meta-analysis of financial education programs (Kaiser et al. 2022) concentrated exclusively on such randomized control trials (including some programs based on our own financial education initiatives). After reviewing 76 financial education programs across 33 countries in six continents, three major conclusions emerge. First, financial education positively affects both financial knowledge and behavior. Second, the impact is three to five times larger than found by older studies (Fernandes et al. 2014), mainly due to the inclusion of many more recent efforts to improve financial literacy and perhaps awareness of the low levels of financial literacy we were able to document. Third, using the scale by Kraft (2020), financial education programs are found to be cost-effective and the effects are similar to other education programs, such as those related to health, energy, or the environment.

Discussion

Research on financial literacy has grown exponentially in recent decades. If one searches for the term "financial literacy" in the Social Science Citations Index (now published as the Clarivate Web of Science™), there were essentially no citations from 1994 to 2004. As recently as 2011, this search uncovered fewer than 500 papers per year. But by 2015, the total number of papers on the topic topped 1,500, and by 2018 it exceeded 3,000 (Kaiser et al. 2022). This expansion underscores the widespread recognition of the new field of financial literacy. This has several implications for research, teaching, and policy and programs.

Implications for Research

After 20 years of research on this topic, financial literacy has now become an official field of study in the economics profession, with its own JEL code (G53). We now have metrics for measuring financial literacy that range from a few to

---

[4]For instance, see Bernheim, Garrett, and Maki (2001), Bernheim and Garrett (2003), Frisancho (2023), Bruhn et al. (2016), and Urban et al. (2015).

many questions, which have been adopted across the world, and which are good predictors of financial behavior. Moreover, we have shown that financial literacy can be incorporated into standard intertemporal models of financial behavior; in fact, not doing so will limit our understanding of the determinants of wealth. A new academic journal, the *Journal of Financial Literacy and Wellbeing*, published by Cambridge University Press, provides further evidence of the expansion of this field. The number and sophistication of financial literacy programs is growing rapidly around the world;[5] in particular, many have moved beyond very short interventions—such as a single retirement seminar or sending employees to a benefits fair—to more robust designs. An invaluable next step would seek to understand both the short-term and long-term costs and benefits of such treatments (compare discussions in Bruhn et al. 2023; Frisancho 2023).

We also recognize that it may not always be optimal for people's behavior to change, as our theoretical model indicated (Lusardi, Michaud, and Mitchell 2017, 2020). In fact, the best response may be sometimes to do nothing, something that evaluation programs should take into account to provide an accurate assessment of the impact of financial education.

Implications for Teaching

Early research on financial education confirmed the importance of having it in school (Bernheim, Garrett, and Maki 2001), and now that the crucial role of financial literacy skills has become widely recognized, a variety of personal finance classes is offered in colleges and universities, particularly in the United States. We started offering such courses at our own universities in 2013, motivated and guided by our research findings.[6] Of course, business schools regularly teach corporate finance to aspiring chief financial officers, but it is now clear that regular consumers must also be exposed to rigorous preparation if they are to successfully manage their money, save and invest properly, and decumulate their assets in retirement; in other words, they must become their own chief financial officers.

Our research also provides insight into what such courses should teach. For example, the Big Three and the P-Fin Index tell us that most people fail to grasp key fundamental financial concepts, particularly financial risk management. Moreover, people must make decisions that require them to know about specific financial instruments and contracts; they must also be aware of their rights and obligations in the financial marketplace. Several years ago, the Council for Economic Education (2013) established National Standards for Financial Literacy, detailing what should

---

[5] Moreover, there are now numerous seminar series, conferences, and major economics and finance meetings dedicated to financial literacy. For example, researchers regularly present their work at the annual US conference jointly organized by the US Board of Governors of the Federal Reserve System and the Global Financial Literacy Excellence Center. The Mannheim Institute for Financial Education has also introduced European researchers to the topic, and the European Finance Association annual meetings have tracks related to financial literacy. The OECD also holds conferences and symposia related to financial literacy at a semi-annual frequency.

[6] We teach courses at both the undergraduate and graduate level. Our course syllabi are available in the online Appendix.

be covered in personal finance courses in US schools. We have extended those standards, borrowing from both the theory and the evidence on financial literacy, to design our own personal finance courses. Lusardi, together with John Shoven and Michael Boskin from Stanford University, also organized the first academic conference for the economics profession dedicated to the teaching of personal finance, held in September 2022 at Stanford University.[7] The most recent meta-analysis of financial education programs also found that financial education had substantial positive impacts on financial knowledge in both developing and in developed countries, and versions of financial literacy courses can now be found everywhere (Kaiser et al. 2022).

Some US states have also mandated financial literacy in high schools, a step which will help the young acquire basic financial skills (widen the access and reduce the cost), and also avoid getting into financial trouble early in life (Urban et al. 2015; Barua, Koh, and Mitchell 2017). Not surprisingly, financial education is most effective when there is a rigorous curriculum, a specific course devoted to personal finance (rather than embedding these concepts into other classes), and trained teachers (Tennyson and Nguyen 2005). Of course, it is also critical to reinforce this training periodically, to offset the effect of human capital depreciation (Lusardi, Michaud, and Mitchell 2020).

Implications for Policy and Programs

This research also has implications for policy and programs. Given low global levels of financial literacy, it will be critical to step up the effort to improve financial knowledge. Indeed, more than 80 countries have set up national committees entrusted with the design and implementation of national strategies for financial literacy (OECD/INFE 2017); one of us chaired the Financial Education Committee charged with improving financial literacy in Italy. Several nations have already established ambitious and innovative programs. New Zealand has been a pioneer in building a national website for financial education, and today, many nations dedicate a week or month to financial literacy (for example, April in the United States, November in Canada, and October in Italy). This has become a powerful way to increase awareness about the importance of financial literacy/education.[8] Since 2012, the Programme for International Student Assessment (PISA) run by the OECD has added financial literacy to the set of topics that 15-year-old students need to know to participate in modern society and succeed in the labor market, evidence that financial literacy is now considered an essential skill, like reading, writing, and knowledge of science.[9] Additionally, the European Commission (2020) has acknowledged the importance of financial literacy as a key step for a capital markets union, and it has now collected data on financial literacy similar to the Big Three across the 27 EU nations (European Commission 2023). Policy leaders

---

[7] More information about this initiative is provided at https://personalfinanceteaching.org/.
[8] For individual programs receiving awards, see https://maiawards.org/.
[9] Lusardi led the team that designed these questions, and concepts relating to the Big Three are included in the assessment.

including António Guterres (2022), Secretary General of the United Nations, have noted that financial skills must be part of the toolkit that all young people must acquire to make good decisions about their economic futures.

Because acquiring financial knowledge is a lifelong process and the crucial financial challenges vary by age, financial education must also be provided after people leave high school, The workplace is one prominent setting: Clark (2023) reviews the literature on the importance of workplace financial education, showing its relevance across types of employers and sectors. Analysis of P-Fin Index data shows that Americans spend an average of seven hours a week dealing with their personal finance issues, and three of these hours occur during worktime. Moreover, the least financially literate are four times more likely to spend ten or more hours per week thinking about and dealing with issues and problems related to personal finances (Hasler et al. 2023; Yakoboski, Lusardi, and Hasler 2022). Given that the costs of financial education programs need not be high, employers are increasingly finding it beneficial to provide financial education for their employees.

Moreover, as there are large differences in financial knowledge across people, it will be critical to provide tailored programs so as to best address the needs of specific subgroups. For example, some of the observed gender differences in financial literacy may be due not only to knowledge, but also to self-confidence. Programs targeting financial literacy for women could therefore try to promote both.

Poor financial literacy also has negative externalities reaching beyond households themselves. Consumers who fail to understand risk and risk management may underinsure (Brown, Kapteyn, and Mitchell 2016; Brown et al. 2017, 2021; Gottlieb and Mitchell 2020). Families can suffer if they lack buffer savings to hedge even small shocks, much less against economic crises or the recent pandemic (Demerztis, Domínguez-Jiménez, and Lusardi 2020; Hurwitz, Mitchell, and Sade 2021). In volatile economic times, if people do not understand inflation or the power of interest compounding, they may not set aside enough for precautionary reasons and retirement. To reduce the political demand for taxpayer support of such families, it could be less costly to try to prevent these behaviors.

Policymakers may also gain an additional benefit from promoting financial literacy. For example, the financially illiterate are less able to appreciate pension reforms, and in turn, to vote for them (Fornero and Lo Prete 2023). Finally, since financial literacy also has such important implications for wealth and wealth inequality, investing in financial literacy may help reduce gaps that arise because of unequal access to financial education.

Conclusions

People who wish to save for retirement, budget, invest for the long term, and draw down their assets in old age can access a wide range of resources, yet they are often unable to make good decisions about the broad range of financial choices they face. Robust interventions are needed to address the persistently low and widespread lack of financial literacy. Indeed, the topic of financial literacy is ripe to be

integrated into syllabi, textbooks, and microeconomic as well as macroeconomic courses at both the undergraduate and graduate levels. Moreover, teaching personal finance in both high school and college is an ideal way to make financial education more widely accessible, while workplace-based programs also have a role to play.

We propose that measures of financial literacy be explicitly added to national statistics indicators, as part of an overall picture of a country's wellbeing, together with data on consumption and savings. Indeed, financial literacy can be an important policy target; for example, Finland recently launched a national strategy for financial literacy, pledging to become the country with the highest level of financial literacy in the world by 2030 (Raijas 2021).

Finally, economic models can and should be amended to incorporate the fact that many people lack the necessary knowledge to participate in financial markets effectively and to use financial instruments properly. Inasmuch as we make financial decisions with potentially lifelong consequences every day, the field of financial literacy has become an integral component of economics research and teaching, and a valuable tool for policymakers.

References

Angrisani, Marco, Sergio Barrera, Luisa R. Blanco, and Salvador Contreras. 2021. "The Racial/Ethnic Gap in Financial Literacy in the Population and by Income." *Contemporary Economic Policy* 39 (3): 524–36.

Barua, Rashmi, Benedict Koh, and Olivia S. Mitchell. 2017. "Does Financial Education Enhance Financial Preparedness? Evidence from a Field Experiment in Singapore." *Journal of Pension Economics and Finance* 17 (3): 254–77.

Behrman, Jere R., Olivia S. Mitchell, Cindy K. Soo, and David Bravo. 2012. "How Financial Literacy Affects Household Wealth Accumulation." *American Economic Review* 102 (3): 300–304.

Bernheim, B. Douglas, and Daniel M. Garrett. 2003. "The Effects of Financial Education in the Workplace: Evidence from a Survey of Households." *Journal of Public Economics* 87 (7–8): 1487–1519.

Bernheim, B. Douglas, Daniel M. Garrett, and Dean M. Maki. 2001. "Education and Saving: The Long-Term Effects of High School Financial Curriculum Mandates." *Journal of Public Economics* 80 (3): 435–65.

Bhutta, Neil, Jesse Bricker, Andrew C. Chang, Lisa J. Dettling, Sarena Goodman, Joanne W. Hsu, Kevin B. Moore, Sarah Reber, Alice Henriques Volz, and Richard A. Windle. 2020. "Changes in U.S. Family Finances from 2016 to 2019: Evidence from the Survey of Consumer Finances." *Federal*

Reserve Bulletin 106 (5): 1–42.

Boyle, Patricia, Olivia S. Mitchell, Gary Mottola, and Lei Yu. 2023. Declining Financial and Health Literacy among Older Men and Women. Unpublished.

Brown, Jeffrey R., Arie Kapteyn, Erzo F. P. Luttmer, and Olivia S. Mitchell. 2017. "Cognitive Constraints on Valuing Annuities." *Journal of the European Economic Association* 15 (2): 429–62.

Brown, Jeffrey R., Arie Kapteyn, Erzo F. P. Luttmer, Olivia S. Mitchell, and Anya Samek. 2021. "Behavioral Impediments to Valuing Annuities: Complexity and Choice Bracketing." *Review of Economics and Statistics* 103 (3): 533–46.

Brown, Jeffrey R., Arie Kapteyn, and Olivia S. Mitchell. 2016. "Framing and Claiming: How Information-Framing Affects Expected Social Security Claiming Behavior." *Journal of Risk and Insurance* 83 (1): 139–62.

Bruhn, Miriam, Luciana de Souza Leão, Arianna Legovini, Rogelio Marchetti, and Bilal Zia. 2016. "The Impact of High School Financial Education: Evidence from a Large-Scale Evaluation in Brazil." *American Economic Journal: Applied Economics* 8 (4): 256–95.

Bruhn, Miriam, Gabriel Garber, Sergio Koyama, and Bilal Zia. 2022. "The Long-Term Impact of High School Financial Education: Evidence from Brazil." Banco Central do Brasil Working Paper 563.

Bucher-Koenen, Tabea, Rob J. Alessie, Annamaria Lusardi, and Maarten van Rooij. 2021. "Fearless Woman: Financial Literacy and Stock Market Participation." NBER Working Paper 28723.

Clark, Robert L. 2023. "Effectiveness of Employer-Provided Retirement Planning Programs." *Journal of Financial Literacy and Wellbeing* 1 (1): 154–68.

Clark, Robert L., Annamaria Lusardi, and Olivia S. Mitchell. 2017. "Financial Knowledge and 401(k) Investment Performance." *Journal of Pension Economics and Finance* 16 (3): 324–47.

Clark, Robert L., Annamaria Lusardi, and Olivia S. Mitchell. 2021. "Financial Fragility during the COVID-19 Pandemic." *AEA Papers and Proceedings* 111: 292–96.

Clark, Robert, Annamaria Lusardi, Olivia S. Mitchell, and Hallie Davis. 2021. "Factors Contributing to Financial Well-Being among Black and Hispanic Women." *Journal of Retirement* 9 (1): 71–97.

Council for Economic Education. 2013. *National Standards for Financial Literacy.* New York: Council for Economic Education.

Cupák, Andrej, Pirmin Fessler, Alyssa Schneebaum, and Maria Silgoner. 2018. "Decomposing Gender Gaps in Financial Literacy: New International Evidence." *Economics Letters* 168: 102–06.

Demertzis, Maria, Marta Domínguez-Jiménez, and Annamaria Lusardi. 2020. "The Financial Fragility of European Households in the Time of COVID-19." Bruegel Policy Contribution 2020/15.

Driva, Anastasia, Melanie Lührmann, and Joachim Winter. 2016. "Gender Differences and Stereotypes in Financial Literacy: Off to an Early Start." *Economics Letters* 146: 143–46.

European Commission. 2020. "Capital Markets Union 2020 Action Plan: A Capital Market Union for People and Businesses." https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/capital-markets-union/capital-markets-union-2020-action-plan_en.

European Commission. 2023. "Eurobarometer Survey Reveals Low Levels of Financial Literacy across the EU." European Commission, July 18. https://finance.ec.europa.eu/news/eurobarometer-survey-reveals-low-levels-financial-literacy-across-eu-2023-07-18_en.

Fernandes, Daniel, John G. Lynch Jr., and Richard G. Netemeyer. 2014. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." *Management Science* 60 (8): 1861–83.

Finke, Michael, John S. Howe, and Sandra J. Huston. 2017. "Old Age and the Decline in Financial Literacy." *Management Science* 63 (1): 213–30.

Fornero, Elsa, and Anna Lo Prete. 2023. "Financial Education: From Better Personal Finance to Improved Citizenship." *Journal of Financial Literacy and Wellbeing* 1 (1): 12–27.

Frisancho, Veronica. 2023. "Is School-Based Financial Education Effective? Immediate and Long-Lasting Impact on High School Students." *Economic Journal* 133 (651): 1147–80.

Gottlieb, Daniel, and Olivia S. Mitchell. 2020. "Narrow Framing and Long-Term Care Insurance." *Journal of Risk and Insurance* 87 (4): 861–93.

Guterres, António. 2022. "Transforming Education: An Urgent Political Imperative for our Collective Future." Speech. United Nations Transforming Education Summit. *https://www.un.org/en/transforming-education-summit/sg-vision-statement.*

Hasler, Andrea, Annamaria Lusardi, Nikhil Yagnik, and Paul Yakoboski. 2023. "Resilience and Wellbeing in the Midst of the COVID-19 Pandemic: The Role of Financial Literacy." *Journal of Accounting and Public Policy* 42 (2): 107079.

Hastings, Justine S., Brigitte C. Madrian, and William L. Skimmyhorn. 2013. "Financial Literacy, Financial Education, and Economic Outcomes." *Annual Review of Economics* 5 (1): 347–73.

Hastings, Justine, and Olivia S. Mitchell. 2020. "How Financial Literacy and Impatience Shape Retirement Wealth and Investment Behaviors." *Journal of Pension Economics and Finance* 19 (1): 1–20.

Hsu, Joanne W. 2016. "Aging and Strategic Learning: The Impact of Spousal Incentives on Financial Literacy." *Journal of Human Resources* 51 (4): 1036–67.

Hurwitz, Abigail, Olivia S. Mitchell, and Orly Sade. 2021. "Longevity Perceptions and Saving Decisions during the COVID-19 Outbreak: An Experimental Investigation." *AEA Papers and Proceeding* 111: 297–301.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban. 2022. "Financial Education Affects Financial Knowledge and Downstream Behaviors." *Journal of Financial Economics* 145 (2A): 255–72.

Kim, Hugh H., Raimond Maurer, and Olivia S. Mitchell. 2021. "How Financial Literacy Shapes the Demand for Financial Advice at Older Ages." *Journal of the Economics of Ageing* 20: 100329.

Klapper, Leora, and Annamaria Lusardi. 2020. "Financial Literacy and Financial Resilience: Evidence from Around the World." *Financial Management* 49 (3): 589–614.

Kraft, Matthew A. 2020. "Interpreting Effect Sizes of Education Interventions." *Educational Researcher* 49 (4): 241–53.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell. 2017. "Optimal Financial Knowledge and Wealth Inequality." *Journal of Political Economy* 125 (2): 431–77.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell. 2020. "Assessing the Impact of Financial Education Programs: A Quantitative Model." *Economics of Education Review* 78: 101899.

Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics* 54 (1): 205–24.

Lusardi, Annamaria, and Olivia S. Mitchell. 2011a. "Financial Literacy and Planning: Implications for Retirement Well-Being." In *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, edited by Olivia S. Mitchell and Annamaria Lusardi, 17–39. Oxford, UK: Oxford University Press.

Lusardi, Annamaria, and Olivia S. Mitchell. 2011b. "Financial Literacy and Retirement Planning in the United States." *Journal of Pension Economics and Finance* 10 (4): 509–25.

Lusardi, Annamaria, and Olivia S. Mitchell. 2011c. "Financial Literacy around the World: An Overview." *Journal of Pension Economics and Finance* 10 (4): 497–508.

Lusardi, Annamaria, and Olivia S. Mitchell, eds. 2011d. "Special Issue on Financial Literacy around the World." *Journal of Pension Economics and Finance* 10 (4): 471–630.

Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52 (1): 5–44.

Lusardi, Annamaria, and Olivia S. Mitchell. 2017. "How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness." *Quarterly Journal of Finance* 7 (3): 1–31.

Lusardi, Annamaria, and Olivia S. Mitchell. 2023. "Replication data for: The Importance of Financial Literacy: Opening and New Field." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193204V1.

Lusardi, Annamaria, Olivia S. Mitchell, and Noemi Oggero. 2020. "Debt and Financial Vulnerability on the Verge of Retirement." *Journal of Money, Credit, and Banking* 52 (5): 1005–34.

Mitchell, Olivia S., and Annamaria Lusardi eds. 2020. *Remaking Retirement: Debt in an Aging Economy.* Oxford, UK: Oxford University Press.

OECD/INFE. 2017. *G20/OECD INFE Report on Adult Financial Literacy in G20 Countries.* Paris: OECD.

Okamoto, Shohei, and Kohei Komamura. 2021. "Age, Gender, and Financial Literacy in Japan." *PLoS ONE* 16 (11): e0259393.

Raijas, Anu. 2021. "The First National Strategy of Financial Education in Finland." Presentation to the Financial Literacy Seminar Series, George Washington University. https://gflec.org/wp-content/uploads/2021/09/Presentation.pdf.

Tennyson, Sharon, and Chau Nguyen. 2005. "State Curriculum Mandates and Student Knowledge of Personal Finance." *Journal of Consumer Affairs* 35 (2): 241–62.

Urban, Carly, Maximilian Schmeiser, J. Michael Collins, and Alexandra Brown. 2015. *State Financial Education Mandates: It's All in the Implementation.* Washington, DC: FINRA Investor Education Foundation.

van Rooij, Maarten, Annamaria Lusardi, and Rob Alessi. 2011. "Financial Literacy and Stock Market Participation." *Journal of Financial Economics* 101 (2): 449–72.

Yakoboski, Paul J., Annamaria Lusardi, and Andrea Hasler. 2022. *How Financial Literacy Varies among U.S. Adults.* TIAA Institute GFLEC Personal Finance Index. New York: TIAA Institute.

Transmission Impossible? Prospects for Decarbonizing the US Grid

Lucas W. Davis, Catherine Hausman, and Nancy L. Rose

Encouraged by the declining cost of grid-scale wind and solar, recent analyses conclude that the United States could reach net zero[1] carbon dioxide emissions by 2050 at relatively low cost using currently available technologies (Princeton 2021; Williams et al. 2021; National Academies 2021). For example, Williams et al. (2021) find that the United States could reach net zero carbon dioxide emissions by 2050 at a cost of about $1 per person per day. These scenarios rely on, for instance, electrifying vehicles and home heating while transitioning the electricity grid to zero-carbon sources.

While the cost of renewable *generation* has declined dramatically, we focus in this paper on the large expansion in *transmission* that would be required. We first document recent US investments in renewables, then examine some of the issues already emerging in US electricity markets due to insufficient transmission capacity. For example, in some areas, wholesale electricity prices are now negative during more than 20 percent of all hours. While the United States is building more

■ *Lucas Davis is Jeffrey A. Jacobs Distinguished Professor at the Haas School of Business, University of California—Berkeley, Berkeley, California. Catherine Hausman is Associate Professor, Gerald R. Ford School of Public Policy, University of Michigan, Ann Arbor, Michigan. Nancy L. Rose is Charles P. Kindleberger Professor of Applied Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts. Their email addresses are lwdavis@berkeley.edu, chausman@umich.edu, and nrose@mit.edu.*

[1] Net zero carbon refers to the recognition that to the extent that some CO_2 producing activities continue—whether continued burning of some carbon-based fuels or industrial or agricultural processes that emit CO_2 or other greenhouse gases as by-products—there will be a need for sufficient offsetting carbon capture and sequestration.

transmission, the current pace of investment is well below what would be required for the net zero future.

We describe several challenges which make it difficult to build new transmission. The US electricity grid is a disorganized patchwork that is the result of over a century of mostly disconnected individual utilities making independent decisions. There is no central authority for approving new transmission projects, so typically new projects must be approved by a combination of federal, state, and local authorities, and it can be hard to achieve consensus. Moreover, even when stakeholders agree on the need for transmission, there are disagreements about how to pay for project costs. In addition, siting and permitting challenges and NIMBY ("not in my back yard") concerns make it expensive and time-consuming to negotiate right-of-way permissions.

Finally, we describe potential policy responses. The public good characteristics of electricity transmission provide an economic argument for enhanced federal authority over siting decisions. We also point to the potential for increasing the capacity of existing transmission corridors. We then discuss potential substitutes, including storage and dynamic pricing. Neither of these is a panacea, but the challenges of expanding transmission capacity imply that the benefits from these substitutes are higher than they would be otherwise.

Throughout the paper, we focus on the United States, though both the need for transmission and the barriers to transmission expansion exist in other countries. The United States is neither the most ambitious country with regard to decarbonization nor does it have the best options for renewable resources. But it is a significant contributor to greenhouse gas emissions and a valuable case study for understanding the complicated challenges of building new transmission as well as for testing potential policy responses.

## The Role of Renewable Generation in a Net Zero Carbon Future

Renewable electricity generation plays an outsized role in all scenarios for weaning economies off fossil fuels and moving to a net zero carbon future. This has broad implications for electricity system investments.

The Decline in Renewable Costs

The last decade has seen a dramatic decrease in the cost of grid-scale wind and solar generation.[2] Investment in these technologies grew rapidly during this period, and economies of scale, learning-by-doing, and other factors resulted in large cost declines. With wind power, one of the biggest changes was a move toward

---

[2]We focus throughout on grid-scale wind and solar generation, as opposed to rooftop solar or other distributed generation. The costs of distributed renewable generation (including residential and commercial rooftop solar) have fallen, but they remain well above the costs of grid-scale renewables (NREL 2022), and the decarbonization scenarios mentioned earlier rely on grid-scale rather than distributed generation (National Academies 2021).

Figure 1
Decreasing Cost of Grid-Scale Renewables
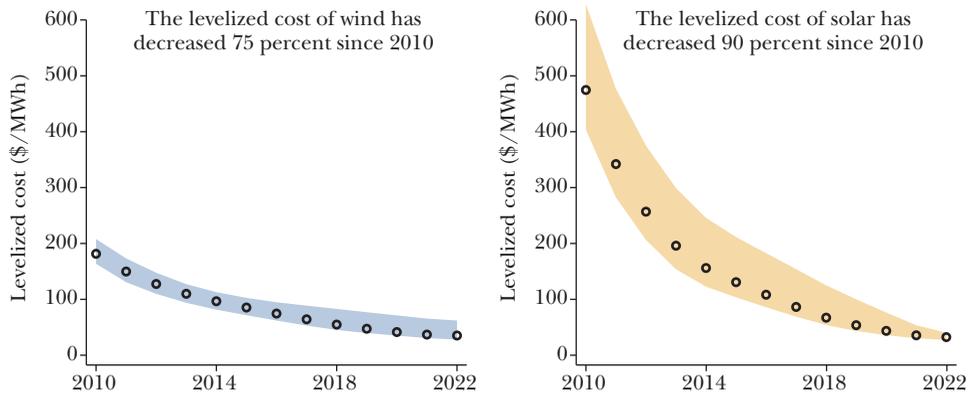
Source: This figure was created by the authors using levelized costs calculations from the US Energy Information Administration (2010–2022), and reflects lifetime project costs including construction, financing, and operations.
Note: The circles indicate the US average levelized cost in each year without tax credits for onshore wind and solar photovoltaics. The range indicates regional variation. A small amount of smoothing has been applied to emphasize the overall pattern rather than idiosyncratic year-to-year fluctuations. All values in the paper have been deflated to reflect year 2022 dollars.

much larger turbines (Covert and Sweeney 2022). With solar photovoltaics, the cost declines were the result of a series of incremental improvements in the manufacturing process, including better and more automated manufacturing equipment, supply chain optimization, and more efficient use of materials (Nemet 2019).

Figure 1 plots typical costs for US grid-scale wind and solar. In 2010, electricity generation from wind and solar cost $200 and $500 per megawatt-hour, respectively. Costs of generation declined sharply initially and then continued to decline throughout the period, falling below $40 by the end of the period. Between 2010 and 2022, costs declined 75 percent for wind and 90 percent for solar, such that today wind and solar are on par or cheaper than fossil fuels using a levelized cost basis— that is, a present discounted value of the costs over the lifetime production of the electricity-generating investment. In a 2022 report from the US Energy Information Administration (2010–2022), levelized costs of generating electricity per megawatt-hour are $38 for wind and $36 for solar, compared to $37 for "combined cycle" natural gas generation. For more details on levelized costs, a useful starting point is Borenstein (2012), and see Joskow (2011) on the challenges of comparing levelized costs across technologies.

The Need for Continued Capacity Growth in Renewable Energy Generation

Encouraged by these declining costs and concern about climate change, the United States is investing heavily in wind and solar. Figure 2 plots the percentage of total US electricity generation that comes from grid-scale wind and solar, from

Figure 2
Growing Percentage of US Electricity from Grid-Scale Renewables
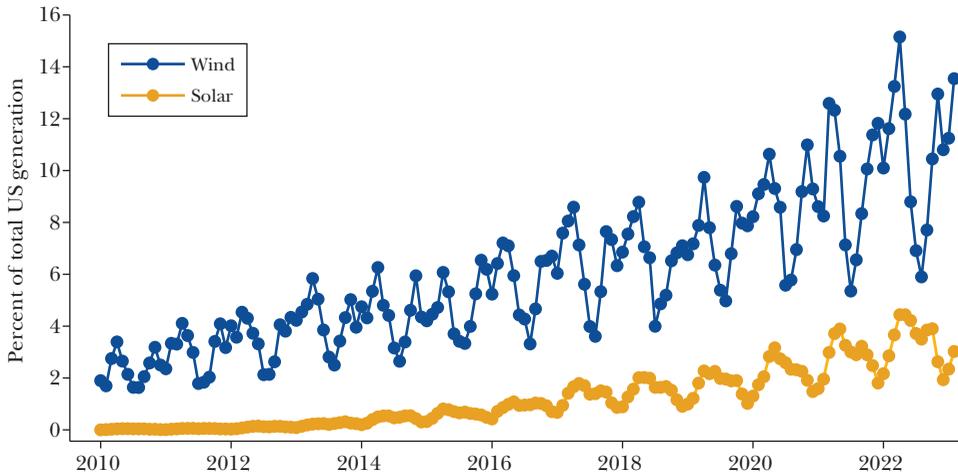
Source: This figure was created by the authors using monthly net generation by category from US Energy Information Adminstration (2023).
Note: Wind and solar are grid-scale generation as a percentage of total grid-scale generation from all sources. The seasonality reflects that during summer months (June–August), wind generation is 10 percent lower than other months whereas solar generation and total generation are 18 percent and 6 percent higher, respectively.

US Energy Information Administration (2023). In 2010, wind was less than 3 percent and solar was negligible. Over this period, total generation increased four-fold for wind and 120-fold for solar, such that by 2022, 10 percent of US electricity generation came from wind and 3 percent came from grid-scale solar. The other major categories of electricity generation in the United States are natural gas (40 percent), coal (20 percent), nuclear (18 percent), and hydroelectric (6 percent). Small-scale solar (for example, from rooftops) is not reflected in Figure 1 and is estimated to be smaller, only 1 percent.

This growth in renewable generation is impressive, but decarbonization scenarios require substantially more; in 2022, the two largest fuel sources for US electricity generation remained natural gas and coal. In recognition, 30 states have adopted renewable portfolio standards aimed at accelerating this transition toward renewables (Barbose 2021). California, for example, has a goal of 60 percent renewables by 2030 and 100 percent renewables by 2045. While the state-level renewable portfolio standards typically also include geothermal, hydroelectric, and in some cases, nuclear, the vast majority of new renewable capacity between 2020 and 2050 is expected to come from wind and solar (US Energy Information Administration 2022a).

There is also growing enthusiasm for taking advantage of lower-carbon electricity generation not just to replace existing electricity generation, but also to reduce the carbon intensity of other sectors. The most significant component of the "electrify

everything" movement would take the form of increased adoption of electric vehicles. In the United States alone, 350 million gallons of motor gasoline are used each day, according to the US Energy Information Administration (2022b), and electric vehicle proponents envision replacing much of this petroleum consumption with electricity, presumably generated from low- or zero carbon sources. Again, California has staked out a particularly aggressive goal, with the California Air Resources Board proposing a ban on new gasoline-fueled cars by 2035 (as reported by Friedman 2022).

Electrification of all energy uses in buildings is also receiving increased attention. Vast amounts of natural gas and other fossil fuels are consumed on-site in the United States for heating and other end uses, and a growing number of policies are aimed at transitioning much of this over to electricity (Davis 2023). New York City recently banned natural gas in new buildings, joining over 40 cities in California, Washington, Massachusetts, and Rhode Island that have either banned natural gas for new buildings or implemented "electric-preferred" building codes.

Both replacing existing fossil-fueled electricity generation and building new generation to meet the increased demands of widespread electrification of vehicles and buildings as part of a decarbonization transition will require extremely large increases in renewables. For example, in the baseline scenario in the Princeton (2021) study, US renewable capacity triples from existing levels by 2030, increases nine-fold by 2040, and rises by a factor of 16 by 2050. Similarly, US renewables generation increases by a factor of 22 by 2050 under the baseline decarbonization scenario in Williams et al. (2021).

The Mismatch between Population Centers and the Prime Locations for Renewable Power

Conventional sources for generating electricity can be sited on suitable land close to population and load centers with fuel transported to generating plants. In contrast, renewable generation capacity such as wind and solar must be sited where those natural resources are found. Indeed, the cost-competitiveness of wind and solar depends on them being located at favorable sites, which are not evenly distributed across the United States.

The best wind resources are located in the middle of the country. In states like Nebraska, Kansas, and Oklahoma, wind "capacity factors" are 40 percent or more, meaning that wind turbines produce 40 percent of what they would produce if they operated at maximum capacity 24 hours a day, 365 days a year.[3] In contrast, capacity factors in the rest of the United States tend to be less than 30 percent. To date, investments in wind generation have been heavily concentrated in states with the best wind resources. Texas, by itself, has 26 percent of US wind generation, more than the next three highest states combined: Iowa, Oklahoma, and Kansas. During

---

[3] See Lawrence Berkeley National Laboratory (2022) for wind capacity factors by state. For maps of US wind and solar resources, see NREL (2023a) and NREL (2023b), respectively.

2022, the top ten states accounted for 75 percent of US wind generation, while having only 32 percent of the US population.[4]

The best solar resources are heavily concentrated in the Southwest and Southeast regions. Solar capacity factors in these states are nearly twice as high as in northern states. Typical capacity factors for grid-scale solar are 29 percent in Arizona and 28 percent in California, compared to only 17 percent in New Jersey and Massachusetts (US Energy Information Administration 2019). Not surprisingly, investments in solar generation have been heavily focused on these states with the best solar resources. California is to solar what Texas is to wind, with 27 percent of US grid-scale solar, almost as much as the next three states combined (Texas, North Carolina, and Florida). During 2022, the top ten states accounted for 81 percent of US grid-scale solar generation, while having only 42 percent of the population.

This geographically uneven distribution of renewables points to the importance of electricity transmission. Not only is renewable generation potential distributed unevenly across states, but even within states with good renewable generation potential, the best locations often are far from major population centers and far from existing transmission infrastructure. It is not enough to generate renewable electricity at a competitive cost—getting this electricity to where it needs to go is increasingly just as important.

The Need for Transmission, Past and Present

The electricity supply chain relies, of course, not only on generation but also on delivery to consumers. As we described above, the costs of low-carbon generation have substantially fallen—but that has not solved the challenge of how to get that low-carbon electricity to potentially quite distant end users. Our focus in this paper is primarily on transmission; that is, high-voltage, large-volume transportation of electricity over medium to long distances.[5] As we describe in this section, there is abundant evidence that the United States already does not have enough transmission capacity to integrate the growing levels of renewable generation.

One Sign of Insufficient Transmission: Renewables Curtailment

In a functioning local electricity grid, quantity supplied needs to equal quantity demanded. If quantity supplied in a local market exceeds quantity demanded, then—at least in the absence of significant storage resources in most areas and

---

[4]This information and the information about solar in the following paragraph are authors' calculations based on net generation by state from US Energy Information Administration (2023) and state-level populations as of 2022 from US Census Bureau (2023).

[5]Whereas *transmission* refers to high-voltage large-volume transportation of electricity between the source of generation and high-voltage substations, *distribution* refers to low-voltage, lower-volume transportation between substations and the final customer. Although we do not focus on electricity distribution, there are important related challenges with regard to, for example, integrating residential electrification and electric vehicle adoption. For example, see Elmallah, Brockway, and Callaway (2022).

the presence of transmission constraints that limit the export of electricity to more distant demand centers—the quantity of electricity supplied to the grid needs to be reduced, or "curtailed."

One sign that many regions in the United States have insufficient transmission capacity is that renewables curtailment is becoming increasingly common, despite the zero or near-zero marginal cost of these resources. Renewables generation can be immediately and temporarily reduced—for example, the pitch controls on a wind turbine can be used to rotate the blades and generate less electricity, or solar arrays can be disconnected from the grid—and then quickly restored when this generation is needed again. For this reason, dumping power supplied by renewables can be easier than adjusting generation from fossil fuel and nuclear plants, which have operational constraints limiting the speed with which they can ramp generation up and down.

Peak loads happen at different times in different places, and at the same time renewables are being curtailed in some locations, there are often other locations nearby that would have benefited from access to this excess supply. Transmission constraints prevent these mutually beneficial trades from occurring, and create divergence in local prices for electricity.

Figure 3 plots monthly renewables curtailment in two major US electricity markets, selected because of their high saturation of solar and wind, respectively. The California Independent System Operator (CAISO) oversees electricity transmission in California, and the Southwest Power Pool (SPP) manages transmission for large parts of 15 states including Arkansas, Colorado, Iowa, Kansas, Louisiana, Minnesota, Missouri, Montana, Nebraska, New Mexico, North Dakota, Oklahoma, South Dakota, Texas, and Wyoming. Curtailment has increased dramatically in both markets.

At the beginning of the period, curtailment was negligible. Between 2015 and 2022, solar curtailment increased 18-fold in CAISO and wind curtailment increased 37-fold in SPP. Total solar curtailment in CAISO in 2022 was 1,734 gigawatt hours, equivalent to 4.4 percent of total solar generation. Total wind curtailment in SPP in 2022 was 11,124 gigawatt hours, equivalent to 10.3 percent of total wind generation.

These rising levels of curtailment reduce the incentive for additional renewables investments in these prime locations. After incurring large capital costs and finally getting a project online, the last thing a renewables developer wants to learn is that their generation is not needed during a large number of hours each year. Curtailment reduces the private and societal value of renewables investments and points to the broader challenge of integrating increasing levels of wind and solar.

Another Sign of Insufficient Transmission: Negative Wholesale Electricity Prices

Many US markets now routinely evidence an even more severe indicator of insufficient transmission capacity, namely, an increasing prevalence of negative wholesale electricity prices. In most markets, producers stop supplying a good when the price reaches zero or below. However, electricity generators often have technological and institutional reasons to continue supplying power even when the price

Figure 3
Increasing Curtailment of Renewables
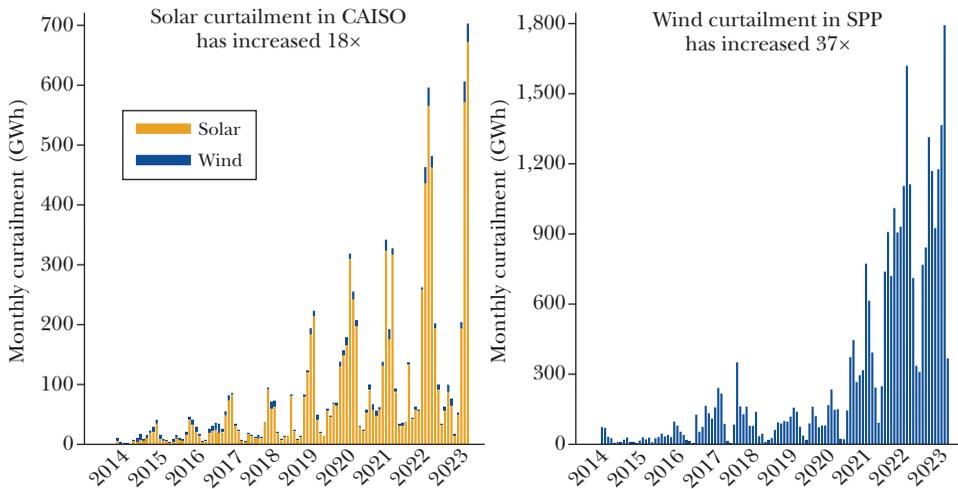
Source: This figure was created by the authors using data on renewables curtailment from the California Independent System Operator (CAISO 2023) and the Southwest Power Pool (SPP 2023).
*Note:* In CAISO, solar curtailment in 2022 was 1,734 gigawatt hours, which was 4.4 percent of total grid-scale solar generation. In SPP, wind curtailment in 2022 was 11,124 gigawatt hours, which was 10.3 percent of total wind generation. SPP has members in 15 states (Arkansas, Colorado, Iowa, Kansas, Louisiana, Minnesota, Missouri, Montana, Nebraska, New Mexico, North Dakota, Oklahoma, South Dakota, Texas, and Wyoming).

of power turns negative. For example, constraints on ramping power plants up or down (and other operational limitations), combined with the lack of cost-effective electricity storage, make electricity different from other goods. Quite simply, it is often not feasible to ramp down generation of a fossil-fuel or nuclear plant, even in response to several consecutive hours of negative prices. In these situations, generators may choose to pay to inject their power into the grid.

The design of policies meant to encourage renewable generation can further influence responses to negative prices. For example, wind generators tend to resist being curtailed because not producing means they do not receive the federal production tax credit (Aldy, Gerarden, and Sweeney 2023). Solar generators, in contrast, often receive an investment tax credit that does not depend on how much they produce and, in general, are thus more willing to curtail generation when prices are negative.

Figure 4 plots the frequency of negative electricity prices during all hours in 2022. This map shows prices from more than 50,000 individual locations across the seven major US electricity markets. Some parts of the United States, including most notably the Southeast, do not have electricity markets, so no information is reported

Figure 4
Frequency of Negative Electricity Prices in 2022
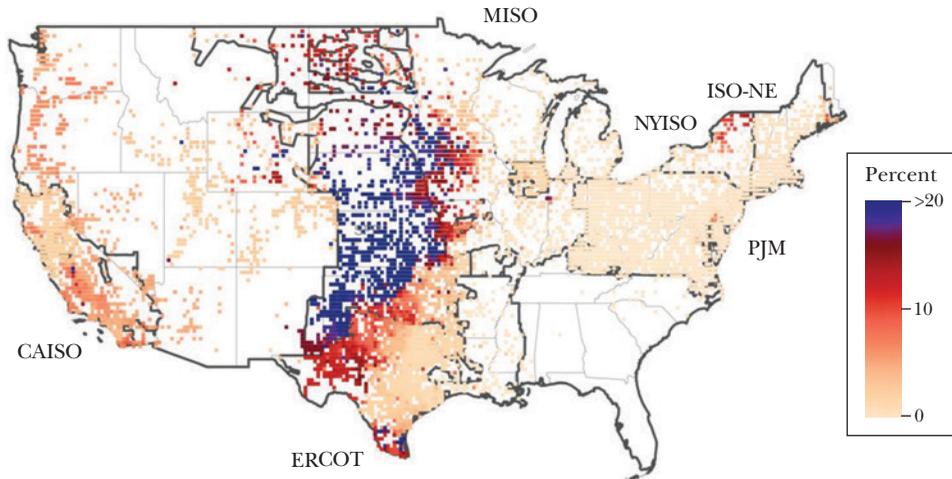
Source: This figure was reproduced with permission from Millstein, O'Shaughnessy, and Wiser (2023).
Note: The figure plots the frequency of negative local marginal electricity prices during all hours in 2022. The underlying price data in the ReWEP tool was compiled through the commercial product "Velocity Suite" based on prices from over 50,000 individual local nodes across the seven major US independent system operators. To verify the map, we spot-checked the negative price frequency at hundreds of locations in MISO and SPP (roughly, North Dakota to Michigan to Oklahoma) using hourly wholesale market price data.

for those places. Negative prices happen throughout the United States, but have become particularly common in the middle of the country where so much of the investment in wind generation has occurred.[6]

This phenomenon is relatively new. As recently as 2015, negative wholesale prices occurred in less than 2 percent of all hours and locations. Since 2015, the frequency of negative prices has increased steadily, reaching over 6 percent in 2022. Strikingly, there are hundreds of locations, mostly in the middle of the country, that now experience negative electricity prices during more than 20 percent of all hours. Seel et al. (2021) documents the increase in negative prices over this period and shows a strong correlation between wind generation and negative prices.

Negative prices indicate the need for increased transmission investments because at the exact same time these negative prices are occurring in some locations, there often are other locations not far away with customers willing to pay for

---

[6]In related work, Bushnell and Novan (2021) document that renewables have decreased electricity prices in California, with a distinct hourly pattern with large decreases in midday prices combined with modest increases during "shoulder hours" like in the evenings when the sun is setting. They also show that while wind and solar tend to generate at different times, their combined profile does not exactly match the timing of demand.

additional electricity supply. These price impacts are related to the broader operational and market design challenges associated with integrating renewables into electricity markets. For example, Joskow (2019) argues that serious market reforms are necessary to provide incentives for investment in "dispatchable" generation, storage, or other resources necessary to manage the intermittency of renewables. For additional discussion, see Gowrisankaran, Reynolds, and Samano (2016), Joskow (2019), and Mallapragada et al. (2022). While market design improvements could reduce these operational challenges, they do not eliminate the value of more transmission.

A related infrastructure problem is the difficulty of building new interconnection lines. In all US markets, new generation projects need to request and receive authorization before being connected to the electric grid. Typically, this process includes a series of studies aimed at evaluating how the new generation would impact grid operations and stability, with particular attention to any necessary transmission system upgrades or other additional physical infrastructure that would be required. As of 2021, there was almost one terawatt of solar and wind projects in US interconnection queues, and the amount of time that projects remain in these queues has been increasing steadily (Rand et al. 2022). In addition, there are widespread reports of renewables projects being withdrawn from interconnection queues because of concerns about transmission congestion, particularly in the Midwest (for example, as reported in Tomich 2019 and Penrod 2022).

Looking Ahead: Can the United States Build the Transmission Network It Needs?

Given the evidence that the United States does not have enough transmission capacity to efficiently utilize even the current level of renewables, one might reasonably ask what happens as we move toward a net zero carbon future. Figure 5 helps to put the necessary expansion of transmission in context, juxtaposing the assumed levels of total US transmission capacity under three prominent decarbonization scenarios with the historical record since 2005. These studies outline a range of different scenarios with varying levels of increased renewables, electrification of vehicles and buildings, carbon capture, and other features (Princeton 2021; Williams et al. 2021; National Academies 2021). The exact combination of strategies varies across scenarios, but a key feature of all three studies is that they are renewables-heavy and therefore assume an unprecedented level of investments in electricity transmission over the next several decades.

From 2005 to 2020, transmission capacity shown in Figure 5 grew by 27 percent. Although this increase may seem modest, it was actually a break in historic trends. The annual average increase in transmission capacity over 2015–2020 was greater than the annual average over the previous 30 years (US Department of Energy 2015), which itself had followed decades of falling investment in the transmission system (Hirst and Kirby 2001).

Figure 5

Decarbonization Likely Requires Vastly More Transmission
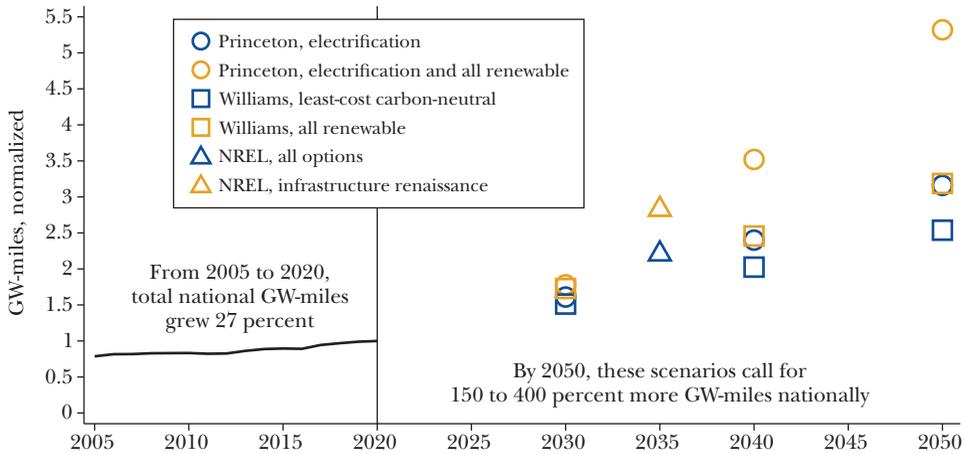
Source: The left-hand side of the figure was created by the authors using data on transmission miles from FERC Form 1, 2005–2020 (Catalyst Cooperative 2022). Details on the conversion from FERC's miles data to our reported gigawatt-miles are provided in this paper's data archive.
Note: This figure juxtaposes historical US electricity transmission capacity (normalized to one in 2020) with the future capacity called for in three prominent decarbonization studies (Princeton 2021; Williams et al. 2021; Denholm et al. 2022).

By the late 2010s, major US utilities were spending more than $20 billion annually on new transmission investments (US Energy Information Administration 2021d). Some of the drivers of this growth were to integrate renewables, replace aging infrastructure, improve storm hardening, and improve reliability (US Energy Information Administration 2018, 2021e). However, the majority of these projects were fairly small-scale—for instance, within the service territory of one utility and less than 100 miles (Catalyst Cooperative 2022)—and did not involve the kind of cross-state or cross-region coordination that is particularly challenging.

The six scenarios plotted in Figure 5 all assume dramatic increases in US transmission capacity between 2020 and 2050. Even the least aggressive scenario entails more than a doubling of transmission capacity by 2050. The other scenarios involve three-, four-, or five-fold increases in transmission, depending on the extent to which decarbonization relies exclusively on renewables, as opposed to, say, nuclear power or carbon capture.

Under all of the assumed scenarios, the annual growth in transmission greatly exceeds the annual growth 2005–2020. It is difficult to overstate the scope of such an increase. In total inflation-adjusted dollars, a transmission expansion of this magnitude could cost more than historically massive investments like the national highway system—in which almost $600 billion (in 2022 dollars) was spent on 43,000

lane-miles of highway over a 35-year period (Erickson 2012). Moreover, as we discuss in the following sections, there are additional factors, above and beyond financial cost, that make large-scale transmission projects particularly challenging.

The Overhang of Fragmented Regulated Utilities

While it makes sense today to think about generating wind power in Oklahoma to deliver to consumers in Tennessee, this type of long-distance cross-state transmission was not how the US electric grid was designed. The main exceptions, when portions of the grid were designed to move power over long distances, typically were associated with efforts to bring remote hydroelectric power to high-population load centers, like the electricity lines connecting power generated in the Pacific Northwest to California, or to share the power from a large scale generation plant (particularly nuclear plants) across customers of more than one utility (Joskow 2021).

A bit of history is helpful for understanding why it is so difficult to build large-scale transmission projects in the United States. The US electricity grid is not a single, centrally-designed entity; instead, it is a disorganized patchwork resulting from more than a century of mostly disconnected individual electric utilities making decisions for their own monopoly franchise territories. Fossil and nuclear fuel can be transported to generating plants, enabling utilities to choose generation sites based on proximity to load, subject to available land and environmental restrictions. This enabled transmission networks built to connect generating plants to load centers largely within a utility's service area; interconnections across networks were relatively limited and typically motivated more by resiliency and network reliability concerns. This system had little need to transport significant volumes of power across service territories or to connect utilities across different regions of the country (Joskow 2021).

Moreover, the cost of service regulation traditionally applied in the United States incentivized utilities to go it alone rather than cooperate with others. If a utility builds a power plant, the cost is a capital expense for which the utility earns a profitable rate-of-return, as the allowed rate of return typically exceeds the utility's cost of capital. Buy electricity from someone else, however, and the cost is operating expense for which the utility earns no profit.[7] Gold (2019) tells the story of a utility executive speaking to a developer for additional transition lines: "Why would we buy from you and make no money? We'd rather run our own plants and make money that way." Economists have long argued that rate-of-return regulation can create a bias toward capital-intensive investments (Averch and Johnson 1962), and

---

[7] Traditional vertically integrated electric utilities receive a rate of return for both generation and transmission investments (for example, see Joskow 2005). However, generation has historically been a much larger component than transmission (Joskow and Schmalensee 1988). Moreover, buying electricity from someone else typically involves using someone else's transmission line, in which case any transmission charges are an operating expense rather than a capital investment.

this preference for building rather than buying is part of the reason we have such a fragmented grid to begin with (Cicala 2021).

Utilities have traded electricity bilaterally for a long time and through organized wholesale markets in many parts of the United States since about 2000. These wholesale markets have tended to improve market efficiency (Cicala 2022), but such trades have occurred subject to the constraints of a transmission system historically built almost entirely by individual utilities operating by themselves, not designed for long-distance market integration, and certainly not optimized from a centralized perspective. The United States has over 50 "Balancing Authorities," which are the system operators responsible for managing the power flows within their network and ensuring compliance with the operating criteria for the synchronized networks to which they are connected (Joskow 2021). These range from relatively small single-utility systems to large Regional Transmission Operators (RTOs) and Independent System Operators (ISOs). In regions with deregulated wholesale generation markets, generation dispatch and the operation of the grid has been turned over to ISOs or RTOs, even as individual utilities continue to own most transmission lines. While RTOs/ISOs oversee transmission planning and investment decisions, most of their focus until very recently was on network reliability. Despite efforts by the Federal Energy Regulatory Commission (FERC) to extend planning criteria to include economic efficiency and policy objectives through its FERC Order 1000, transmission investment continued to be driven primarily by compliance with reliability standards (Pfeifenberger et al. 2021; Joskow 2021).

The most obvious and immediate consequence of this history is that the US electric grid is inherently limited in terms of overall capacity for long-distance movement of power. There is virtually no capacity to move electricity between the three so-called "interconnections" in the continental United States—Eastern, Western, and Texas—despite evidence that the benefits of such connections would greatly exceed the costs (McCalley et al. 2012; Bloom et al. 2022). Apart from those relatively few high-capacity lines built long ago to transmit hydro power to distant customers within interconnections, there is a limited ability to move electricity, and this particularly binds during peak periods when those connections would be most valuable.

The lack of centralized decision-making means that proposals for large transmission projects typically require high degrees of consensus among affected parties. A new transmission line connecting several states, for example, usually requires approval by the utility commissions in each state as well as, in many cases, the major affected utilities as well. Depending on the state, siting authority usually rests with the public utility commission, another agency, or, in some cases, multiple agencies. The review typically includes economic analyses, environmental reviews, and public hearings, after which the state must decide if the project is in the public interest (for additional details, see FERC 2020). But near-universal consensus can be difficult to achieve, because most projects create both winners and losers.

How Additional Transmission Creates Winners and Losers

When you take energy from a location with a low price and you move it to another location with a high price, you create economic value.[8] However, there are inevitably winners and losers from market integration. Commercial and industrial customers, ratepayer advocates, and other electricity buyers in renewables-rich areas tend to oppose new transmission because, for them, electricity flowing to other areas pushes prices upward. Similarly, owners of existing fossil-fuel power plants and other electricity sellers in receiving areas tend to oppose new transmission because receiving energy from other areas pushes prices downward. Even if a project creates large aggregate net benefits, those made worse off are likely to protest.

Utility regulation does not solve these challenges. Traditional vertically-integrated utilities in receiving areas tend to oppose new transmission because it makes their power plants less valuable, and because cheap electricity from other places reduces the need for new local investments in generation. Moreover, regulated utilities are under only weak incentives to deliver lower-priced electricity to their customers, so they tend not to be particularly motivated by potential cost savings.

In addition, some of the economic value from increased transmission of electricity generated with carbon-free methods comes in the form of reduced environmental externalities, but there is typically no stakeholder at the table in these negotiations to represent the interests of reducing global greenhouse gas emissions.[9] For the same reason, there typically is no direct financial incentive for projects with particularly beneficial environmental impacts. At least in theory, pricing carbon (and other environmental externalities) would help align incentives. But, of course, there is no price on carbon in most US states, and even if there were, it would not eliminate the kind of coordination issues described here where it can be difficult to build near-perfect consensus for any project that creates winners and losers.

Free Riders and Cost Allocation

Even when stakeholders agree on the overall need for particular transmission investment (or at least, do not oppose such an investment), disagreements often emerge over how project costs should be divided. To understand why cost allocation is such a challenge, it is helpful to compare electricity transmission to natural gas pipelines, and to think about how the network externalities with electricity transmission make it harder to finance. US interstate electricity transmission and natural gas

---

[8]Gonzales, Ito, and Reguant (2022) document a vivid example of price convergence following recent electricity transmission expansions in Chile. Prior to the expansions, the two largest electricity markets in Chile were separate and disconnected. Market integration led to price convergence and enabled additional renewables investments which they find would not have been profitable without the expansions.

[9]In related research, Fell, Kaffine, and Novan (2021) use data from two major US electricity markets to show that the environmental benefits from wind generation are 30 percent larger when transmission is uncongested. This result happens in the markets they study primarily by offsetting more fossil generation in population-dense locations. Their results imply that a major recent transmission project in Texas (CREZ) increased the environmental benefits from Texas wind generation by $366 million annually.

pipelines are both regulated by Federal Energy Regulatory Commission, but the two markets are very different. With natural gas, a pipeline owner has complete control over who injects natural gas at one end and who extracts it at the other. US interstate natural gas pipelines have been built mostly on a decentralized, contract-based framework, with private pipeline developers building projects paid for by pipeline customers. The simplicity, speed, and flexibility of this approach enabled the growth of a vast US network of interstate pipelines, and it has been nimble enough to respond to changes in market conditions (Adamson 2018).

This decentralized, quasi-competitive approach to contracting has not worked well with electricity transmission. Part of the challenge has to do with the physical laws of power flow. Electricity is injected and withdrawn at many locations in the grid, and there are network externalities such that changes at any location affect the entire grid, with consequences for transmission constraints, market power, and other issues (Borenstein, Bushnell, and Stoft 2000; Griffin and Puller 2009; Joskow 2012; Davis and Hausman 2016; Ryan 2021). These physical features of electricity transmission make it more like a public good, with the benefits from increased transmission experienced among a more diffuse set of beneficiaries. As such, electricity transmission is susceptible to free riding. A new transmission line can help relieve constraints and increase reliability throughout the broader region, including relatively distant parts of the grid, but quantifying those benefits and getting beneficiaries to recognize and pay for them can be challenging.

Not coincidentally, there are continued calls to reform the system of cost allocation used to finance electricity transmission (Hogan 2018; Olmos, Rivier, and Pérez-Arriaga 2018). A series of orders from Federal Energy Regulatory Commission have attempted to increase incentives for new investments in electricity transmission, including guidance that costs be allocated "roughly commensurate" with estimated benefits (Adamson 2018; Joskow 2020a; Joskow 2020b). However, it can be difficult to precisely quantify the benefits of transmission and, even while following FERC cost allocation principles, two parties can reach very different conclusions about the magnitude of benefits. As FERC (2020) puts it, "Given this complexity and the general contentious nature of cost allocation issues, cost allocation determinations may continue to be prone to disagreement and litigation that present a challenge to development of transmission facilities, including high-voltage transmission." In practice, these and other challenges have frustrated FERC's efforts to encourage new transmission investment, in sharp contrast to the more dynamic natural gas pipeline sector.

Local Siting and Permitting Challenges

Siting and permitting issues—determining the line's route and securing the necessary land use authorizations—can stymy construction. Long-distance transmission projects typically require permission from hundreds of different landowners. Transmission project developers must convince landowners to allow construction on their properties, a process that is uncertain, expensive, and time-consuming. Landowners often oppose high-voltage transmission lines due

to concerns about visual impacts, perceived health effects, site preservation, and other issues (Vajjhala and Fischbeck 2007; Cohen et al. 2016; Mueller, Keil, and Bauer 2017). Saul, Malik, and Merrill (2022) report the saga of the TransWest Express line, which sought to send power from Wyoming wind farms to California but was held up by land-use concerns.

An unusually large number of agencies are involved in this process. It is common for the siting application of a new line to reference coordination with multiple federal agencies (for example, the Army Corps of Engineers, Fish and Wildlife Service, Forest Service, Federal Aviation Administration, and Department of Agriculture) and state-level authorities (for example, agencies responsible for environmental protection, transportation, agriculture, and historic preservation), along with county, municipal, and tribal authorities. Examples of laws that may apply to land use are the Clean Water Act, the Rivers and Harbors Act, the Bald and Golden Eagle Protection Act, the Endangered Species Act, the National Historic Preservation Act, and the Federal Aviation Administration Act (FERC 2020). And if a project crosses federal land, then the project must also satisfy National Environmental Policy Act (NEPA) requirements, including detailed environmental impact assessments (FERC 2020). Permitting Institute (2023) provides a remarkable flowchart that references these and other agencies in illustrating the complicated permitting process that may be required for new US transmission projects. Moreover, the challenge is not just the numerous agencies—but the potential throughout for these processes to be co-opted by private economic interests.

Related research on local opposition to energy projects shows how it can significantly increase costs. For example, Jarvis (2022) finds that local opposition to wind projects in the United Kingdom pushes these projects toward less-desirable locations, increasing total project costs by between 10 percent and 29 percent. Indeed, the United Kingdom's investments in off-shore wind are viewed, in part, as a response to how difficult it is to overcome local opposition to on-shore wind projects (Jones and Eiser 2010). Of course, local opposition to large infrastructure projects is not unique to energy infrastructure. Brooks and Liscow (2023) find that spending per mile on interstate highways increased three-fold between the 1960s and 1980s. Increasing incomes and housing prices explain just over half the increase, which they interpret as the rising cost of "citizen voice."

Electricity transmission sometimes ends up on the ballot. Maine residents voted in 2021 to reject a 145-mile transmission line that would have connected Canadian hydropower with electricity consumers in New England (as reported by Kamp 2021). A similar project was rejected in New Hampshire by a government committee in 2018 (as reported by Ailworth and Kamp 2018). Opponents in both cases compared the project to an extension cord stretched through their forested landscape, highlighting visual impacts, local site preservation, and other issues, while also emphasizing that much of the benefit would go to other states. There is often interaction between local siting issues and political support for green policies. For example, Stokes (2016) finds a backlash by voters who live within three kilometers of newly-sited wind turbines in Ontario, Canada.

Resolving these tensions between the broader public good and local land use concerns is one of the key barriers preventing faster growth in energy infrastructure. Whether it is an electricity transmission line to connect hydro power from Quebec to consumers in New England, new connections to renewables in the Midwest, or a new wind farm in the United Kingdom, the benefits of energy infrastructure are usually widely dispersed, while the land use concerns are highly localized.

Paying for New Transmission Lines

Even after all the relevant stakeholders have approved a project, the transmission lines still need to be built. Transmission lines are large capital investments: building a high-capacity long-distance transmission line can cost $3 million or more per mile (WECC 2019). Unlike the costs for grid-scale renewables, there is little evidence that the costs of transmission lines have declined over time.[10] In fact, the most expensive transmission projects in recent years have had much higher costs per mile than the most expensive projects in the early 2000s (authors' calculations, based on Catalyst Cooperative 2022). Whereas renewables have benefited from economies of scale and learning-by-doing, the technology for electricity transmission is largely unchanged from what it has been for decades.

There is also the related question about whether transmission lines will be used enough to justify the required capital costs. With a typical fossil fuel or nuclear power plant, it is possible to size transmission to guarantee that the lines are used at close to full capacity, 24 hours a day, 7 days a week. This is not the case for renewables: a transmission line connecting a solar-rich area to the rest of the grid may be operated only 30 percent of the time, for example. This capacity factor problem tends to mean that electricity systems based on renewables need more total transmission capacity than systems based on fossil fuels (Sioshansi and Denholm 2012).

What Can Be Done?

In this section we discuss several approaches for accelerating US electricity transmission projects. We also briefly discuss two potential substitutes for additional transmission: storage and dynamic pricing. While neither substitute would eliminate the need for additional transmission, both could play an important role

---

[10] Evidence on costs comes from industry sources like the Western Electricity Coordinating Council, which coordinates the entire Western interconnection system, the Southwest Power Pool, which manages the grid for the central United States, and the Midcontinent Independent System Operator, which manages the high-voltage grid across 15 US states and Manitoba. For examples, see WECC's 2012 "Capital Costs for Transmission and Substations," WECC's 2014 "Capital Costs for Transmission and Substations," WECC's 2019 "Transmission Capital Cost Tool—with E3 Updates, SPP's 2016 "The Value of Transmission," SPP's 2021 "The Value of Transmission," and MISO's 2020 "Transmission Cost Estimation Guide."

accommodating renewables growth, even if the United States proves unable to build new transmission at the pace suggested by recent decarbonization studies.

Approaches for Increasing Transmission

Probably the most discussed potential policy reform is enhanced federal authority for siting and permitting new transmission lines. Whereas the current approval process is split between federal, state, and local agencies, the public good characteristics of electricity transmission provide a clear economic argument for greater centralization of these decisions (Brown and Botterud 2021). Increased transmission capacity creates economic benefits that are widely diffused across states, and a federal agency can take these spillovers into account when evaluating projects in a way that individual states or utilities are not incentivized to do. The Federal Energy Regulatory Commission has long been the central authority for siting natural gas pipelines, and a similar approach could be adopted with electricity transmission (Borenstein and Kellogg 2021; Cicala 2021). MIT (2011) argues that enhanced authority should be reserved for interstate transmission lines, given the especially large coordination challenges for those projects. Reforms in Europe have to some degree followed this prescription (Joskow 2021).

Enhancing federal authority for siting will not be easy. The Bipartisan Infrastructure Law of 2021 and the Inflation Reduction Act of 2022 take steps in this direction, for example, by giving the Federal Energy Regulatory Commission the authority to overturn state objections to transmission projects (Sud, Patnaik, and Glicksman 2023). But it is too soon to say to what extent FERC will be willing or able to exert this new authority. Previous attempts to enhance federal authority—for example, with FERC Order 1000—have run into considerable challenges with state and local regulators unwilling to cede their authority (Joskow 2020b). Additional legislation proposed in 2022 by Senators Chuck Schumer and Joe Manchin would have gone further, taking steps to address many of the barriers we discussed, including improved cost allocation rules, enhanced federal authority for permitting, simplified NEPA procedures, and simplified multi-agency coordination (Goggin and Gramlich 2022), but this legislation faced significant political hurdles and did not pass.

There also are opportunities for shifting where new transmission investments are happening. Rather than focus on new transmission projects in greenfield locations, more of the emphasis could be shifted toward upgrading lines in existing transmission corridors. High-voltage lines can be upgraded to expand capacity, for example, from 230kV to 345kV. There is also the potential to increase the capacity of existing high voltage alternating current (HVAC) lines by converting them to high-voltage direct current (HVDC) lines or hybrid AC/DC lines. Reed et al. (2019) explain that such conversions can increase total transmission capacity by up to four times. Capacity expansion projects still require large capital investments, but can be easier from the perspective of local siting concerns. New transmission projects could also be placed along waterways, railroads, highways, and other corridors that have already been designated for public infrastructure use (Cicala 2021). This approach of using public infrastructure corridors can be

easier than negotiating right-of-way permissions with a large number of individual landowners (FERC 2020) and has the potential to significantly reduce local siting concerns relative to projects that break new ground.[11]

Increasing Storage

Perhaps the closest substitute for additional transmission is more capacity for electricity storage. Whereas transmission allows for arbitrage across locations, storage allows for arbitrage across time. In an electricity market with ample storage, you would not expect to see renewables curtailment or frequent negative prices. Moreover, when storage is co-located with renewables, transmission can be used for a greater fraction of all hours throughout the year, improving the financial viability of transmission investments.

Electricity storage is expensive, so historically there has been very little of it. By far the largest form of electricity storage worldwide is pumped hydro, which refers to facilities with two water reservoirs at different elevations as well as pumps for moving water uphill. When electricity is scarce, water is released down through turbines to generate electricity. Then, when electricity is plentiful, water is pumped back up. Until recently, pumped hydro represented more than 90 percent of US grid-scale storage, but capacity is growing relatively slowly, mostly from upgrades to existing facilities (Uría-Martínez, Johnson, and Shan 2021).

The fastest-growing form of electricity storage relies on lithium ion batteries. For example, the 400-megawatt Moss Landing project in Monterey, California has been called the largest battery storage facility in the world (Gearino 2021). US electric utilities have spent billions over the last few years building battery storage projects like Moss Landing. Total US battery capacity on the grid reached 1,650 megawatts at the end of 2020, and is forecasted to reach 12,000 megawatts by the end of 2023 (US Energy Information Administration 2021c). But even with this recent growth, total battery storage is still small compared to the size of the market. Peak electricity demand in the United States is about 700,000 megawatts (US Energy Information Administration 2021b). Thus 12,000 megawatts by the end of 2023 is not negligible, but still represents less than 2 percent of peak US demand. Of course, during high demand periods, it can be very valuable to have even a relatively small amount of stored electricity available (for an example, see Blaustein 2022). But current battery storage investments can play only a modest role in addressing hour-to-hour imbalances between supply and demand.

Could grid-scale storage scale up dramatically? It is not clear. Grid-scale battery storage costs did fall more than 70 percent from 2015 to 2019 (US Energy Information Administration 2021a), but battery storage is still not cost-effective (Karaduman 2021; Cole, Frazier, and Augustine 2021). There is scope for some

[11]Relatedly, there are also potential opportunities for siting renewable generation facilities on the grounds or near retired power plants, as reported in Shao (2022). Hundreds of US coal plants have closed or are planning to close, and these locations are already connected to the grid with existing transmission infrastructure.

optimism about battery storage costs declining further, given that these technologies potentially benefit from economies of scale and learning-by-doing. Cole, Frazier, and Augustine (2021) predict steep further declines for lithium ion batteries over the next several years. Moreover, engineers are working on a range of differential alternative battery technologies that show promise (MIT 2022). As one example, an MIT-based startup called Form Energy is trying to develop a cost-effective battery using iron, air, and water (as reported in McCarthy 2022; Ramkumar 2022).

The challenges associated with building electricity transmission imply that the private and societal benefits of new storage technologies are larger than they would be otherwise. A breakthrough in battery technology would significantly offset the need for additional transmission. Thus, an indirect but potentially important policy response to the challenges of building new electricity transmission would be to increase US government support for research and development in storage and other substitute technologies. Innovation generates knowledge spillovers, a positive externality. It appears that knowledge spillovers have contributed to the cost declines in wind turbines (Covert and Sweeney 2022), and more broadly that solar, wind, and energy storage create knowledge spillovers (Noailly and Shestalova 2017). Other researchers have argued that path dependency in technological development justifies government investment in research and development on clean technologies (Acemoglu et al. 2016; Aghion et al. 2016).

Dynamic Pricing

Another potential substitute for electricity transmission is dynamic pricing. The vast majority of electricity customers face time-invariant prices, which fail to efficiently communicate real-time changes in market conditions (Borenstein 2005; Borenstein and Holland 2005). If instead customers faced higher prices during peak periods, they would demand less electricity during those periods. In turn, this would allow for smaller investments in generation and transmission capacity, which are driven by peak demand, not just average demand.

Dynamic pricing is relatively rare, particularly in the residential sector. Consumer groups often object to dynamic pricing, raising concerns that customers do not understand complex prices or cannot rapidly respond to price changes, and that some customers would pay more (Joskow and Wolfram 2012). Dynamic pricing can also increase the overall volatility of electricity bills, which is particularly challenging for lower-income households who have less resilience to economic shocks (Borenstein 2013).

Nonetheless, economists have argued for decades that electricity markets would be more efficient with dynamic pricing (Boiteux 1960). In addition, dozens of empirical studies document reductions in electricity demand in response to dynamic pricing including, both the residential and nonresidential sectors (Ito, Ida, and Tanaka 2018; Blonz 2022). Consumers are also becoming more able to respond to real-time price changes as communications technology, and in particular, automation, continues to improve (Jessoe and Rapson 2014; Bollinger and Hartmann 2020). Smart thermostats, smart electric vehicle chargers, and other

automated technologies mean that consumers can preset which adjustments will (or will not) occur when prices rise, and so they do not need to be aware of price changes in order to be price responsive.

The benefits from dynamic pricing are thought to be particularly large in systems with high renewables penetration. Imelda, Fripp, and Roberts (forthcoming) apply a model of investment, supply, storage, and demand to evaluate the economic benefits from dynamic pricing on the island of Oahu, Hawaii. They find that the gains from dynamic pricing are six to twelve times higher for a high-renewables system relative to a system dominated by fossil fuels. Dynamic pricing plays several roles in the optimized system, including not only moving demand from peak to off-peak periods, but also addressing challenges associated with cloudy days and other forms of renewable intermittency. Oahu is a compelling setting, in part, because of its lack of connectivity with other electricity markets. In some sense, examining Oahu is to ask what an electricity system would need to look like if building electricity transmission were completely impossible. This is, of course, more extreme than the situation faced on the continental United States, but it nonetheless provides a valuable setting for demonstrating the potential for dynamic pricing in transmission-constrained scenarios.

Conclusion

Many of the technologies that could decarbonize the US economy are getting cheaper, most notably wind and solar generation, but getting to full decarbonization is nonetheless a challenge. The most promising scenarios presented to date rely on massive investments in the electricity sector. And a key sticking point with current technologies is the need to dramatically grow the transmission capacity that would transport all the new low-cost wind and solar generation to homes and businesses. This problem is particularly true in the United States, but most of these constraints exist in other countries as well. As Joskow and Schmalensee (1988) wrote: "The role of the transmission network in transporting power and in coordinating the efficient supply of electricity in both the short run and the long run is the heart of a modern electric power system."

We have highlighted the many signs pointing towards barriers to this expansion and the myriad reasons that it is hard to build new transmission. Some of these are bureaucratic, and others are about economic fundamentals, like the public good nature of transmission and the winners versus losers problem inherent with any market integration. We have described potential policy remedies as well as the potential role of storage and dynamic pricing, probably the two most important substitutes to transmission.

The topic of electricity transmission is ripe for research. While numerous white papers and government reports explore the problem of transmission expansion in the United States, recent academic research in this area is strikingly sparse. Empirical work on the impacts of past transmission policy changes, on transmission

market design, or on the spillovers between transmission and generation markets could contribute to a better understanding of and resolution of the current policy challenge.

References

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2016. "Transition to Clean Technology." *Journal of Political Economy* 124 (1): 52–104.

Adamson, Seabron. 2018. "Comparing Interstate Regulation and Investment in U.S. Natural Gas and Electric Transmission." *Economics of Energy and Environmental Policy* 7 (1): 7–24.

Aghion, Philippe, Antoine Dechezleprêtre, David Hémous, Ralf Martin, and John Van Reenen. 2016. "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry." *Journal of Political Economy* 124 (1): 1–51.

Ailworth, Erin, and Jon Kamp. 2018. "New England Has a Power Problem." *Wall Street Journal*, February 23. https://www.wsj.com/articles/new-england-has-a-power-problem-1519390800.

Aldy, Joseph E., Todd D. Gerarden, and Richard L. Sweeney. 2023. "Investment versus Output Subsidies: Implications of Alternative Incentives for Wind Energy." *Journal of the Association of Environmental and Resource Economists* 10 (4): 981–1018.

Averch, Harvey, and Leland T. Johnson. 1962. "Behavior of the Firm under Regulatory Constraint." *American Economic Review* 52 (5): 1052–69.

Barbose, Galen L. 2021. *U.S. Renewables Portfolio Standards: 2021 Status Update: Early Release.* Berkeley, CA: Lawrence Berkeley National Laboratory.

Blaustein, Anna. 2022. "How California Kept the Lights On during Monster Heat Wave." *Scientific American*, September 16. https://www.scientificamerican.com/article/how-california-kept-the-lights-on-during-monster-heat-wave/.

Blonz, Joshua A. 2022. "Making the Best of the Second-Best: Welfare Consequences of Time-Varying Electricity Prices." *Journal of the Association of Environmental and Resource Economists* 9 (6): 1087–1126.

Bloom, Aaron, Josh Novacheck, Greg Brinkman, James McCalley, Armando Figueroa-Acevedo, Ali Jahanbani-Ardakani, Hussam Nosair et al. 2022. "The Value of Increased HVDC Capacity between Eastern and Western U.S. Grids: The Interconnections Seam Study." *IEEE Transactions on Power Systems* 37 (3): 1760–69.

Boiteux, Marcel. 1960. "Peak-Load Pricing." *Journal of Business* 33 (2): 157–79.

Bollinger, Bryan K., and Wesley R. Hartmann. 2020. "Information vs. Automation and Implications for Dynamic Pricing." *Management Science* 66 (1): 290–314.

Borenstein, Severin. 2005. "The Long-Run Efficiency of Real-Time Electricity Pricing." *Energy Journal* 26 (3): 93–116.

Borenstein, Severin. 2012. "The Private and Public Economics of Renewable Electricity Generation." *Journal of Economic Perspectives* 26 (1): 67–92.

Borenstein, Severin. 2013. "Effective and Equitable Adoption of Opt-In Residential Dynamic Electricity Pricing." *Review of Industrial Organization* 42 (2): 127–60.

Borenstein, Severin, and Stephen Holland. 2005. "On the Efficiency of Competitive Electricity Markets with Time-Invariant Retail Prices." *RAND Journal of Economics* 36 (3): 469–93.

Borenstein, Severin, and Ryan Kellogg. 2021. "Challenges of a Clean Energy Transition and Implications for Energy Infrastructure Policy." In *Rebuilding the Post-Pandemic Economy*, edited by Melissa S. Kearney and Amy Ganz, 234–71. Washington, DC: Aspen Institute Press.

Borenstein, Severin, James Bushnell, and Steven Stoft. 2000. "The Competitive Effects of Transmission Capacity in a Deregulated Electricity Industry." *RAND Journal of Economics* 31 (2): 294–325.

Brooks, Leah, and Zachary D. Liscow. 2023. "Infrastructure Costs." *American Economic Journal: Applied Economics* 15 (2): 1–30.

Brown, Patrick R., and Audun Botterud. 2021. "The Value of Inter-regional Coordination and Transmission in Decarbonizing the US Electricity System." *Joule* 5 (1): 115–34.

Bushnell, James, and Kevin Novan. 2021. "Setting with the Sun: The Impacts of Renewable Energy on Conventional Generation." *Journal of the Association of Environmental and Resource Economists* 8 (4): 759–96.

California Independent System Operator (**CAISO**). 2023. "Managing Oversupply, Wind and Solar Production and Curtailments." CAISO. http://www.caiso.com/informed/Pages/ManagingOversupply.aspx (accessed May 19, 2023).

Catalyst Cooperative. 2022. "PUDL Project, FERC Form 1 Data." https://data.catalyst.coop/ferc1 (accessed April 8, 2022).

Cicala, Steve. 2021. "Decarbonizing the U.S. Economy with a National Grid." Unpublished.

Cicala, Steve. 2022. "Imperfect Markets versus Imperfect Regulation in US Electricity Generation." *American Economic Review* 112 (2): 409–41.

Cohen, Jed, Klaus Moeltner, Johannes Reichl, and Michael Schmidthaler. 2016. "An Empirical Analysis of Local Opposition to New Transmission Lines across the EU-27." *Energy Journal* 37 (3): 59–82.

Cole, Wesley, A. Will Frazier, and Chad Augustine. 2021. "Cost Projections for Utility-Scale Battery Storage: 2021 Update." National Renewable Energy Laboratory, NREL/TP-6A20-79236.

Covert, Thomas R., and Richard L. Sweeney. 2022. "Winds of Change: Estimating Learning by Doing without Cost or Input Data." Unpublished.

Davis, Lucas W. 2023. "What Matters for Electrification? Evidence from 70 Years of U.S. Home Heating Choices." *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01292.

Davis, Lucas, and Catherine Hausman. 2016. "Market Impacts of a Nuclear Power Plant Closure." *American Economic Journal: Applied Economics* 8 (2): 92–122.

Denholm, Paul, Patrick Brown, Wesley Cole, Trieu Mai, Brian Sergi, Maxwell Brown, Paige Jadun et al. 2022. *Examining Supply-Side Options to Achieve 100% Clean Electricity by 2035*. Golden, CO: National Renewable Energy Laboratory.

Elmallah, Salma, Anna M. Brockway, and Duncan Callaway. 2022. "Can Distribution Grid Infrastructure Accommodate Residential Electrification and Electric Vehicle Adoption in Northern California?" *Environmental Research: Infrastructure and Sustainability* 2 (4): 045005.

Erickson, Jennifer. 2012. "Top 10 U.S. Government Investments in 20th Century American Competitiveness." Center for American Progress, January 6. https://www.americanprogress.org/article/top-10-u-s-government-investments-in-20th-century-american-competitiveness/.

Federal Energy Regulatory Commission (**FERC**). 2020. *Report on Barriers and Opportunities for High Voltage Transmission*. Washington, DC: United States Department of Energy.

Fell, Harrison, Daniel T. Kaffine, and Kevin Novan. 2021. "Emissions, Transmission, and the Environmental Value of Renewable Energy." *American Economic Journal: Economic Policy* 13 (2): 241–72.

Friedman, Lisa. 2022. "California Reveals Its Plan to Phase Out New Gas-Powered Cars by 2035." *New York Times*, April 13. https://www.nytimes.com/2022/04/13/climate/california-electric-vehicles.html.

Gearino, Dan. 2021. "In California, the World's Largest Battery Storage System Gets Even Larger." *Inside Climate News*, September 2. https://insideclimatenews.org/news/02092021/inside-clean-energy-california-energy-storage-vistra-corp/.

Goggin, Michael, and Rob Gramlich. 2022. *Emissions Reductions from Electricity Transmission Provisions in Energy Permitting Legislation*. Bethesda, MD: Grid Strategies LLC.

Gold, Russell. 2019. *Superpower: One Man's Quest to Transform American Energy*. New York: Simon & Schuster.

Gonzales, Luis E., Koichiro Ito, and Mar Reguant. 2022. "The Dynamic Impact of Market Integration: Evidence from Renewable Energy Expansion in Chile." NBER Working Paper 30016.

Gowrisankaran, Gautam, Stanley S. Reynolds, and Mario Samano. 2016. "Intermittency and the Value of Renewable Energy." *Journal of Political Economy* 124 (4): 1187–1234.

Griffin, James M., and Steven L. Puller. 2009. "Introduction: A Primer on Electricity and the Economics of Deregulation." In *Electricity Deregulation: Choices and Challenges*, edited by James M. Griffin and Steven L. Puller, 1–28. Chicago: University of Chicago Press.

Hirst, Eric, and Brendan Kirby. 2001. *Transmission Planning for a Restructuring U.S. Electricity Industry.* Washington, DC: Edison Electric Institute.

Hogan, William W. 2018. "A Primer on Transmission Benefits and Cost Allocation." *Economics of Energy and Environmental Policy* 7 (1): 25–46.

Imelda, Matthias Fripp, and Michael J. Roberts. Forthcoming. "Real-Time Pricing and the Cost of Clean Power." *American Economic Journal: Economic Policy.*

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy* 10 (1): 240–67.

Jarvis, Stephen. 2022. "The Economic Costs of NIMBYism: Evidence from Renewable Energy Projects." Unpublished.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge Is (Less) Power: Experimental Evidence from Residential Energy Use." *American Economic Review* 104 (4): 1417–38.

Jones, Christopher R., and J. Richard Eiser. 2010. "Understanding 'Local' Opposition to Wind Development in the UK: How Big is a Backyard?" *Energy Policy* 38 (6): 3106–17.

Joskow, Paul L. 2005. "Transmission Policy in the United States." *Utilities Policy* 13 (2): 95–115.

Joskow, Paul L. 2011. "Comparing the Costs of Intermittent and Dispatchable Electricity Generating Technologies." *American Economic Review: Papers and Proceedings* 100 (3): 238–41.

Joskow, Paul L. 2012. "Creating a Smarter U.S. Electricity Grid." *Journal of Economic Perspectives* 26 (1): 29–48.

Joskow, Paul L. 2019. "Challenges for Wholesale Electricity Markets with Intermittent Renewable Generation at Scale: The US Experience." *Oxford Review of Economic Policy* 35 (2): 291–331.

Joskow, Paul L. 2020a. "Competition for Electric Transmission Projects in the USA: FERC Order 1000." In *Transmission Network Investment in Liberalized Power Markets*, edited by Mohammad Reza Hesamzadeh, Juan Rosellón, and Ingo Vogelsang, 275–322. Cham, Switzerland: Springer.

Joskow, Paul L. 2020b. "Transmission Capacity Expansion Is Needed to Decarbonize the Electricity Sector Efficiently." *Joule* 4 (1): 1–3.

Joskow, Paul L. 2021. "Facilitating Transmission Expansion to Support Efficient Decarbonization of the Electricity Sector." *Economics of Energy and Environmental Policy* 10 (2): 57–91.

Joskow, Paul L., and Richard Schmalensee. 1988. *Markets for Power: An Analysis of Electrical Utility Deregulation.* Cambridge, MA: MIT Press.

Joskow, Paul L., and Catherine D. Wolfram. 2012. "Dynamic Pricing of Electricity." *American Economic Review* 102 (3): 381–85.

Kamp, Jon. 2021. "Maine Voters Reject $950 Million Power Line for Hydro Imports." *Wall Street Journal*, November 3. https://www.wsj.com/articles/maine-voters-reject-950-million-power-line-for-hydro-imports-11635944404.

Karaduman, Ömer. 2021. "Economics of Grid-Scale Energy Storage in Wholesale Electricity Markets" MIT CEEPR Working Paper 2021-005.

Lawrence Berkeley National Laboratory. 2022. *Land-Based Wind Market Report: 2022 Edition.* Oak Ridge, TN: U.S. Department of Energy.

Mallapragada, Dharik S., Cristian Junge, Cathy Wang, Hannes Pfeifenberger, Paul L. Joskow, and Richard Schmalensee. 2022. "Electricity Pricing Problems in Future Renewables-Dominant Power Systems." MIT CEEPR Working Paper 2021-017-R.

McCalley, Jim, James Bushnell, Venkat Krishnan, and Santiago Lemos. 2012. "Transmission Design at the National Level: Benefits, Risks and Possible Paths Forward." *Power Systems Engineering Research Center (PSERC) Publication* 12-01.

McCarthy, Elizabeth. 2022. "Lithium-Ion Roadblocks Drive Development of US-Based Alternatives for Grid Battery Storage." *Utility Dive*, April 5. https://www.utilitydive.com/news/lithium-ion-roadblocks-drive-development-of-us-based-alternatives-for-grid/621565/.

Millstein, Dev, Eric O'Shaughnessy, and Ryan Wiser. 2023. "The Renewables and Wholesale Electricity Prices (ReWEP) Tool Version 2023.1." Lawrence Berkeley National Laboratory. https://emp.lbl.gov/renewables-and-wholesale-electricity-prices-rewep (accessed May 19, 2023).

MIT. 2011. *The Future of the Electric Grid. An Interdisciplinary MIT Study.* Cambridge, MA: Massachusetts Institute of Technology.

MIT. 2022. *The Future of Energy Storage. An Interdisciplinary MIT Study.* Cambridge, MA: Massachusetts Institute of Technology.

Mueller, Christoph Emanuel, Silke Inga Keil, and Christian Bauer. 2017. "Effects of Spatial Proximity to Proposed High-Voltage Transmission Lines: Evidence from a Natural Experiment in Lower Saxony." *Energy Policy* 111: 137–47.

National Academies of Sciences, Engineering, and Medicine. 2021. *Accelerating Decarbonization of the U.S. Energy System.* Washington, DC: National Academies Press.

National Renewable Energy Laboratory (NREL) 2022. "Annual Technology Baseline." NREL. https://atb.nrel.gov/electricity/2022/index (accessed August 12, 2022).

National Renewable Energy Laboratory (**NREL**)**.** 2023a. "Wind Resource Maps and Data." NREL. https://www.nrel.gov/gis/wind-resource-maps.html (accessed May 19, 2023).

National Renewable Energy Laboratory (**NREL**)**.** 2023b. "Solar Resource Maps and Data." NREL. https://www.nrel.gov/gis/solar-resource-maps.html (accessed May 19, 2023).

Nemet, Gregory F. 2019. *How Solar Energy Became Cheap: A Model for Low-Carbon Innovation.* Abingdon, UK: Routledge.

Noailly, Joëlle, and Victoria Shestalova. 2017. "Knowledge Spillovers from Renewable Energy Technologies: Lessons from Patent Citations." *Environmental Innovation and Societal Transitions* 22: 1–14.

Olmos, Luis, Michel Rivier, and Ignacio Pérez-Arriaga. 2018. "Transmission Expansion Benefits." *Economics of Energy and Environmental Policy* 7 (1): 47–62.

Pascale, Andrew, Jesse D. Jenkins, and Emily Leslie. 2021. *Princeton's Net-Zero America Study: Annex F: Integrated Transmission Line Mapping and Costing.* Princeton, NJ: Princeton University Press.

Penrod, Emma. 2022. "Why the Energy Transition Broke the U.S. Interconnection System." *Utility Dive,* August 22. https://www.utilitydive.com/news/energy-transition-interconnection-reform-ferc-qcells/628822/.

Permitting Institute. 2023. Federal Permitting Process Flowchart for a Transmission Project. https://www.permittinginstitute.org/transmission-chart (accessed May 22, 2023).

Pfeifenberger, Johannes P., Kasparas Spokas, J. Michael Hagerty, and John Tsoukalis. 2021. *A Roadmap to Improved Interregional Transmission Planning.* Boston: Brattle Group.

Princeton. 2021. "Net-Zero America: Potential Pathways, Infrastructure, and Impacts." Princeton University. https://netzeroamerica.princeton.edu (accessed April 1, 2022).

Ramkumar, Amrith. 2022. "Long-Duration Battery Startup Form Energy Raises $450 Million." *Wall Street Journal,* October 4. https://www.wsj.com/articles/long-duration-battery-startup-form-energy-raises-450-million-11664877603.

Rand, Joseph, Ryan H. Wiser, Will Gorman, Dev Millstein, Joachim Seel, Seongeun Jeong, and Dana Robson. 2022. *Queued Up: Characteristics of Power Plants Seeking Transmission Interconnection as of the End of 2021.* Berkeley, CA: Lawrence Berkeley National Lab.

Reed, Liza, M. Granger Morgan, Parth Vaishnav, and Daniel Erian Armanios. 2019. "Converting Existing Transmission Corridors to HVDC Is an Overlooked Option for Increasing Transmission Capacity." *Proceedings of the National Academy of Sciences* 116 (28): 13879–84.

Ryan, Nicholas. 2021. "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market." *American Economic Journal: Microeconomics* 13 (2): 202–42.

Saul, Josh, Naureen Malik, and Dave Merrill. 2022. "The Clean-Power Megaproject Held Hostage by a Ranch and a Bird," *Bloomberg,* April 12. https://www.bloomberg.com/graphics/2022-clean-energy-power-lines-transwest-wind-maps-private-property/#xj4y7vzkg.

Seel, Joachim, Dev Millstein, Andrew Mills, Mark Bolinger, and Ryan Wiser. 2021. "Plentiful Electricity Turns Wholesale Prices Negative." *Advances in Applied Energy* 4: 100073.

Shao, Elena. 2022., "In a Twist, Old Coal Plants Help Deliver Renewable Power. Here's How." *New York Times,* July 15. https://www.nytimes.com/2022/07/15/climate/coal-plants-renewable-energy.html.

Sioshansi, Ramteen, and Paul Denholm. 2012. *Transmission Benefits of Co-locating Concentrating Solar Power and Wind.* Golden, CO: National Renewable Energy Laboratory (NREL).

Southwest Power Pool (SPP). 2023. "Variable Energy Resources Curtailments." SPP. https://marketplace.spp.org/pages/ver-curtailments# (accessed May 19, 2023).

Stokes, Leah C. 2016. "Electoral Backlash against Climate Policy: A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy." *American Journal of Political Science* 60 (4): 958–74.

Sud, Rayan, Sanjay Patnaik, and Robert Glicksman. 2023. "How to Reform Federal Permitting to Accelerate

Clean Energy Infrastructure." Brookings Center on Regulation and Markets Policy. https://www.brookings.edu/wp-content/uploads/2023/02/20230213_CRM_Patnaik_Permitting_FINAL.pdf.

Tomich, Jeffrey. 2019. "Renewables 'Hit a Wall' in Saturated Upper Midwest Grid," *E&E News EnergyWire*, December 12. https://www.eenews.net/articles/renewables-hit-a-wall-in-saturated-upper-midwest-grid/.

Uría-Martínez, Rocío, Megan M. Johnson, and Rui Shan. 2021. *US Hydropower Market Report*. Oak Ridge, TN: U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy.

US Census Bureau. 2023. "National Population Totals and Components of Change: 2020-2022, Annual Estimates of the Resident Population for the United States, Regions, States, District of Columbia and Puerto Rico." https://www.census.gov/data/tables/time-series/demo/popest/2020s-national-total.html (accessed May 19, 2023).

US Department of Energy. 2015. *Quadrennial Energy Review. Appendix C: Electricity*. Washington, DC: US Department of Energy. https://www.energy.gov/sites/prod/files/2015/07/f24/ElectricityAppendix.pdf.

US Energy Information Administration. 2018. "Utilities Continue to Increase Spending on Transmission Infrastructure." *Today in Energy*, February 9. https://www.eia.gov/todayinenergy/detail.php?id=34892.

US Energy Information Administration. 2019. "Southwestern States Have Better Solar Resources and Higher Solar PV Capacity Factors." *Today in Energy*, June 12. https://www.eia.gov/todayinenergy/detail.php?id=39832#.

US Energy Information Administration. 2021a. *Battery Storage in the United States: An Update on Market Trends*. Washington, DC: U.S. Department of Energy.

US Energy Information Administration. 2021b. "Electricity Demand in Lower 48 States Reached a High of 720 Gigawatthours on August 12." *Today in Energy*, August 19. https://www.eia.gov/todayinenergy/detail.php?id=49216/.

US Energy Information Administration. 2021c. "U.S. Large-Scale Battery Storage Capacity Up 35% in 2020, Rapid Growth Set to Continue." *Today in Energy*, August 20. https://www.eia.gov/todayinenergy/detail.php?id=49236.

US Energy Information Administration. 2021d. "Utilities Continue to Increase Spending on the Electric Transmission System." *Today in Energy*, March 26. https://www.eia.gov/todayinenergy/detail.php?id=47316.

US Energy Information Administration. 2010–2022. *Levelized Costs of New Generation Resources in the Annual Energy Outlook. Annual Reports 2010–2022*. U.S. Department of Energy. Washington, DC: US Energy Information Administration.

US Energy Information Administration. 2022a. *Annual Energy Outlook 2022*. Washington, DC: US Energy Information Administration.

US Energy Information Administration. 2022b. "Prime Supplier Sales Volumes." https://www.eia.gov/dnav/pet/pet_cons_prim_dcu_nus_a.htm (accessed May 19, 2023).

US Energy Information Administration. 2023. "Electricity Data Browser, Monthly Net Generation for All Sectors." https://www.eia.gov/electricity/data/browser/ (accessed May 19, 2023).

Vajjhala, Shalini P., and Paul S. Fischbeck. 2007. "Quantifying Siting Difficulty: A Case Study of U.S. Transmission Line Siting." *Energy Policy* 35 (1): 650–71.

Western Electricity Coordinating Council (WECC). 2019. "Transmission Capital Cost Tool - with E3 Updates." WECC. https://www.wecc.org/Administrative/TEPPC_TransCapCostCalculator_E3_2019_Update.xlsx (accessed May 19, 2023).

Williams, James H., Ryan A. Jones, Ben Haley, Gabe Kwok, Jeremy Hargreaves, Jamil Farbes, and Margaret S. Torn. 2021. "Carbon-Neutral Pathways for the United States." *AGU Advances* 2 (1): e2020AV000284.

The Economics of Electricity Reliability

Severin Borenstein, James Bushnell, and Erin Mansur

On August 14, 2003, a midsized power plant owned by an Ohio electric utility (FirstEnergy) suffered an unplanned shutdown. Shortly thereafter several poorly-maintained large transmission lines failed. By late afternoon, voltage in its service territory had dropped to dangerous levels. The only way to restore stability would have been to interrupt service to a large portion of the Cleveland area, but no such service interruption was implemented. By 4:00 PM, uncontrolled outages began quickly cascading outward from Ohio, first to Detroit and Toronto, and then to Pennsylvania and New York. The outages eventually reached parts of nine US states and most of Ontario, which suffered intermittent blackouts for more than a week. All told, more than 50 million people were affected. Estimates of total costs were \$4–\$10 billion in the US and 0.7% of monthly GDP in Canada (US-Canada Power System Outage Task Force 2004).

Electricity resembles a service much more than a good. It is very costly to store for even seconds, so it must be produced largely at the same time that it is consumed. Demand varies minute to minute, so the barriers to storage mean that suppliers must be responsive to the fluctuations in demand. Unlike most

■ *Severin Borenstein is E. T. Grether Professor of Business Administration and Public Policy Haas School of Business, University of California, Berkeley, California. James Bushnell is Professor of Economics, University of California, Davis, California. Erin Mansur is Revers Professor of Business Administration, Tuck School of Business, Dartmouth College, Hanover, New Hampshire. Their email addresses are severinborenstein@berkeley.edu, jbbushnell@ ucdavis.edu, and erin.mansur@dartmouth.edu.*

services, however, electricity can be transported over relatively long distances at near-zero cost up to transmission capacity, so production and consumption of the product need not be physically proximate. Still, transmission has capacity constraints, so while large areas can be part of the same market most of the time, sub-areas can become isolated quite rapidly. Unlike natural gas or water, electric grids have no valves that can be used to direct electricity to where it is needed most. As a result, the capacity to move electricity between locations is a very complex physics relationship that depends on the demands and supplies at each node of the grid.

In most markets, a temporary supply shortage leads to high prices, isolated stock-outs, or other nonmarket rationing schemes. The physics of an electrical grid, however, means that a supply-demand imbalance can cause two critical characteristics of electricity—voltage and frequency—to deviate from their required levels, which can damage both appliances using the electricity and generation units producing it. To mitigate that risk, generators have protective devices that disconnect them from the grid when large deviations occur. Those protective disconnections, however, can worsen the voltage or frequency deviations on the grid, potentially causing more disconnections and, ultimately, triggering a cascade. Thus, electricity is almost unique among commodities in the way that a local supply-demand imbalance can cascade into widespread service disruptions, potentially affecting millions of customers located far away from the original market imbalance, as happened in 2003.

Despite this possibility of serious negative spillovers from a local imbalance, electricity grids typically cover very large areas due to the value of supply diversification in maintaining supply-demand balance. Even with conventional generation resources, assuring adequate supply is challenging, due to the risk of generator outages and the uncertainty of peak demand levels. Because outages and demand are imperfectly correlated across regions, connecting them into a common grid reduces the cost of capacity needed to lower the probability of a supply shortfall below any given level.

Due to the unique physics of electricity, maintaining on-demand availability to millions of customers requires a precise juggling of real-time delivery systems. Furthermore, the complexity of electricity flow on a grid with millions of connected sources and sinks means that it is not practical in real time to establish which entities are responsible for a supply shortfall or surplus. Because this interdependence stretches across areas served by different electric utilities, extensive rules have been developed over the decades to manage reserves and operational standards in real time.

Standards for operations and reserves help reduce local imbalances and generally prevent them from cascading to neighboring areas. Most economists who study electricity markets agree that relying purely on market forces to provide these types of real-time services would not be efficient, due to imperfect information and the fluctuating, and potentially massive, externalities from a local supply-demand imbalance. However, for long-run investments in electricity generation capacity to maintain grid reliability, there is less agreement on the role of markets versus regulation.

In one sense, the problem faced in electricity supply is similar to any industry in which a complex web of vertically-related firms must coordinate on production and remuneration to deliver a product to consumers. Electricity, however, is possibly the most challenging situation due to the physics of grid stability, the high cost of storage, the shared network of transmission, the mix of for-profit companies with nonprofit or heavily regulated firms, and the critical role this product plays in the functioning of an economy.

In this paper, we begin with a review of the landscape of US regulations and markets created to assure reliability of this unusual, and critical, product. We then turn to the economics of supply and demand balancing in different kinds of electricity spot markets and current approaches to assuring long-run adequacy of electricity supply.

While electricity supply has been highly reliable in the United States, these markets are facing new challenges as they use less "dispatchable" generation, like natural gas, coal, or nuclear, and more "intermittent" sources that fluctuate exogenously, like wind and solar. In addition, environmental concerns are accelerating the electrification of transportation and building energy use. This is both increasing demand for electricity and raising the stakes for reliability as so many services become dependent upon this single source of energy.

A Short Primer on Electricity Regulation and Deregulation

The electricity industry has four main segments: generation, transmission, distribution, and retailing/billing. The first three involve physical hardware to produce and distribute electricity, while the fourth is a procurement and accounting function. Historically, all these segments were vertically integrated within regional utility companies operating monopoly franchises for serving customers in each of their territories (Joskow 1997).

While some vertically-integrated utilities are owned by local governments or associations of governments, the majority of electricity in the United States was, and still is, sold through investor-owned utilities under regulation by state agencies. There are also some utilities—primarily municipal or cooperative distribution utilities—that purchase all of their power and engage only in distribution and retailing/billing.

Decades of regulation of the electricity industry under cost-of-service principles raised concerns about the incentives provided to regulated utilities and their resulting efficiency (Borenstein and Bushnell 2000). Starting in the late 1990s, several US states began restructuring their power sectors. Electricity generators began to earn market prices, and independent power producers could enter into this market. Furthermore, in many regions, the incumbents (the vertically-integrated, investor-owned utilities) were required to sell off their generation or operate it in a separate entity under market prices.

Figure 1
Independent System Operators
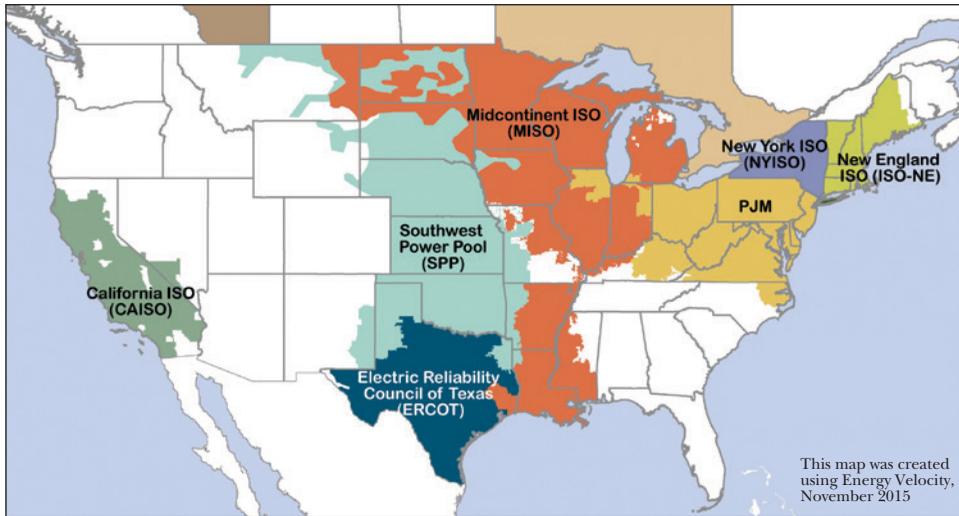
Source: https://www.ferc.gov/sites/default/files/2020-05/elec-ovr-rto-map.pdf.
Note: The colored areas on the map represent the territories of the ISOs. The areas in white remain under traditional vertical integration and local utilities manage these functions independent of neighboring utilities.

In the continental United States, there are three large, physically-connected grids—roughly covering the areas west of the Rocky Mountains, areas east of the Rockies, and Texas—each of which connects many different utility service areas. Prior to the 1990s, most electrical utilities operated their local grids and bore responsibility for reliability within their service areas. There was some limited power trading between utilities. Starting in the 1990s, seven Independent System Operators (ISOs) were created to operate much of the country's high-voltage transmission grid and coordinate a more decentralized power sector, while maintaining supply reliability. These organizations (which are also known as Regional Transmission Organizations) run wholesale electricity markets in which generators bid in supply of electricity and grid balancing services that they provide and retail providers bid in their demands. The ISOs are shown in the map in Figure 1.[1] The ISOs seek to ensure reliable and nondiscriminatory access to transmission systems. ISOs also have responsibility for grid and supply optimization in the short run and transmission and some generation capacity planning in the long run. In Figure 1, the areas in white remain under traditional vertical integration and local utilities manage

---

[1]We are not aware of any systematic difference between organizations that call themselves Regional Transmission Organizations and those that call themselves Independent System Operators. Throughout this paper, we will use ISO to represent either type of organization.

these functions independent of neighboring utilities.[2] Even within ISO systems, incumbent utilities retain control over, and responsibility for, their lower-voltage local distribution circuits.

With deregulation, decisions about the type and amount of investment in new generation capacity shifted from a regulatory forum to a decentralized, market-based process. Power transactions, rather than being internal to a firm or between neighboring utilities, were to be made through a centralized wholesale power market. The Independent System Operators operate these auctions to clear the wholesale energy markets, and they also order minute by minute adjustments in output from generators in order to constantly balance supply and demand at each location. To do this, the ISOs also run markets for "ancillary services," which are short-term commitments by some generators to make capacity available that can increase or decrease output at the request of the ISO. Owners of high-voltage power lines (involving distribution) continue to operate as natural monopolies. As such, they remain under economic regulation at either the state or federal level. As in the decades before, investments in transmission and distribution assets are reviewed by federal and state regulatory bodies and costs are recovered under cost-of-service regulation principles.

Around the same time that many states moved to deregulate electricity generation, a smaller number of states adopted various forms of "retail competition." The phrase means that customers can purchase their electricity from retailers other than the utility that provides local physical distribution of the power. Such competitive retailers need not be in the physical side of the electricity business, and many are not. Instead, they procure electricity from generators under longer-term contracts, or out of the wholesale spot market, and sell electricity to retail customers. In most states with retail competition (including Texas, Ohio, and Massachusetts) the retailers are for-profit companies, but in other states (such as California and Illinois) retail competitors can also, or only, be local government agencies. By the nature of such retail competition, neighboring customers may not be buying power from the same retail provider. All retailers, whether competitive for-profit, competitive nonprofit, regulated investor-owned utility, or government entity, are collectively known as load-serving entities (LSEs).

Importantly, even in states with retail competition, the reliability of a household's electric supply is decoupled from its choice of load-serving entities. The regulated utility distribution company is responsible for delivering power to all retail customers to meet their real-time demand, regardless of which retailer is procuring power for the customer. LSEs are then responsible for covering the wholesale cost of all electricity delivered to their customers. When there is a supply shortfall, stability of the system is maintained through "load shedding," demand reductions achieved

---

[2]Some of the utilities in the white area of the map, while vertically integrated, participate in broader wholesale markets run by Independent System Operators and face some additional regulations as a result. Likewise, some utilities within the territory of an ISO remain vertically integrated, in that they still own significant generation that is subject to state regulation, but by virtue of being part of the ISO, they are required to be part of the ISO's program for assuring adequate generation capacity.

by cutting all power to some customers. Typically, the utility distribution company does this by randomly rotating outages across neighborhoods with no consideration of which retail provider has procured insufficient supply. As of 2021, about 69 percent of all electricity delivered in the United States is in regions that are part of ISOs and about 44 percent is procured in markets with significant retail competition.[3]

Reliability in Electricity Systems

As the experience from the 2003 blackout illustrates, electricity reliability is a function of much more than just adequate investment in generation capacity. In fact, by far the most common cause of electricity service interruptions—blackouts—is a localized failure in the distribution system, such as might be caused by a tree branch falling on a power line. Electricity service interruptions can be categorized as localized distribution outages, larger-area transmission outages, or supply shortfalls, any of which, if not properly managed, can lead to cascading system outages.

When shortages of supply have occurred—with recent examples in California (2020), Texas (2021), Tennessee (2022) and North Carolina (2022)—shortfalls are generally anticipated far enough in advance to manage the shortage without disrupting supply to the vast majority of customers. The Texas energy crisis in February 2021—the largest and most costly service interruption since the 2003 Northeast blackout—was triggered by extremely cold weather that caused many generating units to fail due directly to the impact of low temperatures or indirectly when fuel supplies failed in the frigid weather. The result was that over 25 percent of projected consumption was curtailed. Yet even in this extreme event, the remaining 75 percent of demand continued to be served by the still-operating regional grid (University of Texas Austin Committee 2022). As serious as the Texas crisis was, it did not create cascading outages and far more drastic disruptions, though it came extremely close (Blunt and Gold 2021).

Cascading outages—the most severe and rare type of outage by far—arise when there is a localized shortfall, usually due to the failure of a generation or transmission resource, that is not contained quickly enough by interrupting local customers. This is the distinctive feature of electricity systems: a local supply-demand imbalance effectively can disrupt the grid on a very large scale if not dealt with quickly and properly. In this way a *small* supply shortfall, which in markets for other goods and services would result in rationing supply to a small number of customers, in electricity can result in interrupting service to all demand, not just the amount that is in excess of available supply.

---

[3]Form 861 (Energy Information Administration 2015–2021a) provides data on electricity sales. We define "customers" as being in an Independent System Operator if their local distribution company is within one of the seven ISOs shown in Figure 1. We define significant retail competition as a utility distribution area in which at least 10 percent of retail sales are made by a load-serving entity that is not the local utility distribution company.

Figure 2
Annual Hours of Outages by County
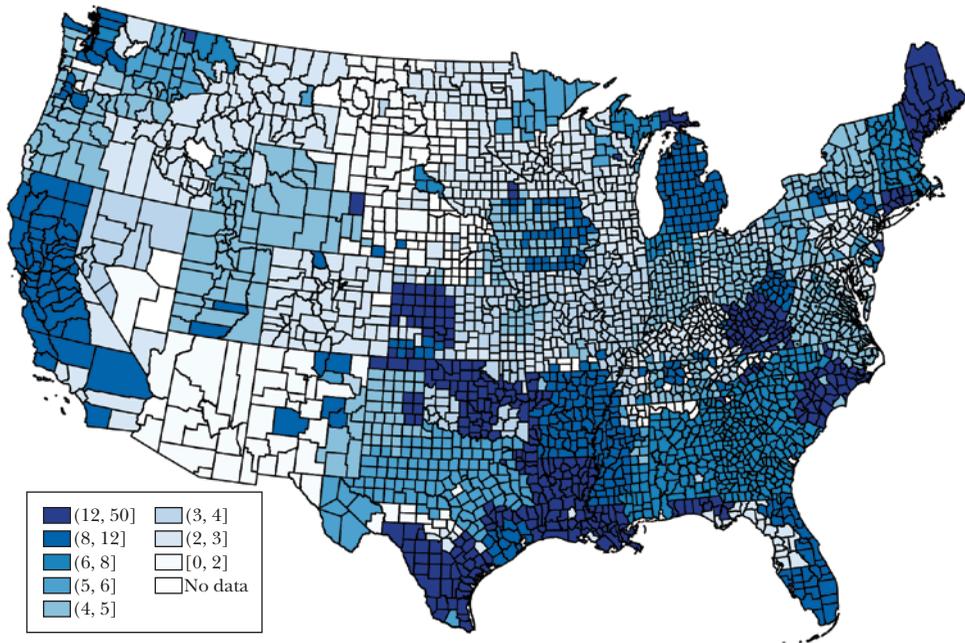
Source: Annual Electric Power Industry Report (Energy Information Administration 2015–2021a).
Note: Data for 2015 to 2020.

Causes and Magnitudes of Outages in the United States

Many electricity outages are unplanned, last-minute responses to weather or issues with the grid as discussed above. However, there are also planned outages that are usually for purpose of maintenance on distribution lines. Recently, planned outages have also been used in California due to the risk of wildfires and the resulting need to de-energize some transmission or distribution lines in order to prevent them from sparking fires (known as "Public Safety Power Shutoffs"). The data in Figure 2 and Table 1 do not distinguish whether the outage was planned or not.

Distribution utilities report information on the frequency and duration of outages.[4] For some electrical utilities, we observe whether an outage was initiated at the distribution system (low voltage) or the transmission network (high voltage). The high-voltage outages might be caused by insufficient generation resources or by problems with the transmission wires. The System Average Interruption Duration Index (SAIDI) measures how many minutes the average customer served by a distribution utility experienced outages for a given year. Another index, the System Average Interruption Frequency Index (SAIFI), measures the frequency of outage

---

[4]See the Annual Electric Power Industry Report (Energy Information Administration 2015–2021a).

events: how many times a year did the average customer at a utility experience an outage.

From 2015 to 2020, customers experienced an average of 1.34 outages a year with an average cumulative duration of 5.67 hours annually. Distribution system outages account for the vast majority (87 percent) of customers' outage minutes, with the balance being due to transmission or system supply shortfalls.[5]

These outages are not distributed evenly throughout the country. While most customers experience only a couple hours of outages annually, the distribution has a long right tail, with some averaging over 15 hours a year. Entergy (a large utility in Louisiana) averages over 38 hours a year, and some small cooperatives are over 100 hours. Figure 2 shows the spatial distribution of outages (in hours per year) by county for the contiguous United States. We see that customers in some states, particularly Louisiana, Maine, Oklahoma, West Virginia, and Connecticut (listed in descending order of duration), experience more than twice as much time without power as the national average. To some extent, high levels of outages are correlated with more extreme weather and more rural locations, but Figure 2 suggests those are not the only drivers. Further study of the locational variation in costs of reliability and the political economy of providing reliability could yield valuable insights.

Available statistics illustrate the fact that the vast majority of local and regional reliability problems in the United States stem from issues related to the delivery, rather than the production, of electricity. The Electric Emergency Incident and Disturbance Report (Form 417, US Department of Energy 2002–2022) lists specific large outages and other major events. Table 1 pools reports over the past 20 years. While the 2003 Northeast blackout and the 2021 Texas crisis are notable, there are large events in most years. In fact, despite the Texas electricity crisis in 2021, that year overall had a similar total number of customers affected and energy losses from outages as other years. Table 1 shows the largest events (reported in millions of customers affected and power losses) from 2002 to 2021 by region and event type. Most are weather-related.

The Economic Cost of Unreliable Electric Supply

While it is clear that power outages are costly to customers, it is much less clear exactly how costly. Within the electricity industry, the cost of an outage is characterized by the "Value of Lost Load" (VOLL), a concept used for planning and investment decisions. Somewhat surprisingly to economists, policy discussions typically are about a single VOLL per kilowatt-hour number, rather than a demand

---

[5] Distribution system outages are the "System Average Interruption Duration Index with Major Event Days Minus Loss of Supply." This index follows the Institute of Electrical and Electronics Engineers (IEEE) standards for measuring an outage duration (SAIDI). We use the measure that includes all outages (a major event day is an "interruption or group of interruptions caused by conditions that exceed the design and operational limits of a system" (see https://www.eia.gov/electricity/annual/html/epa_11_01.html). From this, the utilities remove "loss of supply," which is an outage that was initiated from the high-voltage system.

Table 1
Major Outages by Event Type, State, and Year

Event type	Region	Year	Customers affected (millions)	Power loss (gigawatts)
Cascading Blackout	Northeast US	2003	7.37	78.64
Winter Storm Uri	Texas	2021	2.12	16.41
Hurricane Florence	North Carolina	2018	1.78	15.00
North American derecho	Atlantic/Midwest	2012	8.60	13.61
Hurricane Wilma	Florida	2005	3.24	10.00
Severe Weather	Washington State	2018	4.20	10.00
Transmission	Nevada	2021	1.30	9.00
Severe Weather	Nevada	2020	1.40	8.18
Hurricane Ike	Texas	2008	4.65	8.09
System Operations	Nevada	2021	1.30	8.00
System Operations	AZ/CA	2011	2.00	7.00
Hurricane Isabel	NC/VA	2003	1.80	6.51
Tropical Storm Isaias	New York area	2020	2.94	6.22
Hurricane Frances	Florida	2004	2.78	6.00
Hurricane Matthew	Florida	2016	1.20	5.60
Hurricane Katrina	Louisiana	2005	2.08	5.54
Hurricane Irma	Florida	2017	3.92	4.50
Generation Inadequacy	Texas	2011	1.07	4.00
Wild Fires	California	2014	1.40	3.90
Weather/Transmission	California	2019	0.97	3.19
System Operations	Puerto Rico	2016	1.48	2.75
Cable accidentally cut	California	2005	0.90	2.58
Hurricane Rita	Lower Miss. Valley	2005	1.61	2.30
Severe Weather	Puerto Rico	2011	0.93	2.20
Equipment Trip/Failure	Puerto Rico	2012	0.90	1.80
Severe Weather	Missouri/Illinois	2006	2.50	1.50
Hurricane Charley	Florida	2004	1.20	1.40
Hurricane Jeanne	Puerto Rico	2004	1.42	1.24
Severe Weather	Maryland	2011	0.87	1.11
Hurricane Ivan	Southeast US	2004	0.92	0.92

Note: Power loss is the maximum estimated differential between the quantity demanded by customers and the quantity that utilities were able to deliver.
Source: The Electric Emergency Incident and Disturbance Report (Form 417 [US Department of Energy 2002-2022]).

curve for electricity services with some end-uses producing much greater value than others. Gorman (2022) presents an intellectual history of VOLL and discusses the ways in which it overlaps with standard economic consumer theory and the ways in which it departs. A single VOLL is somewhat consistent with an approach in which retail price is unresponsive to supply/demand balance and rationing is unrelated to the value derived from a particular use by a particular customer. In that case, the aggregate lost gross consumer surplus from a quantity shortfall would, in expectation, be equal to the size of the shortfall multiplied by the average gross consumer surplus across uses, which the VOLL is intended to reflect. Even in that case, however, VOLL fails to account for critical characteristics of outages that would cause the lost consumer surplus to vary, such as weather and other environmental

factors at the time of the outage, the extent of warning customers are given prior to the outage, as well as the size and length of the outage. The 2021 Texas energy crisis, for instance, illustrates that an outage during extreme cold that lasts for multiple days, and in some cases covers large areas so critical electricity services are not available nearby, is likely to be particularly costly.

There is an extensive literature on the economic effects of the availability and reliability of electricity in developing countries, but this literature has not reached consistent findings. Some papers have found relatively modest economic effects in the short run (Dinkelman 2011; Lee, Miguel, and Wolfram 2020; Burlig and Preonas, forthcoming). However, others have found larger effects when the economy more fully adjusts over time (Lipscomb, Mobarak, and Barham 2013; Fried and Lagakos, forthcoming). Blackouts in developing countries have been shown to have economic costs on manufacturers, by altering inputs (Fisher-Vanden, Mansur, and Wang 2015) and changing their scale of operations (Allcott, Collard-Wexler, and O'Connell 2016). Blackouts are transitory shocks that reduce workers' earnings and lead to lower birth weights (Burlando 2014).[6]

There is, however, very little work on the effects of electricity reliability in the United States or in other advanced economies. In part, this is likely because levels of reliability are so high that it is difficult to tease out the longer run impacts of variation in reliability among US states or between countries with developed economies. There are a few studies of specific blackout events, including the Northeast blackout of 2003 and Texas in 2021, but even those extreme events raise substantial estimation challenges. Gorman (2022) discusses some attempts to infer the economic cost of unreliable supply from assumed elasticities of demand, but points out that this omits all of the factors that cause the economic loss to vary across events and customers. Some of the challenges are fairly specific to electricity, such as fixed prices and random rationing, but others are present in a wide range of issues associated with supply shortfalls, such as the correlation of demand with supply shocks and the impact of the shortfall's time span and prior warning of it.

Finally, the electricity industry is currently undergoing a dual transformation that will increase supply from intermittent renewable resources and expand end-use applications of electricity further into transportation and home heating. While it is likely that reliability will continue to be dominated by local weather events and distribution issues, the costs of these outages could rise due to the increasing reliance upon electricity as the sole source of home energy.

---

[6]Additional papers examine another reliability issue not common in industrialized countries, namely, unstable voltage or frequency (Trimble et al. 2016; Zhang 2019; Carranza and Meeks 2021; Berkouwer, Puller, and Wolfram 2021). For example, Meeks et al. (2023) note that voltage fluctuation is a major issue in the Kyrgyz Republic. The authors use a randomized control trial to examine how installing smart meters affects service quality and find that treatment results in less voltage fluctuation and more electricity sales.

Economics of Supply-Demand Balancing in Electricity Spot Markets

During the 1990s and into the 2000s, the United States (and many developed economies) moved to decentralize responsibility for investments in generation capacity. State and regional organizations continued to produce forecasts of demand and of generation resources, but no single entity was tasked with the responsibility of investing in generation capacity. Many economists expected that electricity prices would provide sufficient information and incentive to support investment. Others argued for coordinated procurement mandates, or alternatively for a centralized market for procurement of generation capacity availability, as distinct from the sale of electricity itself.

In this section, we describe why belief in the sufficiency of market-clearing prices—an idea taken for granted in many markets—is the minority view when it comes to electricity (Joskow and Tirole 2007). Our focus is primarily on electricity systems that have deregulated wholesale electricity generation markets, as is the case in most of the United States (and in most developed economies), though we also discuss how the issues manifest in more traditional vertically-integrated service areas.

Little Demand-Side Price Response

The wholesale price of electricity can vary drastically even within a day, due to fluctuating and inelastic demand and supply functions, along with very costly storage. On high-demand days, the wholesale price during the minutes or hour with the tightest supply/demand balance can be ten times or more than the price during lower demand times of the same day. In almost no cases, however, do retail customers see any reflection of those prices. Instead, customers generally face prices that are set months or longer in advance—either a constant price at all times, or higher preset prices during some hours than others. Even such "time-of-use" prices, however, reflect very little of the variation in wholesale prices, because peak demands and fluctuating supply constraints are typically weather-driven and unpredictable months in advance (Borenstein 2005).[7] Retail suppliers, however, are typically required to serve whatever quantity a customer demands at these preset prices, what is known as a "requirements contract."

As a result, the derived demand for electricity in the wholesale market becomes extremely inelastic at a given point in time, regardless of how much consumers would actually respond if they faced retail prices that moved more dynamically with wholesale prices. This absence of real-time price signals to consumers also exacerbates market power concerns, as the inelastic derived demand makes the exercise of market power more profitable in the wholesale market.

---

[7]However, Schittekatte et al. (2022) provide an analysis suggesting that time-of-use pricing may become more reflective of costs under high levels of shiftable loads that may result from electrification of vehicles and buildings.

Price Caps in Wholesale Markets

Producers typically face price caps in the spot market for generating energy. While buyers and sellers generally hedge risk with long-term contracts and trade most of the energy in advance, the prices for those trades are determined knowing that the final spot price is limited. In some cases, price caps may limit the ability of producers to exercise market power. However, they could also result in excess demand if they are set below the competitive level. For example, this could occur if short-run demand is even slightly elastic and fuel prices spike or other factors cause short-run marginal cost to rise above the price cap.

One common argument for capping the price of electricity and wholesale markets is based on the fact that electricity is physically supplied in real time, while financial settlements take place weeks later. For that reason, a buyer might be unaware that they are consuming at an astronomical spot price, and be on the hook much later for that payment.

Correlated Risks in Generation Availability

Supply-demand imbalances become more likely if power plant outages may result from common shocks. For example, a natural gas pipeline accident or extreme cold could limit fuel supply for all gas-fired plants in a region, as happened in Texas in 2021. Similarly, a lull in regional wind associated with extreme heat would limit production from all wind turbines.

For conventional technologies, most unplanned outages are primarily due to uncorrelated shocks, such as equipment failures, though even these events increase with extreme weather. However, as electricity systems decarbonize by increasing generation from intermittent renewable resources—wind and solar—availability will become more correlated across power generation sources either due to known variation like the sun setting or to stochastic events like cloud cover or wind lulls.

Random Rationing Makes Supply Shortfalls a Public Bad

The likelihood of supply-demand imbalances due to the combination of price caps, highly inelastic demand, and correlated risks is heightened when the expected peak demand is nearly as great as the entire aggregate capacity in the system (that is, when the reserve margin is tight). Because such imbalances are addressed by shutting off power by distribution circuit—without regard to willingness to pay to avoid being blacked out or to who helped contribute to the shortage by not investing in, or contracting for, capacity—power shortages are turned into a "public bad" where individual retailers can free ride on one another.

These challenges in wholesale markets imply that there are extraordinary consequences of insufficient capacity investment. Like many capital-intensive industries where firms face uncertain demand in making irreversible investments, power generation can exhibit boom-bust cycles. While other such industries—such as resource extraction and semiconductors—have seen periods of high prices followed by excess entry and a price crash, we do not see many calls for coordinated firm investments in those industries. In fact, where such coordinating entities exist, such as OPEC in

international oil markets, the negative impacts of their collusive activities are typically highlighted. Nonetheless, because of the notable economic challenges discussed here, many argue that there is need for coordinated capacity investment in electricity.

Current Approaches to Long-Run Supply Adequacy

Currently in the United States, there are three general approaches to assuring adequate supply by providing incentives for long-run capacity investment sufficient to meet expected demand, a process called "resource adequacy" within the industry. First, under the traditional electricity industry structure, a monopoly utility invests under either the close regulation or direct ownership of the government. Second, in a deregulated wholesale "energy-only" structure, firms make decentralized and independent investment decisions based largely upon expectations of future electricity prices, similar to the process that drives investment in most other commodity markets. The third approach applies a hybrid of deregulation and centralized planning by imposing capacity procurement requirements on electricity retail service providers operating in deregulated markets.

Of the seven US Independent Systems Operators shown in Figure 1, only ERCOT (Texas) has an energy-only structure, though the approach is also used in Australia, New Zealand, Norway, Sweden, parts of Canada, and other locations. The other US markets follow the third approach and have resource adequacy requirements, in some cases satisfied through a centralized capacity market and in other cases met with bilateral trades. The areas in white remain under traditional vertical integration and manage resource adequacy primarily through the state regulatory oversight process.[8]

It is important to recognize that significant new capacity has been built under all three approaches. Figure 3 shows the cumulative percent of existing capacity (namely those still operating as of 2021) by the year that the capacity was added to the grid. This is not a measure of total capacity in each year, because plants that have retired over this time are not included. The figure shows that substantial investment in new power plants under all three regimes. Natural gas additions account for the majority of new capacity. In restructured regions, wind development has been a notable second category. In contrast, most of the hydropower and nuclear power plants in operation in 2021 were built before this century. Regardless of energy source, we see that about half of the available capacity operating today was built in the past 20 years under all three approaches.

---

[8]Some of the utilities in the white area of the map in Figure 1, while vertically integrated, participate in wholesale markets that have resource adequacy requirements, so the determination of capacity needs is a hybrid of state regulatory oversight and the requirements of the wholesale markets in which they participate. Some utilities within some of the Independent System Operators also remain vertically integrated, in that they still own significant generation that is subject to state regulation, but by virtue of being part of the ISO, they are required to be part of the ISO's resource adequacy program.

Figure 3

Percent of 2021 Existing Capacity Reported Cumulatively by Initial Year of Operation
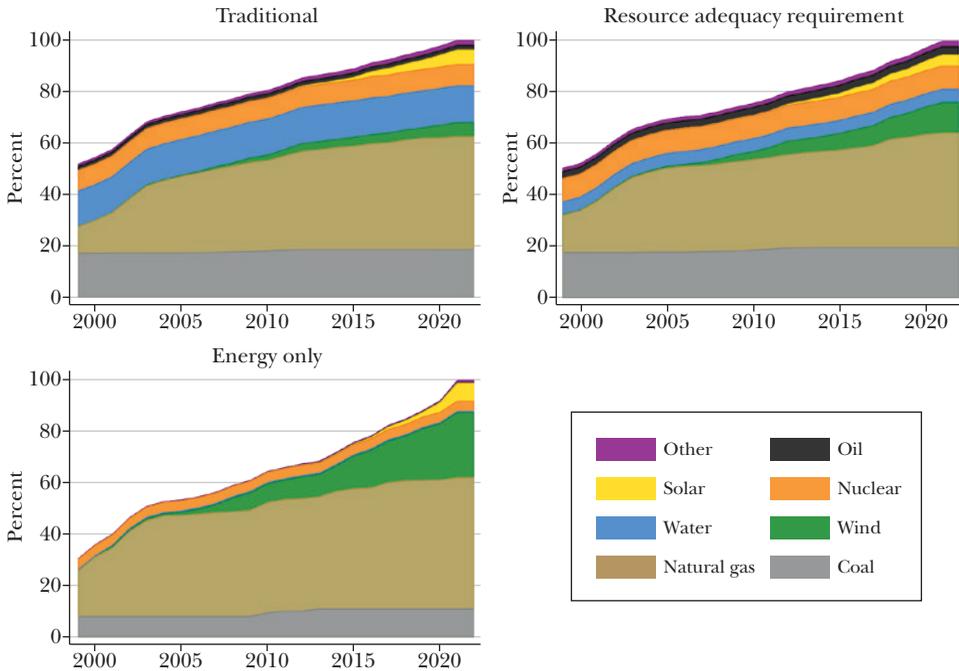
Source: Data are from Form 860 (Energy Information Administration 2021b).
Note: All plants that began operating before 2000 are grouped in 1999.

Traditional Vertical Integration and Regulation

Among the vertically-integrated electric utilities, which still serve about one-third of US demand, the typical resource adequacy process involves joint planning between the utility and its regulators to forecast future demand and establish "needs" for new investment. Specific generation quantities and types, as well as alternatives such as new transmission or demand reduction programs, are negotiated between utility and regulator. Investment in new capacity is then either made directly by the utility or purchased by means of a competitive solicitation overseen by regulators. Once the need for new capacity is established, the recovery of investment costs is largely guaranteed by the regulator, except in the case of extreme cost overruns or gross negligence. The coordination of investment and retirement decisions for both generation and transmission is centralized within a single decision-making process.

At first glance, the combination of regulatory oversight, vertical integration, and monopoly franchise would seem to simplify the process of resource planning. The incentive to free ride on the supply of another retailer is substantially reduced,

though it can still arise to some extent between separate utility distribution areas that are part of the same grid, a practice known in the industry as "leaning."[9]

However, the traditional system lacks incentives for efficient investment given the near-guaranteed recovery of investment costs.[10] The system can create a bias toward capital (Averch and Johnson 1962), although the specific implementation of regulation matters (Joskow 1974). In general, the process can create incentives to overbuild inefficiently, thereby enhancing system reliability but at a potentially inflated cost (Joskow 1997). Indeed, most of the impetus for restructuring the industry inside the United States arose in states with high rates that could be traced to either excessive or inefficient investments in capacity (Borenstein and Bushnell 2000).

Deregulated Wholesale Markets without Resource Adequacy Requirements

As described earlier, the deregulation of generation meant the decentralization of investment decisions in that sector. Previously, regulatory reviews of investment decisions had largely been motivated by a need to justify and approve expenditures that would be added to the capital rate base of a regulated monopoly. With deregulation, the capital invested in generation was no longer guaranteed a regulated rate of return, and the dynamic therefore shifted from a concern over *excess* investment to one of potential inadequacy.

As with markets for most commodities, many deregulated electricity markets around the world rely upon expectations of future prices to provide the signal and incentive for investment in generation capacity. Indeed, wholesale electricity prices are quite sensitive to capacity margins. While prices typically range from $10 to $80 per megawatt-hour, negative prices and prices exceeding $1,000 commonly occur (Table 2).

Because these markets depend on energy prices to signal the need for investment, they tend to feature high price caps and exhibit more volatile spot-market prices. Of all of the Independent Systems Operators, ERCOT has experienced the largest, most frequent, and longest price spikes. While this mirrors the investment process in most other industries, electricity markets face the challenges discussed earlier that exacerbate the size and potential disruption from supply-demand imbalances. Because these imbalances are so costly, grid operators in energy-only markets typically attempt to provide guidance on future demand and other information intended to enable producers to plan more effectively.

Many energy-only wholesale markets around the world are in areas with substantial retail competition. The higher and more volatile energy prices heighten price risk for retailers in energy-only markets. This price risk can provide a stronger

[9]The North American Electric Reliability Corporation (NERC), an industry association, has for decades coordinated standards to prevent leaning between utility distribution areas. Shortly after the 2003 Northeast blackout, the Federal Energy Regulatory Commission gave NERC authority to impose mandatory standards and to enforce penalties for failure to meet them (Nevius 2020).

[10]In theory, investments are evaluated by state regulators based on a "used and useful" criterion. In practice, however, investments are rarely excluded from cost recovery.

Table 2
Summary Statistics of Spot Prices in Wholesale Electricity Markets

	Mean	SD	Min	P10	Median	P90	Max
CAISO (California)	33.87	39.85	−186.32	13.43	29.14	48.93	985.76
ERCOT (Texas)	27.51	44.14	−24.18	14.54	21.99	38.94	5,001.00
ISO-NE (New England)	40.24	45.39	−157.85	14.77	28.52	71.17	2,454.57
MISO (Midwest)	28.19	19.41	−29.94	17.45	24.36	40.80	1,805.60
NYISO (New York)	27.24	26.86	−223.93	10.57	23.01	41.38	927.48
PJM (Mid-Atlantic)	32.17	31.37	−229.98	17.12	26.21	48.11	1,839.28
SPP (Southwest)	23.82	26.62	−57.42	11.46	19.75	35.53	1,592.68

Note: Hourly ISO-wide Average spot prices (in $ per MWh) from January 2013 to December 2020.
Source: FERC Form 714 (for all markets except CAISO) and EnergyOnline (for CAISO).

incentive for retailers to procure, or hedge, their energy in forward markets. Some retailers physically hedge this risk by vertically integrating between generation and retailing functions.[11] Others, however, benefit from bankruptcy laws by offering a fixed retail price and not hedging: if the wholesale spot price ends up low, they make money; if the wholesale price ends up high, they exit. When retailers fail to hedge, however, that reduces the quantity of power purchased through long-term contracts. Such contracts may play an important role in supporting investment in a capital-intensive industry with irreversible investments.

In Texas, the massive reliability problems triggered by winter storm Uri in February 2021 put the energy-only paradigm under greater scrutiny. It is not clear, however, that the typical capacity markets would have coped much better with that disruption. A critical element of winter storm Uri was the inability of much of the existing generation capacity to operate reliably, in some cases due to direct mechanical failures, but in more instances due to the failure of fuel supply infrastructure in the extreme cold. As discussed below, many markets with capacity requirements have weak incentives to ensure reliable performance from the capacity that is procured, and the performance requirements in markets that do feature them are arguably weaker than the $9,000/megawatt-per-hour energy price that was available for any generator that was able to produce during Uri.

Deregulated Wholesale Markets with Capacity Payments

Many restructured power markets have adopted mechanisms that compensate generators for maintaining certain levels of capacity in addition to payments for the electricity produced from that capacity (Joskow 2008). These regions combine

---

[11] Due to the properties of electricity, futures markets have had very limited success. Because electricity prices can vary greatly across nearby locations and times, basis risk greatly reduces the hedging value of a futures contract that specifies a specific time and place of delivery. As a result, forward contracts for electricity, while still fairly standardized in form, do not trade in very liquid markets.

concepts from traditional regulation, where rate-payers paid for the construction and operating costs of power plants, and energy-only markets where customers pay capacity investments through the market prices for electricity. These deregulated wholesale markets have preserved a layer of regulatory planning by creating a distinction between "capacity" and the electrical energy produced by that capacity. In general, capacity needs are set through a coordinated planning process, and in many cases procured centrally and allocated to load-serving entities, while energy is purchased in a more decentralized way. As we will discuss below, however, the distinction between capacity and energy has always been somewhat blurry, and is becoming more so with the advent of new supply technologies.

At a high level, all markets with capacity payments follow a similar process, though there are important distinctions as to how each step in the process is implemented. First, the Independent System Operator coordinates a process to forecast resource capacity need at either a system or individual load-serving entity level. These forecasts range from months to several years into the future. Second, capacity is procured in quantities that are certified to meet forecast needs. The procurement is implemented in some regions by a central entity (such as an ISO) and in others by a mandate applied to individual load-serving entities. And, third, capacity either does or does not perform during periods of tight resource needs. The performance requirements and incentives placed upon the capacity that is procured has varied greatly across regions and over time.

Many extremely contentious regulatory and stakeholder meetings have focused on the process, the amount, and price of capacity that is procured. In the eastern United States, grid operators centrally procure capacity for all load-serving entities, running reverse auctions where producers offer to have capacity available during a specific time period. In California, and much of the Great Plains, load-serving entities (including utilities) are obligated to procure or self-supply an amount of capacity based on the peak demand they serve, similar to an insurance mandate. Somewhat surprisingly, only recently has attention begun to focus on the performance and reliability benefits actually provided by the capacity that is procured.

Part of the argument for capacity payments is the presence of price caps in the associated energy markets, which are in turn justified by concerns over excessive market power in the energy markets. Price caps in electricity energy markets, however, are believed to deny suppliers legitimate scarcity rents at times, creating a so-called "missing money problem" that constrains investment in capacity (Joskow 2006; Cramton and Stoft 2006). Capacity payments are intended to replace those missing scarcity rents (Bushnell 2005).

While the justification for capacity payments can be traced to policies for mitigation of market power, the supply of capacity itself can also be vulnerable to market power. When suppliers are overly concentrated, a mandate to purchase capacity from those suppliers can bestow market power upon them, at least in the short run. This market power can be exacerbated when capacity procurement is divided into localized markets with few sellers (Bowring 2013). Conversely, state governments and regulators have been accused of depressing capacity prices in

an anticompetitive manner by subsidizing local generation through regulatory procurement, tax credits, and other incentives.[12]

The forecasting process entails projecting peak electricity demand needs at either a systemwide or at the level of the load-serving entity. Projections of system-wide demand are more reliable, because they do not require forecasts of the market shares of individual load-serving entities. Partly for this reason, regions that have adopted longer-term capacity requirements—more than a year in advance—tend to do centralized capacity procurement by the Independent Service Operator based on systemwide demand forecasts.

The capacity planning approach has generally relied upon an explicit or implicit assumption that if systems are capable of meeting the hour of highest system demand, they will also be able to operate reliably in all other hours of the planning horizon. In other words, this approach assumes that if a system has enough capacity to meet its peak demand, it will have excess capacity in all other hours. This assumption has always been tenuous when applied to resources for which "nameplate" capacity may not reflect their ability to produce in a particular hour. This is true not only of generation whose output capability fluctuates over time, such as wind and solar, but also "energy-limited" resources such as hydroelectric power and storage.

As these resources have come to comprise a growing share of the mix, the standard planning paradigms have become more stressed. The supply shortages experienced in California during a heat wave in August 2020 provide an illustration of this issue. While California had a capacity requirement in place, it was focused on meeting hours of peak demand, usually in the summertime afternoon. However, the rapid expansion of solar power in California over the last decade left the state with ample supply during peak demand hours, but a potential shortage in the early evening as the sun sets, which became known as the "net demand peak" (net of generation from intermittent renewables). On August 14, 2020, California was forced to implement blackouts around 6:30 PM—more than an hour after demand had peaked—when output was rapidly declining from solar farms, which had been credited with over 3000 megawatt-hours of capacity towards meeting resource adequacy needs for that month (California ISO 2021).[13]

Renewable and hydroelectric resources are not the only ones for which name-plate capacity has at times proven to be a poor measure of reliability contributions. Older fossil power plants have experienced periods of frequent outages, and historic approaches for penalizing such outages have been criticized as too weak. In addition, fuel supply, particularly natural gas, has proven to be a significant contributor

---

[12] Several Independent System Operators have deployed Minimum Offer Price Rules (MOPR), essentially bid floors on supply offers into capacity markets, in an attempt to offset these subsidies. These rules have been controversial in recent years as they have raised the cost of procuring nuclear and renewable resources (Aargaard, Palmer, and Robertson 2022).

[13] Solar generation in California peaked at nearly 11,000 megawatt-hours that day, but was generating 3460 MW at 6:30 PM. It then dropped more than 50 percent by 7 PM and was down to 195 MW by 7:30 PM. Demand, on the other hand, dropped by about 7 percent between 6:30 PM and 7:30 PM.

to reliability problems in several regions, the most notable being the experience of Texas during winter storm Uri in 2021.

Faced with resources whose availability was viewed as unreliable, some regions have adopted more aggressive performance incentives for resources that sell capacity. A major policy question concerns what types of availability problems should be the financial responsibility of the resource and what problems should be considered a *force majeure*. Traditionally, for example, a capacity resource would not be considered responsible for a shortage of natural gas or for low levels of wind. Furthermore, penalties have been relatively modest even for outages that were deemed the responsibility of the resource (Bushnell, Flagg, and Mansur 2017). More recently, some Independent System Operators, such as PJM and ISO-NE, have moved to shift more liability for nonperformance onto the sellers of capacity and have applied steep penalties, on the order of thousands of dollars per megawatt-hour, to resources that are unavailable during a period of regional scarcity. Natural gas shortfalls at some plants in New England during recent cold-weather conditions, however, caused ISO-NE and PJM to trigger performance penalties for only the second time in over five years (Barndollar 2023). By sharply increasing the cost per megawatt-hour of unavailability, such penalties create performance incentives for resources that approach similar levels experienced in energy-only markets. These so-called "performance capacity" policies, however, have not been universally supported. Some critics fear that they shift too much risk to supply resources and could as a result lead to either under-investment or higher capacity prices.

Going Forward: Decarbonization and Technological Change

Electricity policy faces simultaneous challenges: ensuring reliable and affordable power, reducing greenhouse gas emissions, and supporting the expansion of electricity into transportation and heating. Costs have drastically declined for electricity generation from wind and solar—the two technologies most associated with decarbonization—but output volatility from these sources would make supply less reliable if not combined with other resources. Luckily, progress in complementary technologies continues, from energy storage and automated demand response to "firm" carbon-free generation, such as new technologies for nuclear power and geothermal.

While the electrification of transportation and heating will raise overall demand for electricity, the degree to which that strains supply will depend very much on whether these additional demands take place when supply is abundant. Research on electric vehicle usage (Burlig et al. 2021), for instance, indicates that the vast majority of charging is currently done between midnight and 8 AM. Further, both water heating and electric vehicle charging are end-uses of electricity that can probably be shifted across time with relatively little inconvenience to the consumer. For these reasons, electrification will likely create proportionately less additional stress on generation capacity than the increase in energy consumption might imply.

On the other hand, surges in localized demand could stress distribution systems (Elmallah, Brockway, and Callaway 2022). Policies to shift demand will be critical in managing the additional electrical load.

On the supply side, the shift to low-carbon, alternative technologies affects power markets in three important ways: its effect on average wholesale energy prices; its effect on capacity market prices; and the extent to which intermittency of wind and solar, and energy limitations of batteries, create new reliability concerns that are not satisfactorily addressed by conventional resource adequacy crediting. Each of these concerns have been observed in restructured electricity markets. In California, for example, the penetration of utility-scale and rooftop solar has helped drive low, or even negative, energy prices during the middle of the day (Bushnell and Novan 2021). In addition to influencing energy prices, renewable generation has earned an increasing share of capacity payments.

One key policy question, therefore, is whether alternative resources, such as renewable generation or battery storage, can and should provide a comparable form of "capacity" as conventional resources. Such questions get to the heart of a central issue with resource adequacy policy: What exactly constitutes "capacity" under such policies?

Independent System Operators have struggled to define the attributes that constitute capacity, or even to define its units of measurement. Lithium-ion batteries provide a useful illustration. Battery chemistries continue to evolve and with them, their performance characteristics: charging and discharging speeds, round-trip energy loss, and capacity degradation and failure probabilities from different sorts of usage profiles. As an electricity *storage* technology, batteries must also be charged at some point, so their resource adequacy value depends on the ability to charge as well as discharge when they are called upon. The incremental value of storage depends on the dispatchability of the electricity *generating* technologies on the grid. A system with high levels of dispatchable carbon-free generation will derive less value from storage technologies than one more dependent on intermittent renewables.

A number of questions currently debated in the electricity industry highlight the challenges of a paradigm that compensates capacity apart from energy. Should capacity qualifying to provide resource adequacy be limited to resources that can be made available on demand with very high probability, or should it be evaluated based upon a probabilistic expectation of performance? How location specific should capacity procurement be? What performance obligations should be required of the seller? And what should the penalties be for nonperformance?

Performance in an energy-only setting is simply the sale of energy (or ancillary services) in a daily or hourly market. If a unit is operating and selling into the market, it earns revenue. If it is not, then it earns no revenue. Under a capacity payment paradigm, qualified units earn revenue in advance and can keep those earnings, in most cases even if the unit is not available under a long set of possible exemptions. When resource adequacy resources were of similar technologies and were operated by firms with similar incentives, common assumptions about availability did not

distort procurement very much. However, with more diverse resources, Independent System Operators are again revisiting their assumptions about performance and the incentives provided to resources committed through resource adequacy markets.

Resource mix and resource adequacy paradigms also have implications for energy markets. Increasing generation shares from intermittent resources without substantial cost reductions in storage or other complementary technologies will lead to growing wholesale price volatility. More and more hours will have zero or negative prices, and a small number of hours will generate the vast majority of producer rents from the wholesale market. Years could go by with constant excess supply and low prices, and at other times there could be long periods of very high prices, as occurred recently in the Australian energy-only market. While in theory both sides of the market can insure against such fluctuations through either long-term contracting with one another or third-party insurance, in practice such volatility may undermine confidence in wholesale markets. Furthermore, such volatility creates its own opportunities for unproductive strategic behavior, including load-serving entities using bankruptcy as an option when wholesale prices climb and they are inadequately hedged, as well as generators using tight wholesale markets to exercise market power. These possibilities may suggest a continued role for at least some sort of government-mandated level of insurance.

Incentives and Mandates for Performance

A capacity market would have no value if resources were not expected to be able to produce energy when the market was tight. Here we examine how capacity markets are being modified to consider incentives and mandates to achieve performance. In their review of the NYISO capacity market, Harvey, Hogan, and Pope (2013) note the following:

> The larger the total revenues collected through the capacity market rather than the energy or ancillary service market, the greater the concern with the many inherent approximations that appear in the necessary simplifications of the complex problem of constructing forward estimates of resource requirements and defining administrative requirements to provide appropriate performance and investment incentives for capacity suppliers.

When the types of capacity being procured were relatively similar, the simplifications and assumptions created less bias among resource types in procurement. These stresses have become more significant with the increased use of alternative resources to meet capacity needs. The likelihood that a natural gas power plant will deliver during a system emergency, for instance, is less difficult to forecast than a windfarm. This has left the designers of resource adequacy policies with two choices: (1) further refine and categorize the types of capacity to be required; or (2) increase reliance on performance incentives to provide signals about the characteristics and performance abilities of new capacity.

Independent System Operators are taking a diverse approach to this choice. Harvey, Hogan, and Pope (2013) strongly support an emphasis on performance incentives, arguing that "attempting to use capacity market rules to elicit capacity resources with the optimal mix of characteristics to meet load over the operating day has the potential to become more and more difficult as the diversity of the resource mix increases and has the potential to end badly, resulting in both lower reliability and higher consumer cost." In New England, ISO-NE has also shown a preference for strong performance incentives that would be uniformly applied to all resources selling capacity. ISO-NE argues that performance incentives are the key to inducing flexible resources necessary to complement intermittent supply: "Changes to the [forward capacity market] that improve incentives for resource flexibility and availability will provide better incentives for investment in resources that can balance intermittent power supply" (Independent System Operator New England 2012).

Conversely, in California the CAISO, in conjunction with California state agencies, has been incrementally working towards a setting with multiple nested capacity requirements. In addition to a standard resource adequacy requirement that is applied to all participating load-serving entities, the California Public Utilities Commission adopted a "flexible" (or fast responding) capacity procurement requirement in 2014. The requirement for the first time explicitly distinguishes types of capacity by operational characteristics. Other resource adequacy requirements and capacity markets differentiate resources by location, and reduce their qualifying capacity through availability metrics, but do not place explicit limitations based upon an ability to respond on demand to operational orders.

The California approach highlights many of the difficulties inherent in specifying not just a quantity of capacity, but also a range of operational requirements in a resource adequacy context. If fast ramping capability is a key need, must such capability be available for a full hour or smaller intervals? Must resources be available all the time, during peak needs, or during "shoulder ramping periods" (early mornings and late evenings when market demand changes substantially)? The difficulties have been magnified by the need to compare dramatically different resource types, including energy-limited storage, intermittent renewables, conventional generation, and demand flexibility. The comparison of the California approach with the ISO-NE approach illustrates well the trade-off between granular specification of the capacity *characteristics* versus an economic approach to setting *performance* incentives.

The emergence of new resources and technologies is also causing a reassessment of appropriate levels of energy price caps. In the past, prices in the $1,000 per megawatt-hour range could safely be thought to be well above the marginal cost of any generation resource. Debates over price caps therefore centered on the long-run implications of denying suppliers sufficient scarcity rents. However, the growing prominence of batteries and other technologies, along with the potential for more active participation by demand, alters this logic. Opportunity costs and willingness to pay could easily rise above $1,000 per megawatt-hour. Therefore,

higher price caps may be necessary for efficient market clearing in the short run. Such an outcome would dilute one of the main justifications for capacity payments—that price caps deny suppliers necessary scarcity rents.

Technology and Reliability

A bedrock assumption behind resource adequacy standards and policies, though typically unstated, is that customer preferences for supply reliability are uniform and that they are very high. Preferences, however, are not identical and many customers likely would have a willingness to pay for resource adequacy well below the level imposed upon them by these structures. For instance, Cramton and Stoft (2006) map the ubiquitous "one outage in 10 years" standard for system power shortages to an implied Value of Lost Load. Using $80,000 per megawatt-year as the cost of capacity, they translate the one-in-ten standard to a value-of-lost-load of $267 per kilowatt-hour, which is over 1,000 times greater than the average retail price and equivalent to paying more than $1,000 per hour to run a home central air conditioner. While some uses may have such a high value, there are clearly many uses that customers would avoid if faced with such a high price.[14] As described above, the basis for such standards, similar to the basis for resource adequacy policy, is to prevent negative spillovers, or the "free-riding" of one load-serving entity on the resources of others (Spees, Newell, and Pfeifenberger 2013). This is predicated upon the notion that it is impossible to identify and implement the reliability preferences of individuals or communities.

The advancement of technology provides an opportunity to revisit these assumptions. The American Recovery and Reinvestment Act of 2009 provided $4.5 billion for "smart grid" technologies (Joskow 2012), and the 2022 Infrastructure Investment and Jobs Act includes provisions for $13 billion to modernize the electric grid. Even with these investments, grid operators are likely still many years from being able to identify supply and demand at the level of the load-serving entity in real time. Nonetheless, smart meters and other monitoring technologies allow forensic analysis to identify after-the-fact when a load-serving entity was providing insufficient capacity, and levy penalties that could help deter such repeated behavior.[15]

These developments imply that it may be possible to retreat from the axiomatic belief that reliability is a public good. Certainly within short operational time frames, shared responsibility for operating reserves will be necessary for the

---

[14] This is well in excess of most estimates of Value of Lost Load (VOLL). Some may argue, however, that getting to the point of load shedding increases the risk of cascading outages, for which the cost is much greater than one household's lost air conditioning. Still, we know of no research drawing that connection or suggesting how large that increased risk might be.

[15] Ironically, most of the country operated their interconnected control areas in such a fashion before the onset of regional Independent System Operators. Each individual utility was responsible for balancing its load through internal resources and voluntary exchanges with neighboring regions. The temptation to free ride on a neighbor's supply, always technically possible for interconnected control areas, was tempered by NERC oversight and the prospect of serious penalties that could be imposed later for "leaning" on a neighbor's system.

foreseeable future. However, over longer planning horizons it may be possible to identify control areas or individual load-serving entities that have failed to provide adequate resources and impose substantial penalties for their impact on the reliability of other customers. Ultimately it may become possible to interrupt only the customers of the inadequate service providers, although this would require being able to identify culpability for supply shortfalls in near real-time.

Thus, with emerging technologies and creative market design, it may be possible to allow individual load-serving entities to approach their resource acquisition according to their individual choices and beliefs about the market rather than through a standardized set of metrics and rules. Disagreements between local regulators and Independent System Operators about the likely effect of energy efficiency programs, intermittent supply, demand response, or even conventional generation can be put to the test by allowing local load-serving entities to make their choices, but also live with the consequences.

Evolving technologies also enhance prospects for increasing the price elasticity of end-use electricity demand. The combination of widespread smart-meters, growing adoption of home automation, and the expansion of end-uses such as water heating and charging of electric vehicles should lower the technical costs of shifting electricity demand in response to prices. However, longstanding regulatory resistance to the adoption of more dynamic retail electricity prices would also have to be overcome.

At the same time, technological change is making the standard capacity paradigms less and less tenable. With greater resource heterogeneity, it is becoming more difficult to know what combination of resources optimally balances cost minimization and reliability maximization. Relying upon capacity obligations or capacity markets to cost-effectively provide grid stability depends critically on accurately crediting the contribution of different technologies towards resource adequacy. That is challenging even in a technologically static setting, because the value of any one resource depends on the overall mix of resources. It is even more challenging when technologies are changing and operators are constantly learning how best to use them.

References

Aargaard, Todd, Karen Palmer, and Molly Robertson. 2022. "Three Insights from the Debate over Minimum Price Offer Rules in Electricity Markets." *Resources*, July 6. https://www.resources.org/common-resources/three-insights-from-the-debate-over-minimum-offer-price-rules-in-electricity-markets/.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D. O'Connell. 2016. "How Do Electricity Shortages Affect Industry? Evidence from India." *American Economic Review* 106 (3): 587–624.

Averch, Harvey, and Leland L. Johnson. 1962. "Behavior of the Firm under Regulatory Constraint." *American Economic Review* 52 (5): 1052–69.

Barndollar, Hadley. 2023. "Power Plants Face Millions in Penalties after Failure during Christmas Eve Storm." *Providence Journal*, January 7. https://www.providencejournal.com/story/news/2023/01/07/iso-new-england-electricity-power-plant-fines-december-failure-oil-natural-gas-christmas-storm/69785296007/.

Berkouwer, Susanna, Steven Puller, and Catherine Wolfram. 2021. "The Economics of Grid Reliability and Voltage Improvements." Unpublished.

Blunt, Katherine, and Russell Gold. 2021. "Texas Power Grid Was Minutes from Collapse during Freeze, Operator Says." *Wall Street Journal*, February 24. https://www.wsj.com/articles/texas-power-grid-was-minutes-from-collapse-during-freeze-operator-says-11614202063.

Borenstein, Severin. 2005. "Time-Varying Retail Electricity Prices: Theory and Practice." In *Electricity Deregulation: Choices and Challenges*, edited James M. Griffin and Steven L. Puller, 317–56. Chicago: University of Chicago Press.

Borenstein, Severin, and James Bushnell. 2000. "Electricity Restructuring: Deregulation or Reregulation?" *Regulation* 23 (2): 46–52.

Bowring, Joseph. 2013. "Capacity Markets in PJM." *Economics of Energy and Environmental Policy* 2 (2): 47–64.

Burlando, Alfredo. 2014. "Transitory Shocks and Birth Weights: Evidence from a Blackout in Zanzibar." *Journal of Development Economics* 108: 154–68.

Burlig, Fiona, and Louis Preonas. Forthcoming. "Out of the Darkness and into the Light? Development Effects of Electrification in India." *Journal of Political Economy***.**

Burlig, Fiona, James Bushnell, David Rapson, and Catherine Wolfram. 2021. "Low Energy: Estimating Electric Vehicle Electricity Use." *AEA Papers and Proceedings* 111: 430–35.

Bushnell, James. 2005. "Electricity Resource Adequacy: Matching Policies and Goals." *Electricity Journal* 18 (8): 11–21.

Bushnell, James, Michaela Flagg, and Erin Mansur. 2017. "Capacity Markets at a Crossroads." University of California, Berkeley Energy Institute at Haas Working Paper 278.

Bushnell, James, and Kevin Novan. 2021. "Setting with the Sun: The Impacts of Renewable Energy on Conventional Generation." *Journal of the Association of Environmental and Resource Economists* 8 (4): 759–96.

California ISO. 2021. *Final Root Cause Analysis: Mid-August 2020 Extreme Heat Wave*. Folsom, CA: California ISO.

Carranza, Eliana, and Robyn Meeks. 2021. "Energy Efficiency and Electricity Reliability." *Review of Economics and Statistics* 103 (3): 461–75.

Cramton, Peter, and Steven Stoft. 2006. "The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem." MIT Center for Energy and Environmental Policy Research 06-007.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101 (7): 3078–3108.

Elmallah, Salma, Anna M. Brockway, and Duncan Callaway. 2022. "Can Distribution Grid Infrastructure Accommodate Residential Electrification and Electric Vehicle Adoption in Northern California?" *Environmental Research: Infrastructure and Sustainability* 2 (4): 045005.

Energy Information Administration. 2015–2021a. "Annual Electric Power Industry Report, Form EIA-861." https://www.eia.gov/electricity/data/eia861/2015-2021 (accessed August 17, 2022).

Energy Information Administration. 2021b. "Form EIA-860 Detailed Data." https://www.eia.gov/electricity/data/eia860/2021 (accessed June 2, 2022).

Fisher-Vanden, Karen, Erin T. Mansur, and Qiong (**Juliana**) **Wang.** 2015. "Electricity Shortages and Firm

Productivity: Evidence from China's Industrial Firms." *Journal of Development Economics* 114: 172–88.

Fried, Stephie, and David Lagakos. Forthcoming. "Electricity and Firm Productivity: A General-Equilibrium Approach." *American Economic Journal: Macroeconomics*.

Gorman, Will. 2022. "The Quest to Quantify the Value of Lost Load: A Critical Review of the Economics of Power Outages." *Electricity Journal* 35 (8): 107187.

Harvey, Scott M., William W. Hogan, and Susan L. Pope. 2013. *Evaluation of the New York Capacity Market*. Washington, DC: FTI Consulting.

Independent System Operator New England. 2012. "FCM Performance Incentives White Paper."

Joskow, Paul L. 1974. "Inflation and Environmental Concern: Structural Change in the Process of Public Utility Price Regulation." *Journal of Law and Economics* 17 (2): 291–327.

Joskow, Paul L. 1997. "Restructuring, Competition and Regulatory Reform in the U.S. Electricity Sector." *Journal of Economic Perspectives* 11 (3): 119–38.

Joskow, Paul L. 2006. "Competitive Electricity Markets and Investment in New Generating Capacity." AEI-Brookings Joint Center Working Paper 06-14.

Joskow, Paul L. 2008. "Capacity Payments in Imperfect Electricity Markets: Need and Design." *Utilities Policy* 16 (3): 159–70.

Joskow, Paul L. 2012. "Creating a Smarter U.S. Electricity Grid." *Journal of Economic Perspectives* 26 (1): 29–48.

Joskow, Paul, and Jean Tirole. 2007. "Reliability and Competitive Electricity Markets." *RAND Journal of Economics* 38 (1): 60–84.

Lee, Kenneth, Edward Miguel, and Catherine Wolfram. 2020. "Experimental Evidence on the Economics of Rural Electrification." *Journal of Political Economy* 128 (4): 1523–65.

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5 (2): 200–231.

Meeks, Robyn C., Arstan Omuraliev, Ruslan Isaev, and Zhenxuan Wang. 2023. "Impacts of Electricity Quality Improvements: Experimental Evidence on Infrastructure Investments." *Journal of Environmental Economics and Management*. https://doi.org/10.1016/j.jeem.2023.102838.

Nevius, David. 2020. *The History of the North American Electric Reliability Corporation*. Atlanta: North American Electric Reliability Corporation.

Schittekatte, Tim, Dharik S. Mallapragada, Paul L. Joskow, and Richard Schmalensee. 2022. "Electricity Retail Rate Design in a Decarbonized Economy: An Analysis of Time-of-Use and Critical Peak Pricing." NBER Working Paper 30560.

Spees, Kathleen, Samuel A. Newell, and Johannes P. Pfeifenberger. 2013. "Capacity Markets—Lessons Learned from the First Decade." *Economics of Energy and Environmental Policy* 2 (2): 1–26.

Trimble, Christopher Philip, Masami Kojima, Ines Perez Arroyo, and Farah Mohammadzadeh. 2016. "Financial Viability of Electricity Sectors in Sub-Saharan Africa: Quasi-fiscal Deficits and Hidden Costs." World Bank Policy Research Working Paper 7788.

University of Texas Austin Committee. 2021. *The Timeline and Events of the February 2021 Texas Electric Grid Blackouts*. Austin, TX: University of Texas at Austin Energy Institute.

US-Canada Power System Outage Task Force. 2004. *Final Report on the August 14, 2003 Blackout in the United States and Canada: Causes and Recommendations*. Ottawa, ON and Washington, DC: Natural Resources Canada and US Department of Energy.

US Department of Energy. 2000–2022. "Electric Disturbance Events (OE-417) Annual Summaries." https://www.oe.netl.doe.gov/OE417_annual_summary.aspx 2002-2022 (accessed August 15, 2022).

Zhang, Fan. 2019. *In the Dark: How Much Do Power Sector Distortions Cost South Asia?* Washington, DC: World Bank.

The Economics Profession's Socioeconomic Diversity Problem

Anna Stansbury and Robert Schultz

Efforts to track diversity in the workplace rarely consider socioeconomic background explicitly, whether in academia (Kniffin 2007; Oldfield 2007; Lee 2015) or other elite professions (Laurison and Friedman 2020). In the US economics profession, it is well documented that women and racial and ethnic minorities are underrepresented, relative both to the general population and to many other academic disciplines (for example, Bayer and Rouse 2016; Bayer and Wilcox 2019; Lundberg and Stearns 2019; Wessel, Sheiner, and Ng 2019; Lundberg 2020; Bayer, Hoover, and Washington 2020; Ginther and Kahn 2021). However, less is known about the socioeconomic backgrounds of those in the economics profession, largely because such data are scarcer than data on gender and race/ethnicity.

In this paper, we turn to data from the National Science Foundation's Survey of Earned Doctorates, which is an annual census of all individuals who receive a research doctorate from an accredited US institution in a given academic year (for a description of the survey, see https://ncses.nsf.gov/pubs/nsf20301/survey-description). These data allow us to compare the socioeconomic background of economics PhD recipients in the United States with PhD recipients in other disciplines. To proxy for socioeconomic background, we use the highest education level attained by a parent (or guardian) of the PhD recipient. Parental education is one of the three

■ *Anna Stansbury is Assistant Professor of Work and Organization Studies, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts. Robert Schultz is a Researcher for SoFi, San Francisco, California. When this paper was written, he was a graduate Research Assistant, University of Michigan Institute for Social Research, Ann Arbor, Michigan. Stansbury is the corresponding author at amms@mit.edu.*

most commonly used measures of socioeconomic status, alongside household income and parental occupation (neither of which is included in the data).

We begin by documenting four core facts. First, economics PhD recipients are substantially more likely to have highly educated parents, and less likely to have parents without a college degree, than PhD recipients in other disciplines. Second, the gap between economics and other disciplines is particularly stark for PhD recipients born in the United States. In this group, economics stands out as the least socioeconomically diverse even among narrowly defined PhD fields, with a smaller share of first-generation college students than subjects typically considered highly socioeconomically elite like art history or classics. Third, the gap in socioeconomic diversity between economics and other PhD fields has widened substantially since 1970. Fourth, the lack of socioeconomic diversity is particularly stark in higher-ranked economics PhD programs.

We then discuss the relationship between socioeconomic diversity and racial/ethnic and gender diversity in economics. We also discuss possible drivers of the lack of socioeconomic diversity in economics, with some emphasis on the pipeline of undergraduate majors and institutions that produce the bulk of economics PhD students.

The lack of socioeconomic diversity among PhDs in general, and its disproportionate lack among economics PhDs, means that the economics profession is highly unrepresentative of the US population. Relative to the similar-aged US population, we estimate that US-born economics PhD recipients are around five times less likely to have a parent with no college degree, and five times more likely to have a parent with a graduate degree. Such patterns suggest the economics profession is missing out on both talent and perspectives.

Core Facts and Context

Our data from the Survey of Earned Doctorates covers all US PhD recipients from 1970 to 2021. We focus primarily on the most recent period 2010–2021, for which we have data on 611,291 PhD recipients from US institutions, of which 12,911 were in economics.

We measure socioeconomic background with the highest level of parental education. Specifically, we group PhD recipients into three categories: (1) at least one parent or guardian with a graduate degree (any master's, professional, or doctoral degree); (2) at least one parent or guardian with a bachelor's degree but no parent with a graduate degree; and (3) no parent or guardian with a bachelor's degree (this group includes those with a parent or guardian who has an associate's degree or some college, is a high school graduate, or has less than a complete high school education).[1] Parental education is one of the three most commonly used

---

[1] In our data, 13 percent of responses overall, and 14 percent in economics, are missing data on parental education. Select summary statistics for our data are shown in the online Appendix Tables A1–A5.

indicators of socioeconomic background in academic research, alongside parental incomes and occupations (Duncan, Featherman, and Duncan 1972; Hauser 1994). Parental education is a strong predictor of family income, which is associated with students' greater access both to the financial resources and to the opportunities that may enable them to succeed at school and in higher education (Sirin 2005). In addition, even conditional on family income, higher parental educational attainment can provide students with a better understanding and awareness of the opportunities available to them in higher education and the strategies needed to access and succeed in these opportunities. Indeed, a large literature studies the impact of socioeconomic status as measured by parental education on academic achievement, including access to and success in graduate school (for example, Ethington and Smart 1986; Mullen, Goyette, and Soares 2003; Walpole 2003; Morgan et al. 2022).

To compare economics with other academic fields, we primarily use the 14 "major field" categories as defined by the National Science Foundation: agriculture, biological/biomedical sciences, health sciences, engineering, computer and information sciences, mathematics, physical sciences, psychology, social sciences, humanities, education, business management/administration, communication, and other or unknown. In this schema, the "social sciences" category includes economics, so when we use these categories, we will break economics out separately. The National Science Foundation also divides these 14 major PhD fields into 341 smaller categories. Some of these are very small, so when we refer to them, we will commonly look only at those with some minimum number of PhDs in our dataset.

Fact #1: Economics PhD recipients are substantially more likely to have highly educated parents than PhD recipients in other disciplines.

Economics is substantially less socioeconomically diverse than the average PhD field. As shown in Table 1, compared with the 14 major PhD fields, economics has the lowest share of first-generation college graduates and the second-highest share of people with a parent with a graduate degree. Among all PhD recipients over 2010–2021 for whom we have data on parental education, 24 percent of economics PhD recipients had no parent with a bachelor's degree, compared with an average of 30 percent across PhD fields. Conversely, 48 percent of economics PhD recipients had at least one parent with a graduate degree, compared with an average of 43 percent across PhD fields.

As one of several social sciences, our comparison of economics with the 14 large PhD fields may be unfair to economics: there may be other smaller subfields within individual sciences, humanities, or other social sciences which are similarly nondiverse on socioeconomic background. When we restrict the comparison to the 143 narrowly defined fields with more than 1,000 PhD graduates over 2010–2021, economics remains one of the least socioeconomically diverse, ranking 17th lowest in terms of the share with no parent with a bachelor's degree, and 35th highest in terms of the share with at least one parent with a graduate degree.

Table 1

Parental Education of US PhD Recipients, 2010–2021

PhD field	Number of PhD recipients	Share US-born	Share with no parent with bachelor's degree			Share with parent with graduate degree (including PhD)			Share with parent with PhD		
			All	US-born	Foreign-born	All	US-born	Foreign-born	All	US-born	Foreign-born
All	**611,291**	**57%**	**30%**	**26%**	**36%**	**43%**	**50%**	**35%**	**12%**	**12%**	**11%**
Agriculture	15,200	50%	35%	26%	44%	38%	47%	28%	10%	11%	9%
Bio Sciences	98,715	62%	26%	23%	32%	47%	52%	39%	12%	13%	12%
Bus/Management	17,005	43%	32%	30%	33%	41%	48%	36%	13%	15%	12%
Communication	7,047	67%	31%	30%	34%	43%	46%	38%	11%	10%	13%
Comp/Inf Sciences	24,115	30%	28%	19%	32%	43%	60%	36%	14%	19%	12%
Economics	**12,911**	**32%**	**24%**	**13%**	**29%**	**48%**	**65%**	**40%**	**15%**	**20%**	**13%**
Education	55,331	78%	42%	40%	49%	36%	38%	27%	8%	8%	8%
Engineering	107,786	33%	29%	19%	35%	40%	55%	33%	11%	14%	10%
Health Sciences	27,212	64%	36%	35%	39%	39%	42%	34%	9%	9%	9%
Humanities	59,440	74%	25%	22%	33%	52%	55%	42%	13%	13%	14%
Mathematics	21,313	42%	31%	19%	40%	42%	57%	30%	13%	17%	10%
Other or Unknown	11,954	63%	35%	33%	39%	41%	45%	36%	10%	10%	12%
Phys Sciences	67,632	53%	31%	22%	41%	41%	51%	30%	11%	13%	9%
Psychology	42,738	82%	29%	28%	35%	47%	48%	41%	10%	10%	11%
Social Sciences	42,892	65%	30%	28%	36%	47%	51%	38%	13%	13%	12%

Source: Survey of Earned Doctorates.

Notes: "Social Sciences" excludes economics. "Graduate degree" includes master's degrees, professional degrees (including JD, MD, MBA), and PhDs.

Fact #2: The socioeconomic gap between economics and other disciplines is particularly stark for PhD recipients born in the United States.

PhD recipients from American universities come from many different countries, and economics is particularly international: 68 percent of economics PhD recipients 2010–2021 were born outside the United States, as compared to 43 percent of all PhD recipients across fields. The mix of origin countries also differs across PhD fields. Across different countries, parental education is likely to convey different information about socioeconomic background: some countries have higher or lower average educational attainment levels, and the correlation between socioeconomic status and graduate degree attainment differs across countries depending on the degree to which high status professions require graduate qualifications (for example, whether or not law or medicine requires a separate graduate degree). Because of this, for the rest of the paper we analyze US-born and foreign-born PhD students separately.[2]

[2] Online Appendix Figure A1 shows the share of PhD recipients born in the United States, and online Appendix Figure A2 shows the breakdown of foreign-born PhDs by continent of birth, across PhD fields. Data on country of birth are missing for 6 percent of respondents; these are excluded when we analyze US-born and non-US-born separately. Because we do not have data on the location of respondents' childhood or pre-university education, note that some foreign-born individuals in our data may have spent their childhood in the United States. These patterns are very similar if the sample is restricted to US citizens and permanent residents (rather than US-born), as shown in online Appendix Figure A3.

When looking only at US-born PhD recipients, economics is even more striking in its lack of socioeconomic diversity (as shown in Table 1). Across fields, on average 26 percent of US-born PhD recipients had no parent with a bachelor's degree. For economics, this figure was 13 percent—a substantially smaller share than any other major PhD field.

Also, 65 percent of US-born economics PhD recipients had at least one parent with a graduate degree, compared with an average of 50 percent for PhD recipients across all fields. Parents with graduate degrees may have a PhD, or non-PhD graduate degrees like an MD, JD, or MBA. Having a parent with a PhD may confer a different set of academia-specific preferences or advantages, as compared to a parent with a non-PhD graduate degree (which instead may be more reflective of generalized socioeconomic advantage). Economics stands out on both metrics: 20 percent of US-born economics PhD recipients had at least one parent with a PhD, compared to an average of 12 percent across PhD fields; and 45 percent of US-born economics PhD recipients had at least one parent with a non-PhD graduate degree, compared to an average of 38 percent across PhD fields.

Comparing economics to narrow PhD fields, its unrepresentativeness is even more stark. Of the 167 narrow PhD fields for which there were more than 500 US-born PhD recipients over 2010–2021, economics comes last: looking only at US-born PhD recipients, economics has the lowest share of recipients who have no parent with a bachelor's degree and the highest share of recipients who have at least one parent with a graduate degree. This makes economics even less socioeconomically diverse than art history or classics, two fields which are often considered highly socioeconomically elite.[3]

Foreign-born PhD recipients tend to have parents with less formal education, when compared to US-born PhD recipients. Because economics has such a high share of foreign-born PhD recipients, this makes economics' lack of socioeconomic diversity even more surprising: the high share of foreign-born PhD recipients would be expected to push economics to be more socioeconomically diverse than other fields. But even among foreign-born PhDs, economics has a lower share of PhD recipients with no parents with a bachelor's degree than any major field, and the third highest share with a parent with a graduate degree, after psychology and the humanities (shown in Table 1). Again, this reflects overrepresentation both of people with parents with PhDs and of people with parents with non-PhD graduate degrees. Disaggregating to narrower PhD fields, economics has the 14th lowest share of people with no parent with a bachelor's degree and the 29th highest share of people with at least one parent with a graduate degree out of the 112 fields for which there were more than 500 foreign-born PhD recipients over 2010–2021.

Economics' lack of socioeconomic diversity relative to other PhD fields appears to be a common phenomenon across countries, albeit to different degrees. Of the top 20 foreign countries which contribute to producing US economics

---

[3]We list the 25 narrow fields with the lowest and highest shares of first-generation college graduates in online Appendix Table A6. The full list is available in the replication data.

PhDs, economics has a below-average share of first-generation college students as compared to other PhD fields in 19 of these countries. In fact, when comparing economics to US PhD recipients from each of these countries in the other 14 major PhD fields, economics has the lowest share with no parent with a bachelor's degree in six (India, Argentina, Mexico, Canada, Vietnam, France) and the second-lowest share in three more (China, Colombia, and Brazil).[4]

Fact #3: Economics is becoming relatively less socioeconomically diverse over time, as compared to other PhD fields and to the general US population.

In Figure 1, we show trends in parental education shares across PhD fields from 1970–2021 (including computer science only from 1980, due to the small number of PhD recipients before then). For all subjects, the share of PhDs with no parent with a bachelor's degree fell rapidly and the share with at least one parent with a graduate degree rose rapidly, reflecting rapid increases in educational attainment in the overall population. For both US-born and foreign-born PhDs, however, economics saw steeper changes than any other field. Panels A and B show that, in the early 1970s, US-born economics PhDs had similar levels of parental education to the other PhD fields, but over 1970–2021 the share of US-born economics PhDs with no parent with a bachelor's degree fell by nearly 50 percentage points, a substantially larger decline than in other disciplines. This shift came in two parts. First, from the 1980s until around 2000, economics, mathematics, and computer science diverged from other PhD fields. Next, from around 2000, economics diverged even from mathematics and computer science. Panels C and D illustrate a similar time pattern for foreign-born PhD recipients: again, economics became relatively less socioeconomically diverse over time as compared to other PhD fields.

Economics has also been becoming less representative of the general US population. Assuming that the parents of 2010–2021 PhD recipients were mostly between the ages of 50 and 74 in 2021, we can use the educational attainment of the 50–74 year-old US population in 2021 (estimated from the Current Population Survey) to proxy for the parental education of US residents who were the same age as our economics PhD recipients. Using this proxy, we can confirm that US-born economics PhDs are very unrepresentative of the similar-aged US population: only 14 percent had no parent with a bachelor's degree, compared to an estimated 66 percent of the similar-aged US-population. That is, recent US-born economics PhDs are only one-fifth as likely as an average similar-aged American to be from a family where no parent has a college degree. In 1970, in contrast, 57 percent of US-born economics PhDs had no parent with a bachelor's degree, compared to an estimated 92 percent of the similar-aged US population. We also estimate graduate degree prevalence for the similar-aged US population. Specifically, 65 percent of 2010–2021 US-born economics PhD recipients had at least

---

[4]For details, see online Appendix Figure A4.

Figure 1
Parental Education of US PhD Recipients over Time, 1970–2021

Panel A. US-born, no parent with bachelor's
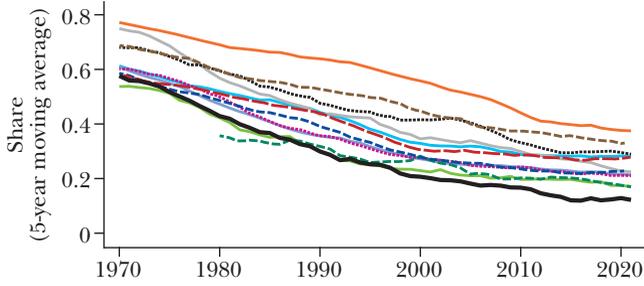
Panel B. US-born, parent with graduate degree
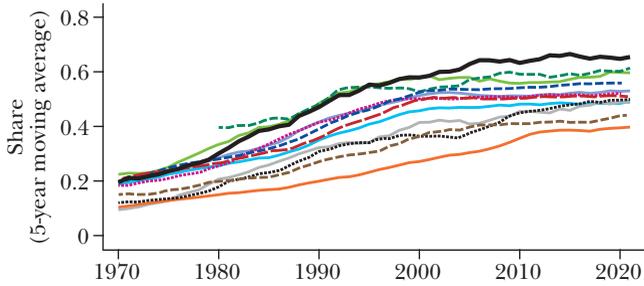
Panel C. Foreign-born, no parent with bachelor's
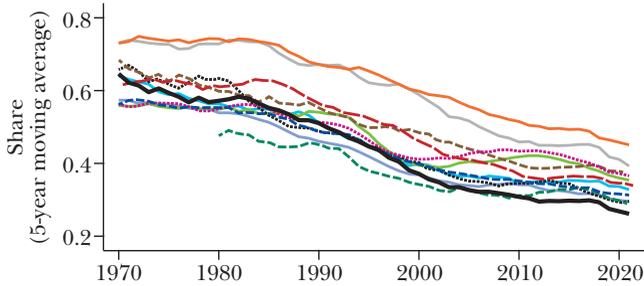
Panel D. Foreign-born, parent with graduate degree
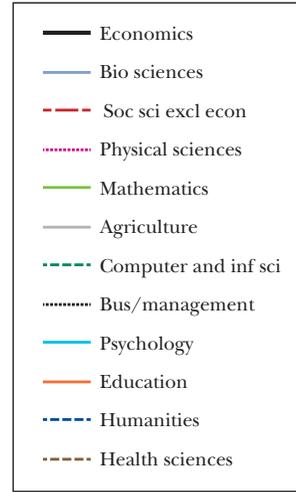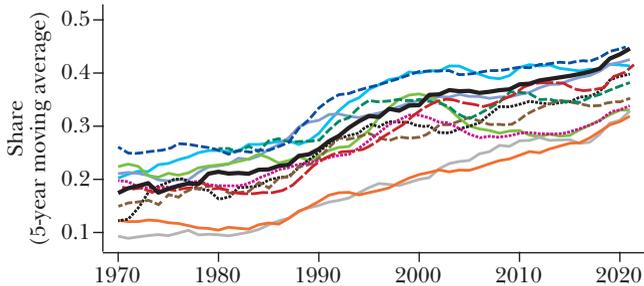
Source: Survey of Earned Doctorates.
Note: Five-year centered moving averages. Computer and information sciences are shown only from 1980 because of small sample sizes before then.

one parent with a graduate degree, compared to an estimated 12 percent of the similar-aged US population. In 1970, 20 percent of US-born economics PhDs had at least one parent with a graduate degree, compared to an estimated 2–3 percent of the similar-aged US population.[5]

Fact #4: The lack of socioeconomic diversity is particularly stark in higher-ranked economics PhD programs.

Half of all tenure-track economics professors in the United States got their PhDs at the 15 top-ranked economics PhD-granting departments, using the 2017 rankings from *US News and World Report* (Jones and Sloan 2020). By analyzing the socioeconomic makeup of PhD recipients from these top-ranked programs, we can therefore get some sense of the socioeconomic makeup of the tenure-track economics professoriate. In Figure 2, we show parental education shares for 2010–2021 PhD recipients from economics PhD programs ranked 1–6, 7–15, and 16 or lower according to the 2017 rankings. In the top six programs, 80 percent of US-born economics PhDs had at least one parent with a graduate degree, while only 5 percent had no parent with a bachelor's degree or higher.[6] The shares are similar for the programs ranked 7–15. US-born PhD recipients from programs ranked 16 and below are substantially more socioeconomically diverse than at the top-ranked programs: 58 percent had at least one parent with a graduate degree, and 17 percent had no parent with a bachelor's degree. Strikingly, however, students at economics PhD programs ranked 16 and below are still less socioeconomically diverse than US-born PhD recipients in any other major field (across all ranks of schools). And while the levels of socioeconomic diversity are higher among foreign-born PhDs across the board, there is still a strong gradient across institution rank: 17 percent in the top six PhD programs had no parent with a bachelor's degree, compared with 33 percent in the programs ranked 16 and below.

---

[5] Note that our method of estimating parental education for the similar-aged US population suffers from two biases which push in opposite directions. First, we implicitly assume perfect assortative matching on education, and less than perfect assortative matching would push the share with no parent with a bachelor's degree lower. Second, we implicitly assume equal numbers of children at a given age for people of different education levels, and higher fertility for those with lower education levels would push the estimated share with no parent with a bachelor's degree higher. An alternate way to estimate a comparator for the similar-aged population is to estimate the share of 2–10 year-old children in 1990 who had no parent in their household with a bachelor's degree, using the US Census. This share was 74 percent (as compared to our estimated share of 69 percent using the education levels of 50–74 year-olds). We use 2–10 year-olds in this calculation because the median age of US-born PhD recipients in 2010–2021 was 31.5, and 31 year-olds in 2010–2021 would have been aged around 2–10 in 1990.

[6] Bolotnyy, Basilico, and Barreira (2022) report similar figures in their study of mental health among economics PhD students at eight highly-ranked departments.

Parental Education of US Economics PhD Recipients by Rank of Program, 2010–2021

Parent: No bachelor's (US-born)

Rank: 1 to 6

Rank: 7 to 15

Rank: 16+

Share with no parent with bachelor's degree, US-born economics PhD graduates 2010–2021

Parent: No bachelor's (foreign-born)

Rank: 1 to 6

Rank: 7 to 15

Rank: 16+

Share with no parent with bachelor's degree, foreign-born economics PhD graduates 2010–2021

Parent: Graduate degree (US-born)

Rank: 1 to 6

Rank: 7 to 15

Rank: 16+

Share with parent with graduate degree, US-born economics PhD graduates 2010–2021

Parent: Graduate degree (foreign-born)

Rank: 1 to 6

Rank: 7 to 15

Rank: 16+

Share with parent with graduate degree, foreign-born economics PhD graduates 2010–2021

Source: Survey of Earned Doctorates.
Note: Rank of economics department from *US News and World Report*, 2017.

Race/Ethnicity, Gender, and Socioeconomic Diversity

In this section, we analyze the relationship between socioeconomic, racial/ethnic, and gender diversity. When studying racial and ethnic diversity in US PhD programs, it is common to focus only on US-born or US citizens and permanent residents (in this journal, for example, Bayer, Hoover, and Washington 2020), because the racial and ethnic makeup of different countries varies as well as the degree to which different racial and ethnic groups are underrepresented. We follow that approach here.

The analysis in the previous section shows that economics is one of the least socioeconomically diverse fields. From prior research, we know that economics also has among the lowest shares of women and of underrepresented racial or ethnic minority PhD recipients.[7] Indeed, this correlation exists across fields:

---

[7]We define an "underrepresented minority" as anyone who reports their ethnicity as Hispanic and/or who reports their race as Black or African American, American Indian or Alaska Native, or Native

among US-born students, PhD fields with lower shares of first-generation college students tend also to have lower shares of underrepresented minority students and lower shares of women, as illustrated in Figure 3. The correlation also holds at a more disaggregated level. Using narrow PhD fields (with more than 500 US-born PhD recipients 2010–2021), regressions of the share who had no parent with a bachelor's degree on the share who were underrepresented minority or female have an R-squared of 54 percent and 32 percent respectively. Economics' progress relative to other fields slowed on all three fronts—first-generation in college share, underrepresented minority share, and female share—at a similar time, from the 1990s to 2000s onwards.[8]

Economics' lack of socioeconomic diversity is, however, a separate axis that is not fully explained by (and does not fully explain) the field's lack of racial and ethnic diversity. The majority of first-generation college-graduate PhD recipients in economics are not underrepresented racial or ethnic minorities: 82 percent of all US-born first-generation economics PhD recipients 2010–2021 were White non-Hispanic. Similarly, the majority of underrepresented minorities receiving an economics PhD are not first-generation college graduates: more than half of Hispanic and of Black non-Hispanic US-born economics PhD recipients 2010–2021 had at least one parent with a graduate degree.

In fact, within every major racial or ethnic group, economics PhD graduates are disproportionately from families with high levels of formal education relative to other PhD disciplines. For PhD recipients who are White non-Hispanic, Black non-Hispanic, other-non-Hispanic, or Hispanic (all races), economics has the lowest share of PhD recipients who are first-generation college graduates, compared to the 14 major fields, and has the highest share with at least one parent with a graduate degree.[9] For Asian non-Hispanic PhD recipients, economics has the third-lowest share of first generation college graduates and the third-highest share of PhD recipients with at least one parent with a graduate degree (after computer/information

---

Hawaiian or other Pacific Islander (following NCSES 2021). The two major racial groups not included in this category are White non-Hispanic and Asian non-Hispanic, as well as those who report two or more races. Asian non-Hispanic is not considered an underrepresented minority because, while Asian Americans are a racial minority in the US population overall, students who self-report their race as Asian are not underrepresented in economics relative to the US population (for example Bayer and Wilcox 2019). Also, (1) the level of aggregation of our data does not allow us to capture disparities in access and inclusion within the Asian-American population, and (2) even if a group is not underrepresented, its members may be treated inequitably or may not be fully included because of their racial or ethnic identity. In the AEA Climate Survey (Allgood et al. 2019), 24 percent of Asian economists report being discriminated against or treated unfairly in the profession based on their race (the comparable figures were 47 percent for Black economists, 16 percent for Hispanic economists, and 4 percent for White economists).

[8]For foreign-born PhD recipients, there is no relationship across fields between female share and first-generation college student share—see online Appendix Figure A7. (We do not analyze underrepresented minority share for foreign-born PhD recipients given different underrepresentedness of different racial and ethnic groups across countries). For an illustration of how trends towards all three kinds of diversity slowed at about the same time, see online Appendix Figures A8–A11.

[9]For details, see online Appendix Figure A11.

Figure 3

First-Generation College Student Share, Female Share, and Underrepresented Minority Share, US-Born PhD Recipients, 2010–2021

Panel A. Female share and first-generation share
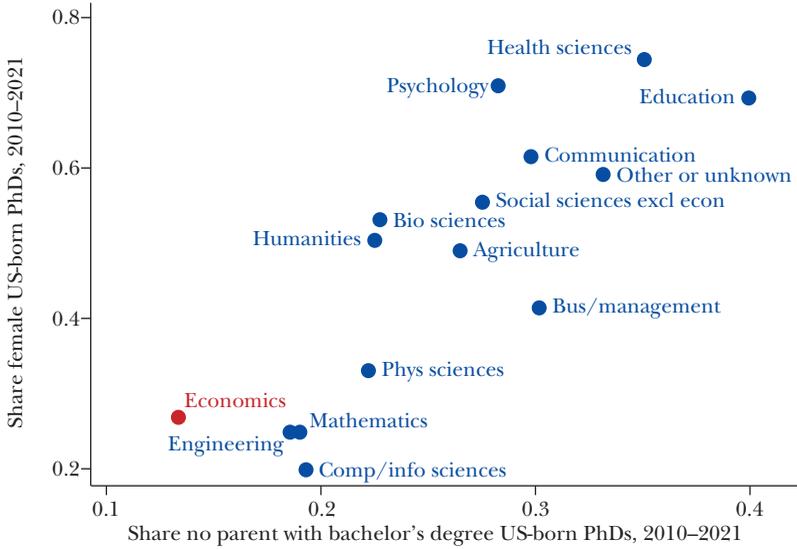
Panel B. Underrepresented minority share and first-generation share
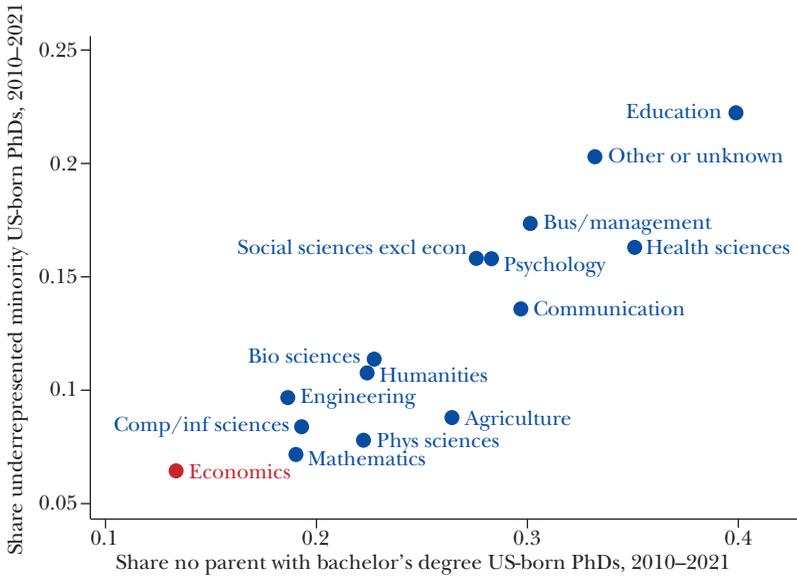
Source: Survey of Earned Doctorates.
Note: Underrepresented minorities include racial or ethnic minorities.

sciences and business/management). That is, economics' socioeconomic diversity problem exists within every major racial/ethnic group.

How do socioeconomic background, race/ethnicity, and gender interact? An intersectional understanding would suggest that students who are both racial/ethnic minorities *and* from socioeconomically less advantaged backgrounds face particularly high barriers. The data are consistent with this insight. While first-generation college students are underrepresented within every racial/ethnic group relative to the general US population, this underrepresentation is particularly stark for Black students: only 0.4 percent of US-born economics PhDs 2010–2021 were Black and had no parent with a bachelor's degree, compared to an estimated 9 percent of the similar-aged US population. Similarly, Hispanic students with no parent with a bachelor's degree made up 1 percent of US-born economics PhDs, compared to an estimated 10 percent of the similar-aged US population.[10]

For students who are advantaged on some of these axes and less advantaged on others, our data illustrate the importance of using the lens of socioeconomic background alongside gender and race to understand access and opportunity. White non-Hispanic people are overrepresented in economics, but White non-Hispanic first-generation college students are heavily underrepresented relative to the similar-aged US population. Similarly, men are overrepresented in economics, but men who are first-generation college students are underrepresented. Indeed, despite the fact that less than one-third of US-born economics PhDs are women, women with at least one parent with a graduate degree are still overrepresented among economics PhDs relative to the similar-aged US population.[11]

Possible Drivers: Brief Explorations

Why is economics less socioeconomically diverse than other PhD disciplines? In this section, we explore some empirical patterns that suggest possible reasons for the lack of socioeconomic diversity in economics.

The Role of Undergraduate Major and Institution

One possibility is that economics' diversity problem at the PhD level arises from a lack of socioeconomic diversity in the undergraduate economics major, which remains by far the largest major for those entering an economics PhD program. Our analysis of data from the 2016 Baccalaureate and Beyond survey, a nationally representative study of bachelor's degree recipients from the National Center for Education Statistics, suggests that economics is less socioeconomically diverse than

---

[10]The share of the similar-aged US population who is Black (Hispanic) and had no parent with a bachelor's degree is estimated by using the share of 50–74 year-old US adults in 2021 who were Black (Hispanic) and had no bachelor's degree, using Census Bureau data. See online Appendix Table A8 for more details.

[11] For details, see online Appendix Figure 12.

Table 2

Parental Education of US Population, US-Born Bachelor's Recipients, and US-Born PhD Recipients

	Share, no parent with a bachelor's degree	*Share, at least one parent with a graduate degree*	*Source*
Economics PhDs (2010–2021, US-born)	13%	65%	Survey of Earned Doctorates, 2010–2021 PhD recipients
Math and social science PhDs (2010–2021, US-born)	24%	54%	
All PhDs (2010–2021, US-born)	26%	50%	
Economics BAs (US-born)	21%	41%	Baccalaureate and beyond study, 2016/2017 BA recipients
Math and social science BAs (US-born)	35%	34%	
All BAs (US-born)	42%	29%	
Similar-aged US population (estimated)	66%	12%	Current population survey

Note: Figures for math and social science include economics. Bachelor's degree is abbreviated to "BA" in labels for brevity; data include any bachelor's degree (including BSc). Data for PhDs are from the Survey of Earned Doctorates, 2010–2021 US-born PhD recipients. Estimates for bachelors' degree recipients ("BAs") are from the Baccalaureate and Beyond 2016/2017 survey. Similar-age US population parental education shares are estimated using US Census Bureau data on education by age calculated from the Current Population Survey. Specifically, we use the share of the 50–74 year-old population in 2021 with no bachelor's degree, or with a graduate degree, to proxy for the parental education shares of the 27–37 year-old US population in 2010–2021.

other majors at the undergraduate level. In 2016, 21 percent of US-born economics bachelor's degree recipients had no parent with a bachelor's degree, compared to 35 percent for math and the social sciences, and 42 percent for the average bachelor's degree recipient, as shown in Table 2. Similarly, 41 percent of US-born economics bachelor's degree recipients had at least one parent with a graduate degree, compared to 34 percent for math and social sciences, and 29 percent for the average bachelor's degree recipient.

A related factor may be that economics PhD programs tend to select from undergraduate institutions with a lower degree of socioeconomic diversity. We group US-born PhD recipients into three groups: those who got their bachelor's degree from an "Ivy Plus" institution (defined as the eight Ivy League schools plus Stanford, MIT, Chicago, and Duke); those who got their bachelor's degree from a private non-Ivy-Plus; and those who got their bachelor's degree from a public institution.[12] Across all US-born PhD recipients over 2010–2021, 59 percent had a

---

[12]We have data on undergraduate institution for all PhDs who did their undergraduate degree in the United States. This represents 97 percent of US-born PhD recipients in our sample 2010–2021. We follow Chetty et al. (2017) in categorizing the twelve "Ivy Plus" schools. Chetty et al. (2017) show that

Figure 4
Undergraduate Institution Type by Field, US-Born PhD Recipients 2010–2021
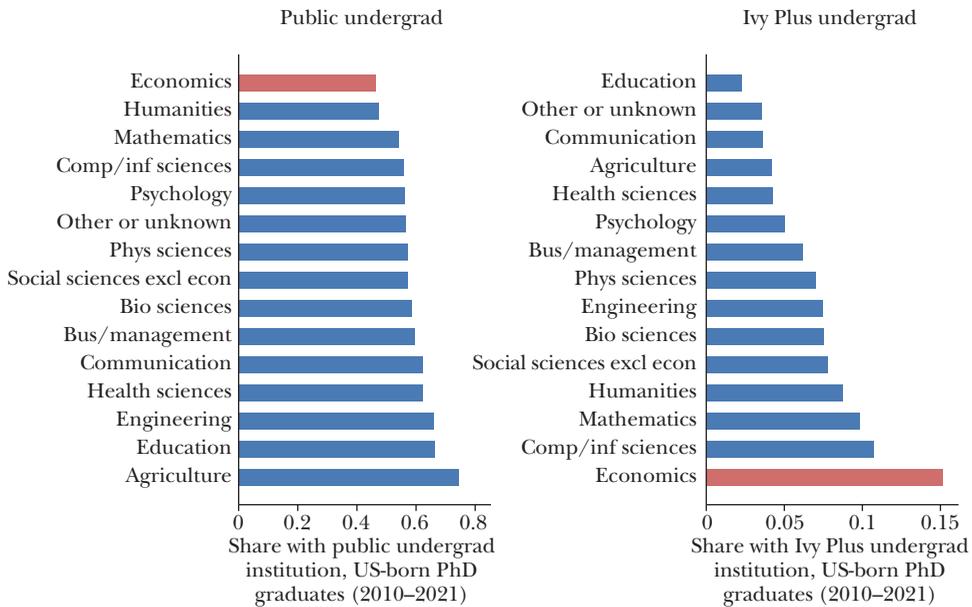
Source: Survey of Earned Doctorates.
Note: "Ivy Plus" institutions are the eight Ivy League schools—Brown, Columbia, Cornell, Dartmouth, Harvard, Princeton, the University of Pennsylvania, and Yale—as well as Duke, Chicago, MIT, and Stanford (following Chetty et al. 2017). Data is shown as a share of all US-born PhD recipients who got their bachelor's degree from a US institution.

bachelor's degree from a public institution and 7 percent from one of the twelve Ivy Plus institutions. Comparing economics to the 14 major PhD fields, we see that economics had the lowest share of PhD recipients who got their bachelor's degree from a public school, at 46 percent, and the highest share of PhD recipients who got their bachelor's degree from an Ivy Plus school, at 15 percent, as illustrated in Figure 4. Note also that the public undergraduate institutions most represented among economics PhDs are typically highly selective (like University of California-Berkeley, University of Wisconsin–Madison, and University of Michigan) (Siegfried and Stock 2007), and likely have a relatively socioeconomically advantaged population as compared with public college students as a whole.

In the top six economics PhD programs, the figures are particularly striking: among US-born students, nearly twice as many got their bachelor's degree from one of the twelve Ivy-Plus institutions (45 percent) as from any public institution

---

14.5 percent of parents of Ivy Plus undergraduates are in the top 1 percent of the US income distribution. This compares to 2.5 percent of parents of undergraduates at the top 26 "highly selective" public colleges as categorized by Barron's.

(24 percent).[13] The disproportionate representation of a small number of elite undergraduate schools among PhD recipients also exists among economics faculty. Jones and Sloan (2020) study the undergraduate institutions of tenure-track faculty at ranked US economics departments and find that 20 percent of the roughly half of economics faculty who got their bachelor's degree in the United States did so at an Ivy League school.

To explore the potential relevance of these mechanisms, we carried out some exploratory regressions.[14] On average, economics PhD recipients are 13 percentage points less likely than the average US-born PhD recipient to have no parent with a bachelor's degree or higher and 15 percentage points more likely to have a parent with a graduate degree. When we add fixed effects for undergraduate majors, these gaps are reduced by about one-third. When we also adjust for undergraduate institution attended, the two factors together reduce the baseline percentage by about half. This suggests that the relative lack of socioeconomic diversity among the undergraduate majors and institutions that economics PhD programs draw from can explain about half of its socioeconomic diversity gap (relative to other fields) at the PhD level. Adding race/ethnicity and gender controls into the regressions has a much smaller effect, suggesting that the lack of racial and ethnic diversity in economics is not the main driver of its lack of socioeconomic diversity. This is unsurprising, given that our analysis in the previous section shows that economics is the least socioeconomically diverse PhD field *within* most major racial and ethnic groups.

Because we found a large role for undergraduate major and undergraduate institution, in the rest of this section we focus our discussion of potential drivers of the socioeconomic diversity problem in economics into two portions: (1) the lack of socioeconomic diversity at the undergraduate level in the majors and institutions which feed into economics PhDs, and (2) the lack of socioeconomic diversity as it appears in the pipeline from undergraduate education to PhD completion. Informed by other research, we discuss some possibilities below. This discussion is necessarily speculative and there are likely additional factors at play.

Why Might the Economics Major Lack Socioeconomic Diversity?

The economics major lacks socioeconomic diversity relative to other majors in part because economics is a larger major at schools with more socioeconomically advantaged student bodies. Data from the US Department of Education Integrated Postsecondary Education Data System (IPEDS) database illustrates, for example, that 3.3 percent of bachelor's graduates at private R1 universities in 2018/2019 were economics majors, as were 5.8 percent of graduates at private liberal arts colleges,

---

[13] For details, see online Appendix Figure A14.
[14] We regress an indicator for whether someone is a first-generation college student (or has a parent with a graduate degree) on an indicator for their PhD field being economics. We then progressively add fixed effects and note how the coefficient on economics changes. More details are available in online Appendix Table A9.

relative to 2.2 percent of bachelor's graduates at public R1 universities.[15] The R1 universities are the 146 or so most research-intensive institutions in the United States.

But even within schools, data suggests that first-generation college students are less likely to major in economics than other students. Using the Baccalaureate and Beyond data, we regress a binary indicator for first-generation status on whether or not someone is an economics major, controlling for undergraduate institution fixed effects: this regression tells us that within a given undergraduate institution, the average economics major is ten percentage points less likely to be a first-generation college student, as compared to a randomly drawn student from the same university.[16] Similarly, Hammock, Routon, and Walker (2016) found that economics majors at 463 US colleges from 1994 to 1999 were more likely than average to come from more educated and more affluent homes, and Bleemer and Mehta (2022) report that economics majors at the University of California, Santa Cruz (2008–2012) came from ZIP codes with mean incomes 8 percent higher than the average freshman's home ZIP code.

Why is economics a less socioeconomically diverse undergraduate major than average even within individual universities?

A first factor may be *access to the major*. Bleemer and Mehta (2021) show that grade-point-average cutoffs for economics majors (common at large public universities) disproportionately decrease access to the economics major for students with lower socioeconomic status, as a result of lesser academic preparation and opportunity before college.

Second, if less socioeconomically advantaged students are less likely to have had exposure to the study of economics at high school, or from their families or communities, there may be a *lack of information about the economics major*. Indeed, two field experiments suggest that informational messages to first years about the economics major can substantially increase uptake of economics courses and majors by first-generation college students (Pugatch and Schroeder 2021; Bayer, Bhanot, and Lozano 2019).

Third, *the way economics is taught at undergraduate level*—and in introductory economics courses in particular—may dissuade students from less advantaged socioeconomic backgrounds who see unrealistic or limited portrayals of topics that are of particular importance to them or in which they have more first-hand experience. Some attempts at curricular reform to broaden introductory economics classes and connect them more closely to the real world have increased the number of first-generation college students in undergraduate economics courses

---

[15] For details, see online Appendix Table A10.

[16] Specifically, this reports the results from a regression of a binary indicator for first-generation (no parent with a college degree) on a binary indicator for being an Economics major, with bachelor's institution fixed effects and standard errors clustered by bachelor's institution, using the 2008/2018 Baccalaureate and Beyond survey data for 2008 bachelor's degree recipients. The regression sample is restricted to US-born undergraduates only. The total sample is 14,310 individuals.

(Bayer et al. 2020; Owen and Hagstrom 2021).[17] In addition, the language commonly used in economics—and the values implicit in this language, whether intentionally or unintentionally—may be off-putting to students from less advantaged socioeconomic backgrounds. Such terms include "unskilled," "low type," "low skill," or "low ability," commonly used to refer to people with little formal education.

The Pipeline to the Economics PhD

PhD recipients in all disciplines are, on average, substantially more socioeconomically advantaged than bachelor's degree recipients. But this disparity is particularly stark in economics. This points to a role for factors which reduce socioeconomic diversity in the pipeline from undergraduate to PhD.

One possibility is *the path to a successful economics PhD application (or completion) may be more obscure or inaccessible* than in other disciplines. Undergraduate students from socioeconomically advantaged backgrounds are more aware of and interested in the option of doing a PhD post-college and/or are likely to be more aware of the requirements to get accepted into a PhD program; in contrast, prospective students from lower socioeconomic-status backgrounds are less likely to have informational resources about options in college or options for graduate school (Gardner and Holley 2011; Posselt and Black 2012; Brown, Wohn, and Ellison 2016). For example, a successful economics PhD application requires specific math classes or even a math major (Jones et al. 2020). As Bayer et al. (2020, p. 198) note, following Sharpe (2017): "[I]t is not intuitive to undergraduates that an economics major is not sufficient preparation for a doctoral economics program." Students from less socioeconomically advantaged backgrounds may be less aware of these somewhat unintuitive requirements until late in their undergraduate degree. Jeitschko (2019) finds evidence that few undergraduates understand what obtaining a PhD in economics entails, and that this information asymmetry is more pronounced for women, underrepresented minorities, and first-generation college students. The importance of following a narrowly defined path may be particularly high in economics relative to other disciplines: in a study of PhD admissions at elite universities, Posselt (2016) argues that economics admission committees are particularly risk-averse, making admission more difficult for applicants with less traditional or prestigious trajectories. The recent rise of pre-doctoral research assistantships as a path to an economics PhD (as illustrated in Bryan 2019) may have the potential to help less advantaged students catch up on missed prerequisites during their undergraduate degree, but this depends to a large degree on what selection process of these pre-doctoral programs. A recent survey of US pre-doctoral students finds that

[17] To the extent that academics in economics tend to hold relatively more conservative views than academics in other majors (Gross and Simmons 2007), and to the extent lower-income students are more likely to hold less conservative views, this may also dissuade people from less advantaged socioeconomic backgrounds from studying economics. Bartlett, Ferber, and Green (2009), for example, found that undergraduates with a conservative political orientation are more likely to major in economics than in other fields.

the large majority have a parent with a graduate degree and few are first-generation college graduates (Huang, Liang, and Russel n.d.).

A second possibility is *disparate access to professional relationships.* If the path to an economics PhD relies more heavily on access to specific relationships or networks as compared with other PhD disciplines, this may explain some of the field's socioeconomic disparities. Access to mentoring and opportunities often relies on students initiating relationships with faculty (going to office hours or asking for opportunities outright), but students with limited family experience in higher education are often unaware that this is an option or expectation, or may be less comfortable in interactions with faculty; for example, Jack (2016, p. 1) notes that at an elite university, low-income undergraduates who attended local high schools tend to be "more resistant to engaging authority figures in college and tend to withdraw from them" (see also Smith, Mao, and Deshpande 2016; Yee 2016). Moreover, effective mentoring relationships are more likely to form along demographic lines (for example, Blake-Beard, Murrell, and Thomas 2006), making the existing demographic makeup of the profession somewhat self-perpetuating in the absence of intentional corrective action. This may be exacerbated by implicit or explicit bias toward students from disadvantaged socioeconomic backgrounds. While socioeconomic background may be less immediately detectable than race or gender, individuals from advantaged socioeconomic backgrounds may have more experience comporting themselves in ways that are considered professionally advantageous or impressive (for a discussion in the context of elite occupations in the United Kingdom, see Friedman and Laurison 2020). Indeed, research suggests that individuals are able to detect socioeconomic status from people's voices (Kraus, Park, and Tan 2017; Kraus et al. 2019) and facial cues (Bjornsdottir and Rule 2017) among other characteristics (like dress, behavior, or name), and a recent correspondence study suggests that individuals with signals on their resume indicating socioeconomic advantage are more likely to receive callbacks for graduate jobs in elite US professional service firms (Rivera and Tilcsik 2016).

Third, *financial circumstances and incentives* might cause strong economics students from less advantaged socioeconomic backgrounds to be less likely to pursue a career in academia. After all, academia offers lower pay and less financial stability in its early years, compared with private sector jobs that students with good undergraduate qualifications can access (for example, Millett 2003; Hoffer et al. 2003; Walpole 2003), and perhaps particularly so for economics-related fields. Thus, the opportunity cost of a PhD relative to a career in the private sector—for students who might be interested in and qualified for both—may be greater in economics than in other disciplines. On the other hand, the career prospects after an economics PhD are better than for most other similar PhD disciplines in terms of both expected salary and expected job security (Freeman 1999; Fourcade, Ollion, and Algan 2015). This would tend to push in the opposite direction—if students are aware of these benefits.

Fourth, *the culture of the economics profession* may be unwelcoming for those from less advantaged socioeconomic backgrounds. The factors highlighted in our

discussion about diversity in the undergraduate major—including language used in economics which is likely disproportionately alienating to people from less advantaged backgrounds—may apply also in the pipeline from undergraduate to PhD. To the extent that people tend to be excluded from and feel alienated in groups where they are highly underrepresented (for example, Rubin 2012), the unrepresentativeness of the economics profession itself could create a climate that dissuades and discourages people from less advantaged socioeconomic backgrounds.

Finally, it is interesting to note that the fields that are particularly nondiverse in terms of gender, race, and socioeconomic background among US-born students—economics, mathematics, and computer science—are also those that have a larger share of PhDs who are not US-born.[18] This is consistent with the hypothesis that PhD recruitment from a larger international pool of students may increase competition among domestic students, which may have the unintended consequence of reducing access for students with less competitive profiles—who may, because of lack of access, information, and/or opportunities, be disproportionately from underrepresented groups.

Conclusion

We have documented the fact that recipients of US PhDs in economics are from a narrower—and more privileged—range of socioeconomic backgrounds than US PhDs from other disciplines. The underrepresentation in economics programs of individuals whose parents do not have a college education reflects underrepresentation of a very broad swathe of the population: people whose parents do not have a college degree spans those raised in poverty to those raised in well-off middle-class households (headed for example by non-college-educated businesspeople, tradespeople, or health professionals, to name a few occupations). Accordingly, the lived experiences—and associated insights—that are missing from economics likely span this entire range. Moreover, a substantial share of those who have the talent and interest to become a PhD economist are not finding a way into the profession.

When considering why economics draws students disproportionately from socioeconomically advantaged backgrounds, there are three useful patterns to distinguish. First, the population of *PhDs in all subjects* is substantially more socioeconomically advantaged than the population of bachelor's degree recipients (who in turn are more socioeconomically advantaged than the population as a whole). Second, US-born students in a *subset of quantitative PhD disciplines*—economics, mathematics, and computer science—are more socioeconomically advantaged than

---

[18] For details, see online Appendix Figure A15. This insight also holds true when looking at narrow PhD fields: a regression of the share of US-born PhD recipients who have no parent with a bachelor's degree on the share of all PhD recipients who are foreign born, for the 143 narrow fields with more than 1,000 PhD recipients 2010–2021, has a coefficient of 0.2 and an R-squared of 24 percent.

average across PhD disciplines. And third, US-born *economics PhD recipients* have since 2000 been even more socioeconomically advantaged than those in mathematics and computer science. As the profession increases its research into, understanding of, and action to tackle its lack of socioeconomic diversity, all three of these factors should be borne in mind. Our empirical explorations in this paper suggest that factors at the undergraduate level (including the decision to become an economics major) as well as factors in the pipeline from undergraduate to PhD play a role.

We have focused on recent graduates of economics PhD programs, but the lack of socioeconomic diversity is likely to be even steeper among the existing population of US economics professors. We can see this from our analysis of PhDs from the top 15 US economics PhD programs, who make up 50 percent of the tenure-track economics professoriate (according to Jones and Sloan 2020). Over 2010–2021, 79 percent of US-born economics PhDs in the top 15 programs had least one parent with a graduate degree. Only 6 percent were first-generation college graduates. A larger share did their undergraduate degree at one of the twelve "Ivy Plus" institutions than at any public institution (35 percent, compared with 27 percent). Professors in a range of other fields also come from much more socioeconomically advantaged backgrounds than PhD recipients (Morgan et al. 2022).

The poor performance of the profession of economics in expanding gender and racial diversity has attracted much-needed attention over recent years (in this journal, for example, see Buckles 2019; Boustan and Langan 2019; Lundberg and Stearns 2019; Bayer, Hoover, and Washington 2020). With our documentation and discussion of the lack of socioeconomic diversity in economics, we hope to spur similar, complementary efforts, toward documenting diversity of socioeconomic background in the economics profession, analyzing the causes (and consequences) of this unrepresentativeness, and developing solutions.

References

Allgood, Sam, Lee Badgett, Amanda Bayer, Marianne Bertrand, Sandra E. Black, Nick Bloom, and Lisa D. Cook. 2019. *AEA Professional Climate Survey: Final Report.* Nashville: American Economic Association.

Bartlett, Robin L., Marianne A. Ferber, and Carole A. Green. 2009. "Political Orientation and the Decision to Major in Economics: Some Preliminary Observations." *International Review of Economics Education* 8 (1): 13–31.

Bayer, Amanda, Syon P. Bhanot, and Fernando Lozano. 2019. "Does Simple Information Provision Lead to More Diverse Classrooms? Evidence from a Field Experiment on Undergraduate Economics." *AEA Papers and Proceedings* 109: 110–14.

Bayer, Amanda, Gregory Bruich, Raj Chetty, and Andrew Housiaux. 2020. "Expanding and Diversifying the Pool of Undergraduates Who Study Economics: Insights from a New Introductory Course at Harvard." *Journal of Economic Education* 51 (3–4): 364–79.

Bayer, Amanda, Gary A. Hoover, and Ebonya Washington. 2020. "How You Can Work to Increase the Presence and Improve the Experience of Black, Latinx, and Native American People in the Economics Profession." *Journal of Economic Perspectives* 34 (3): 193–219.

Bayer, Amanda, and Cecilia Elena Rouse. 2016. "Diversity in the Economics Profession: A New Attack on an Old Problem." *Journal of Economic Perspectives* 30 (4): 221–42.

Bayer, Amanda, and David W. Wilcox. 2019. "The Unequal Distribution of Economic Education: A Report on the Race, Ethnicity, and Gender of Economics Majors at U.S. Colleges and Universities." *Journal of Economic Education* 50 (3): 299–320.

Bjornsdottir, R. Thora, and Nicholas O. Rule. 2017. "The Visibility of Social Class from Facial Cues." *Journal of Personality and Social Psychology* 113 (4): 530–46.

Blake-Beard, Stacy, Audrey J. Murrell, and David A. Thomas. 2006. "Unfinished Business: The Impact of Race on Understanding Mentoring Relationships." Harvard Business School Working Paper 06-060.

Bleemer, Zachary, and Aashish Mehta. 2021. "College Major Restrictions and Student Stratification." Unpublished.

Bleemer, Zachary, and Aashish Mehta. 2022. "Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major." *American Economic Journal: Applied Economics* 14 (2): 1–22.

Bolotnyy, Valentin, Matthew Basilico, and Paul Barreira. 2022. "Graduate Student Mental Health: Lessons from American Economics Departments." *Journal of Economic Literature* 60 (4): 1188–1222.

Boustan, Leah, and Andrew Langan. 2019. "Variation in Women's Success across PhD Programs in Economics." *Journal of Economic Perspectives* 33 (1): 23–42.

Brown, Michael Geoffrey, Donghee Y. Wohn, and Nicole Ellison. 2016. "Without a Map: College Access and the Online Practices of Youth from Low-Income Communities." *Computers and Education* 92–93: 104–16.

Bryan, Kevin A. 2019. "Young 'Stars' in Economics: What They Do and Where They Go." *Economic Inquiry* 57 (3): 1392–1407.

Buckles, Kasey. 2019. "Fixing the Leaky Pipeline: Strategies for Making Economics Work for Women at Every Stage." *Journal of Economic Perspectives* 33 (1): 43–60.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." NBER Working Paper 23618.

Dal Bó, Ernesto, Pedro Dal Bó, and Jason Snyder. 2009. "Political Dynasties." *Review of Economic Studies* 76 (1): 115–42.

Daly, Mary C. 2018. "Getting from Diversity to Inclusion in Economics." Federal Reserve Bank of San Francisco (FRBSF) Economic Letter 2018-15.

Duncan, Otis D., David L. Featherman, and Beverly Duncan. 1972. "Socioeconomic Background and Achievement." New York: Seminar Press.

Ethington, Corinna A., and John C. Smart. 1986. "Persistence to Graduate Education." *Research in Higher Education* 24 (3): 287–303.

Fourcade, Marion, Etienne Ollion, and Yann Algan. 2015. "The Superiority of Economists." *Journal of Economic Perspectives* 29 (1): 89–114.

Freeman, Richard B. 1999. "It's Better Being an Economist (But Don't Tell Anyone)." *Journal of Economic Perspectives* 13 (3): 139–45.

Friedman, Sam, and Daniel Laurison. 2020. "The Class Ceiling: Why It Pays to Be Privileged." *Social Forces* 99 (1): e5.

Gardner, Susan K., and Karri A. Holley. 2011. "'Those Invisible Barriers Are Real': The Progression of First-Generation Students through Doctoral Education." *Equity and Excellence in Education* 44 (1): 77–92.

Ginther, Donna K., and Shulamit Kahn. 2021. "Women in Academic Economics: Have We Made Progress?" *AEA Papers and Proceedings* 111: 138–42.

Gross, Neil, and Solon Simmons. 2007. "The Social and Political Views of American Professors." Unpublished.

Hammock, Michael R., P. Wesley Routon, and Jay K. Walker. 2016. "The Opinions of Economics Majors before and after Learning Economics." *Journal of Economic Education* 47 (1): 76–83.

Hauser, Robert M. 1994. "Measuring Socioeconomic Status in Studies of Child Development." *Child Development* 65 (6): 1541–45.

Hoffer, Thomas B., Scott Sederstrom, Lance Selfa, Vince Welch, Mary Hess, Shana Brown, Sergio Reyes, Kristy Webber, and Isabel Guzman-Barron. 2003. *Doctorate Recipients from United States Universities: Summary Report 2002.* Chicago, IL: National Opinion Research Center at University of Chicago.

Huang, Zong, Pauline Liang, and Dominic Russel. "Survey of Pre-doctoral Research Experiences in Economics." https://z-y-huang.github.io/predoc_survey/slides.pdf.

Jack, Anthony Abraham. 2016. "(No) Harm in Asking: Class, Acquired Cultural Capital, and Academic Engagement at an Elite University." *Sociology of Education* 89 (1): 1–19.

Jaeger, David A. 1997. "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers." *Journal of Business and Economic Statistics* 15 (3): 300–309.

Jeitschko, Thomas D. 2019. "Toward the Next Generation of Scholarship: Challenges and Opportunities for Full Participation in PhD Training in Economics." *AEA Papers and Proceedings* 109: 250–54.

Jones, Adam, Peter Schuhmann, Daniel Soques, and Allison Witman. 2020. "So You Want to Go to Graduate School? Factors that Influence Admissions to Economics PhD Programs." *Journal of Economic Education* 51 (2): 177–90.

Jones, Todd R., and Arielle Sloan. 2020. "Staying at the Top: The Ph.D. Origins of Economics Faculty." Annenberg Institute at Brown University EdWorkingPaper 20-324.

Jonsson, Jan O., David B. Grusky, Matthew Di Carlo, Reinhard Pollak, and Mary C. Brinton. 2009. "Microclass Mobility: Social Reproduction in Four Countries." *American Journal of Sociology* 114 (4): 977–1036.

Kniffin, Kevin. 2007. "Accessibility to the PhD and Professoriate for First-Generation College Graduates: Review and Implications for Students, Faculty, and Campus Policies." *American Academic* 3: 49–79.

Kraus, Michael W., Jun Won Park, and Jacinth J. X. Tan. 2017. "Signs of Social Class: The Experience of Economic Inequality in Everyday Life." *Perspectives on Psychological Science* 12 (3): 422–35.

Kraus, Michael W., Brittany Torrez, Jun Won Park, and Fariba Ghayebi. 2019. "Evidence for the Reproduction of Social Class in Brief Speech." *Proceedings of the National Academy of Sciences* 116 (46): 22998–23003.

Laurison, Daniel, and Sam Friedman. 2020. "The Class Ceiling in the United States: Class-Origin Pay Gaps in Higher Professional and Managerial Occupations." Unpublished.

Lee, Elizabeth M. 2015. "'Where People Like Me Don't Belong': Faculty Members from Low-Socioeconomic-Status Backgrounds." *Sociology of Education* 90 (3): 197–212.

Lundberg, Shelly, ed. 2020. *Women in Economics.* London: CEPR Press.

Lundberg, Shelly, and Jenna Stearns. 2019. "Women in Economics: Stalled Progress." *Journal of Economic Perspectives* 33 (1): 3–22.

May, Ann Mari, Mary G. McGarvey, and Robert Whaples. 2014. "Are Disagreements among Male and Female Economists Marginal at Best? A Survey of AEA Members and Their Views on Economics and Economic Policy." *Contemporary Economic Policy* 32 (1): 111–32.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh. 2015. "What Happens Before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations." *Journal of Applied Psychology* 100 (6): 1678–1712.

Millett, Catherine M. 2003. "How Undergraduate Loan Debt Affects Application and Enrollment in Graduate or First Professional School." *Journal of Higher Education* 74 (4): 386–427.

Morgan, Allison C., Nicholas LaBerge, Daniel B. Larremore, Mirta Galesic, Jennie E. Brand, and Aaron Clauset. 2022. "Socioeconomic Roots of Academic Faculty." *Nature Human Behaviour* 6 (12): 1625–33.

Mullen, Ann L., Kimberly A. Goyette, and Joseph A. Soares. 2003. "Who Goes to Graduate School? Social and Academic Correlates of Educational Continuation after College." *Sociology of Education* 76 (2): 143–69.

National Center for Science and Engineering Statistics (NCSES). 2020. *2019 Doctorate Recipients from U.S. Universities.* NSF 21-308. Alexandria, VA: National Science Foundation.

National Center for Science and Engineering Statistics (NCSES). 2021. *Women, Minorities, and Persons with Disabilities in Science and Engineering.* NSF 21-321. Alexandria, VA: National Science Foundation.

Oldfield, Kenneth. 2007. "Expanding Economic Democracy in American Higher Education: A Two-Step Approach to Hiring More Teachers from Poverty- and Working-Class Backgrounds." *Journal of Higher Education Policy and Management* 29 (2): 217–30.

Owen, Ann L., and Paul Hagstrom. 2021. "Broadening Perceptions of Economics in a New Introductory Economics Sequence." *Journal of Economic Education* 52 (3): 175–91.

Posselt, Julie R. 2016. *Inside Graduate Admissions: Merit, Diversity, and Faculty Gatekeeping.* Cambridge, MA: Harvard University Press.

Posselt, Julie Renee, and Kim R. Black. 2012. "Developing the Research Identities and Aspirations of First-Generation College Students: Evidence from the McNair Scholars Program." *International Journal for Researcher Development* 3 (1): 26–48.

Pugatch, Todd, and Elizabeth Schroeder. 2021. "A Simple Nudge Increases Socioeconomic Diversity in Undergraduate Economics." IZA Discussion Paper 14418.

Rivera, Lauren A., and András Tilcsik. 2016. "Class Advantage, Commitment Penalty: The Gendered Effect of Social Class Signals in an Elite Labor Market." *American Sociological Review* 81 (6): 1097–1131.

Rubin, Mark. 2012. "Social Class Differences in Social Integration among Students in Higher Education: A Meta-analysis and Recommendations for Future Research." *Journal of Diversity in Higher Education* 5 (1): 22–38.

Schultz, Robert, and Anna Stansbury. 2022. "Socioeconomic Diversity of Economics PhDs." Peterson Institute for International Economics Working Paper 22-4.

Sharpe, Rhonda Vonshay. 2017. "We've Built the Pipeline: What's the Problem and What's Next?" *CSWEP News* 2: 10–11.

Siegfried, John J., and Wendy A. Stock. 2007. "The Undergraduate Origins of PhD Economists." *Journal of Economic Education* 38 (4): 461–82.

Sirin, Selcuk R. 2005. "Socioeconomic Status and Academic Achievement: A Meta-analytic Review of Research." *Review of Educational Research* 75 (3): 417–53.

Smith, Laura, Susan Mao, and Anita Deshpande. 2016. "'Talking across Worlds': Classist Microaggressions and Higher Education." *Journal of Poverty* 20 (2): 127–51.

Stansbury, Anna, and Robert Schultz. 2023. "Replication data for: The Economics Profession's Socioeconomic Diversity Problem." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193683V1.

Stock, Wendy A., and John J. Siegfried. 2015. "The Undergraduate Origins of PhD Economists Revisited." *Journal of Economic Education* 46 (2): 150–65.

Strayhorn, Terrell L. 2010. "When Race and Gender Collide: Social and Cultural Capital's Influence on the Academic Achievement of African American and Latino Males." *Review of Higher Education* 33 (3): 307–32.

Urban Institute. IPEDS Directory, Education Data Portal (Version 0.18.0). https://educationdata.urban.org/documentation/, made available under the ODC Attribution License (accessed August, 08, 2023).

US Census Bureau, "Educational Attainment in the United States 2021", available at https://www.census.gov/data/tables/2021/demo/educational-attainment/cps-detailed-tables.html (accessed August 17, 2023).

US Department of Education. National Center for Education Statistics, Baccalaureate and Beyond Longitudinal Study, 08/18 and 16/17 waves. Restricted access dataset, available from National Center for Education Statistics.

US Department of Education. National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2018–19, Summary Tables: Degrees/certificates by race/ethnicity, and field of study, 2018–19. Available at https://nces.ed.gov/ipeds/SummaryTables (accessed August 8, 2023).

Walpole, MaryBeth. 2003. "Socioeconomic Status and College: How SES Affects College Experiences and Outcomes." *Review of Higher Education* 27 (1): 45–73.

Wessel, David, Louise Sheiner, and Michael Ng. 2019. "Gender and Racial Diversity of Federal Government Economists." Washington, DC: Brookings Institution.

Wu, Alice H. 2020. "Gender Bias among Professionals: An Identity-Based Interpretation." *Review of Economics and Statistics* 102 (5): 867–80.

Yee, April. 2016. "The Unwritten Rules of Engagement: Social Class Differences in Undergraduates' Academic Strategies." *Journal of Higher Education* 87 (6): 831–58.

Early Career Paths of Economists Inside and Outside of Academia

Lucia Foster, Erika McEntarfer, and
Danielle H. Sandler

Over 20 years ago, Richard Freeman (1999) noted how much better the job market was for economists relative to other mathematical fields. He conjectured this was due in part to strong public and private sector demand for economists outside of academia. However, Freeman cautioned, "don't tell anyone." Wider knowledge of the comparatively attractive job market might attract more graduate students, creating less favorable job market conditions. In the decades that followed, the number of new PhDs in economics did increase. Yet the job market remains strong, in part due to growing demand for economists outside of academia.

Currently just over half of PhD economists begin their careers as professors in research-oriented universities or teaching colleges. Nearly half begin other careers that encompass a wide range. Some work for technology companies, others alongside lawyers in consulting firms. There is vast career diversity within the public sector: economists there develop macroeconomic forecasts, work on antitrust cases, advise policymakers, and improve economic measurement. But despite a large share of nonacademic placements, study of the economics profession has mostly focused on academia. This imbalance means new PhD economists navigate

■ *Lucia Foster is Chief of the Center for Economic Studies and Chief Economist, Erika McEntarfer is Lead Economist for the Longitudinal Employer-Household Dynamics (LEHD), and Danielle H. Sandler is a Principal Economist, all at the US Census Bureau, Washington, DC. Parts of this paper were written while McEntarfer was on detail as a Senior Economist at the Council of Economic Advisers, Washington, DC. Their email addresses are lucia.s.foster@ census.gov, erika.mcentarfer@census.gov, and danielle.h.sandler@census.gov.*

significant unknowns during their job search; nonacademic job offers represent a diverse set of career paths about which they know comparatively little. The stakes of the decision are high: growing compensation gaps between academic and private sector jobs means the choice of career path has significant financial consequences. Career mobility is also a concern; many job candidates express fears they will be unable to return to academia if they accept a nonacademic offer.

Our objective here is to fill some of these knowledge gaps by providing a broader look at economists' early career paths—both in and outside of academia. We do this by developing a novel large panel dataset of earnings and employment after completing an economics PhD. This data allows us to look beyond the snapshot of initial placements to examine employment and earnings dynamics over the first ten years of a career. This ten-year period encompasses the standard tenure clock and early career job moves and thus provides more information about the ultimate career path.[1] While our data cannot speak to whether a particular career path is right for a specific individual, we can make more explicit how different placements shape future financial outcomes and how much medium-term job stability or mobility to expect after initial placement.

We also hope to shed additional light on the role that the rank of a PhD program plays in determining early career outcomes, both in and outside of academia. For undergraduates applying for admission to graduate schools, attending a top-ranked doctoral program appears a necessary, if not sufficient, condition for success at the top of the profession. In the United States, most tenure-track faculty at PhD-granting institutions earned their PhDs from one of the top 15 PhD programs (Jones and Sloan 2021). In addition, five top PhD programs have produced half of the members of the highly selective National Bureau of Economic Research (Kleemans and Thornton 2021). Graduate students from higher-ranked programs obtain more second-round interviews on the job market and receive higher initial academic and nonacademic salary offers (Stock and Alston 2000).[2] Yet these facts mask a common observation of professional life: there is significant dispersion in outcomes for graduates of the same PhD program. Our data allow us to characterize some of that dispersion and also to explore whether salary gaps observed in initial placements grow or shrink over time. We were also curious whether PhD program rank is more important for some career paths and less important for others. Our data suggest that, at least with respect to early career earnings growth, PhD program matters more inside of academia than outside of it.

---

[1] Data limitations severely limit the study of economists' careers after initial placements, especially outside of academia. Bedard, Lee, and Royer (2021) summarize the current literature as relying on cross-sectional data, short early career panels, or panels specific to a narrow set of institutions.

[2] The hierarchical nature of the economics profession has been extensively discussed (for example, in this journal by Fourcade, Ollion, and Algan 2015). Rankings play a crucial role in the lives of economists: the ranking of their PhD-granting department, the rankings of the journals in which they publish, the monthly rankings of individuals in Research Papers in Economics (RePEc), and the constantly updated implicit rankings via Google Scholar.

Before we describe our main findings, some background about our panel data seems useful. To identify PhD economists, we use the Survey of Earned Doctorates, an annual census conducted by the National Center for Science and Engineering Statistics (NCSES). This survey covers individuals receiving a research doctorate from an accredited US institution for that academic year; for example, the 2017 data cover the academic year from July 1, 2016, to June 30, 2017. The coverage of PhD-granting departments in the survey is high, including over 90 percent of new economics PhDs in the United States. We then link respondents to administrative data on jobs sourced primarily from the Longitudinal Employer-Household Dynamics (LEHD) data, which cover over 95 percent of private-sector workers and almost all state and local government employment. We supplement this with Internal Revenue Service form W-2 data to obtain earnings for jobs not covered by LEHD, particularly federal workers. We link the Survey of Earned Doctorates person-level data to the jobs-level data using Protected Identification Keys assigned to individuals by the Census Bureau's Person Identification Validation System.

The resulting panel dataset covers approximately 12,000 new PhD economists who completed their degrees between 2001 and 2017 and were employed in the United States following graduation. While this panel is very large, it contains notable exclusions. In particular, economists trained outside of the United States are not represented in our data, nor are those trained in the United States who are employed outside of the country. More than one-third of economics PhDs from US institutions do not end up working in the United States. This in part reflects the international composition of economics graduate education. The majority of economics PhDs granted by US institutions are awarded to temporary visa-holders; about half of these students have initial placements in the United States.[3] We assume the vast majority of graduates from US economics PhD programs for whom we do not observe subsequent earnings in our data have placements outside of the United States; where possible we compare our findings to available data that includes this group.[4]

---

[3]The share of economics PhDs granted by US institutions awarded to temporary visa-holders was 52 percent in 2017 (NCSES 2018, Table 62). Fifty-five percent of these temporary visa-holders with definite post-graduation plans had initial placements in the United States, compared to 95 percent for US citizens and permanent residents (NCSES 2018, Table 51). The percentage of new PhD economists with definite postgraduation plans whose initial placement is in the United States has decreased over time, from 77 percent in 2007 to 72 percent in 2017 (NCSES 2008, Table S38; 2018, Table 51).

[4]The experiences of those who are not US citizens in economics PhD programs at US universities deserve further attention; that is, their motivations and expectations in coming to the United States for an economics PhD program, their experiences in the program (including their access to advisers, research assistant positions, and financial aid), and the jobs they take in other countries after graduation. The experiences of US citizens who end up moving abroad (a much smaller group) would also be of interest.

Result 1: The share of academic placements is falling, even among graduates of top-ranked economics PhD programs.

The traditional focus of the PhD job market is on academic jobs. However, academia now accounts for a declining share of new PhD economist placements. We observe this declining trend in our panel, but it is also notable in the broader census of new PhDs. The Survey of Earned Doctorates indicates a drop in initial placements among economists in academia and corresponding rise in placements in industry. In 2007, of those new PhD economists with definite employment commitments (including those both inside and outside the United States), 60 percent went into academia, 14 percent into government, 17 percent into industry, and 4 percent into nonprofits. By 2017, those numbers changed to 53 percent for academia, 13 percent for government, 26 percent for industry, and 6 percent for nonprofits (NCSES 2008, Table S38; 2018, Table 63). Published results focusing on US placements (which are more comparable to our analyses) do not provide as much temporal or sectoral detail, but reveal that the share going into academia in 2017 was 48 percent (NCSES 2018, Table 51).

Turning to our data, we are able to provide some finer detail on economists' initial placements in Table 1. Just over half of the economists in our panel data begin their careers in academia, but this share is falling over time, declining from 64 percent for 2001–2005 graduates to 56 percent for 2014–2017 graduates. Meanwhile, the share with initial placements in industry and government grew: placements in consulting firms increased from 9 percent to 13 percent, in government from 5 percent to 11 percent, and in technology firms from 4 percent to 6 percent.[5] The share of economists with initial placements in finance is relatively stable, at around 9 percent. With the exception of growth in government placements, which may be unique to the United States (the global financial crisis spurred demand for economists in the public sector), broad patterns in initial placements and trends look very similar in the Survey of Earned Doctorates.

We were curious whether the share of nonacademic placements varied by rank of the PhD-granting institution. Graduates from higher-ranked programs are more likely to receive multiple offers; these graduates are more likely to have their choice of options. Table 1 shows placements for 2014–2017 graduates by program rank. Surprisingly, the share of academic placements varies little across PhD program ranks. Among recent graduates of the top 5 PhD programs in economics, just over

---

[5] We define these sectors using NAICS codes as follows: academia (61, Educational Services), consulting (54, Professional, Scientific, and Technical Services but excluding 541511-541519), government (92, Public Administration), finance (52, Finance and Insurance), technology (454110, E-Commerce; 511210, Software Publishers, 517311-519190, Internet Publishing; 541511-541519 Computer System Design; 485310 Taxi Service; and 721100, Traveler Accommodation), and other industry (all other NAICS codes). Over 95 percent of those in academia are employed in NAICS 6113, Colleges and Universities, with the remainder employed in community colleges and other educational services. If an economist works more than one job during the calendar year, we determine the sector using the employer that was the primary source of earnings in that year.

Table 1
Initial Placements by Sector and Career Mobility across Sectors

Initial placement	Across cohorts (2001–2017)				By program rank (2014–2017 cohort)			Ten years later (2001–2007 cohort)		
	2001–2005	2006–2009	2010–2013	2014–2017	Top 5	Top 6–20	Other	Stay	Move	Destination
Academia	63.64	58.62	60.00	55.56	55.56	50.00	54.17	84.21	5.26	Gov't
Consult	9.09	10.34	11.67	12.50	13.33	14.29	10.42	50.00	13.33	Academia
Government	4.55	8.62	10.00	11.11	6.67	11.43	10.42	80.00	16.00	Academia
Finance	9.09	8.62	8.33	9.72	4.44	5.71	12.50	80.00	8.00	Academia
Other Ind.	9.09	6.90	6.67	6.94	4.44	7.14	6.25	45.00	25.00	Academia
Tech	3.64	3.45	5.00	5.56	8.89	7.14	4.17	40.00	20.00	Academia
N	2,200	2,900	3,000	3,600	450	700	2,400	3,100		

Source: Authors' calculations from matched SED/LEHD microdata.
Note: Shares are rounded using Census rounding rules, and columns may not sum to 100 percent. Share academic placements are monotonically decreasing over time in unrounded data. "Move" and "Destination" columns show only the most frequent industry destination and the share of initial placements working in that sector ten years later. Program rank defined by rank of PhD-granting institution.

half (56 percent) have their initial placement in academia, which is roughly similar to those in programs outside the top 20 (54 percent). The middle category of top 6–20 programs have fewer academic placements (50 percent) and the highest share of government placements (11 percent).

Some further explanation of how we categorize PhD program rank in Table 1 and throughout this paper is needed here. We begin by recognizing that existing rankings of PhD programs are imperfect. While there is some consensus in the profession about which programs are "good" with respect to ordinal rankings, there are, as they say, "20 departments in the top 10." Any attempt to bin programs using these rankings is subject to the criticism that they are measured imperfectly; we do our best. Our approach uses Research Papers in Economics' (RePEc) US departmental rankings and bins them into three categories: top 5, top 6–20, and other.[6] Careful examination of multiple ranking regimes and earnings fixed effects informed this approach. Our preliminary investigation indicated that the fixed effect on earnings from being a graduate of a particular program (conditional on sector of employment and other observable characteristics) was largest among schools typically ranked among the top 5, varied more considerably among schools outside the top 5, and was measurably different than zero in few programs outside of the top 20. A feature of most ranking methods is that they generally concur on the very top programs (a top 5 bin does appear to collect a set of schools widely believed to be among the best) as reflected in earnings and rankings. The choice

[6]We examined several different published department ranks, and our main results were robust to ranking methodology. This likely indicates they are mostly equally imperfect methods of ranking schools by student outcomes.

to place the cutoff for the second bin at 20 is more arbitrary—rankings and earnings fixed effects are notably noisier in this range and our results would not much change if we had set the cutoff at 25.

Turning back to Table 1, we find that PhD department rank does impact the type of nonacademic career path chosen. In particular, graduates of top 5 programs are much less likely to seek careers in government (including the Federal Reserve Board), although this share is increasing. There are so few government placements among top 5 candidates in the first decade of our data that we have to drop this sector in some of our later analysis. Consulting firms are the most frequent nonacademic placement for candidates from top 5 and top 6–20 programs (13 percent and 14 percent respectively). We also find that the recruitment of economists by tech firms has increased in recent years, particularly from top 5 departments.[7] Nine percent of US placements of economics PhDs from the top 5 economics departments were at technology firms in 2014–2017, a higher share than placements in either finance or government. The recent dominance of technology firms in recruiting candidates from top programs may partly explain the outsized interest in these jobs, which remain a relatively small share of the overall economist job market. Given the recent layoffs at tech firms, it will be interesting to see whether these patterns change in the future.

Result 2: Early career paths for PhD economists show considerable mobility; in particular, the door between academia and industry swings both ways.

We find considerable job mobility for economics PhDs after initial placements, with many economists changing jobs, often multiple times, in the first ten years of their career. Movements across sectors are also not uncommon. Table 1 shows transitions across sectors from initial placement to placement ten years later for PhD economists for 2001–2007 graduates. Among career tracks, academia is the stickiest, 84 percent of initial placements in academic jobs are still in academia ten years later. Academia is also a draw for those in industry—a sizable share of economists move from industry jobs to academia in the early years of their career. Despite conventional wisdom that initial academic placements are preferred because the "door only swings one way," the door very much does not.

We provide a more granular look at early career paths in Figure 1. Panel A of Figure 1 graphs early career moves for new PhDs whose initial placement was at a research university.[8] As shown in panel A, just over 40 percent of those who start

---

[7] In this journal, Athey and Luca (2019) note that Amazon, which hired over 150 economists from 2014–2019, employs more economists than the largest economics departments.

[8] Specifically, we define a research university here as a top-70 department as ranked by RePEc. We choose the top 70 departments as it is a threshold that includes most PhD-granting departments and is expansive enough to allow us to capture some upward mobility. Also, while the population of graduates from top programs is large, initial placements in research universities are much smaller. Thus, it is largely for

Figure 1

Early Career Moves after Initial Placement

2001–2010 Graduates

Panel A. Initial placement in a top-70 academic department, 2001–2010 graduates
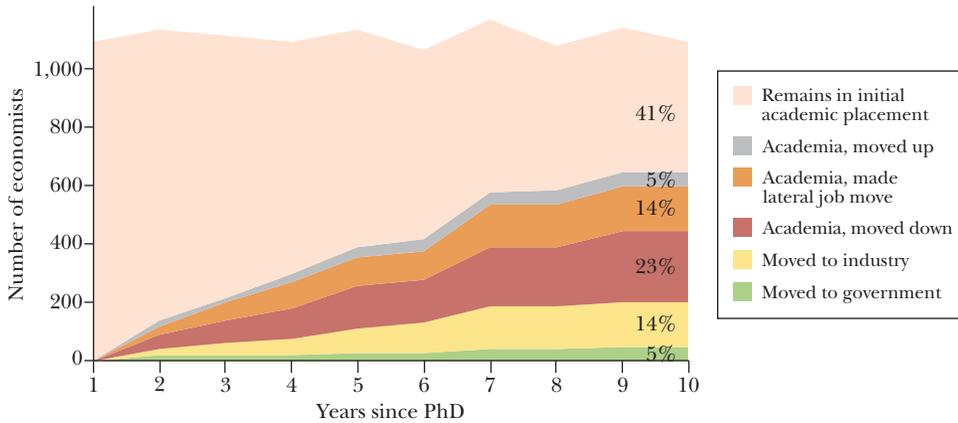
Panel B. Initial placement in consulting or finance, 2001–2010 graduates
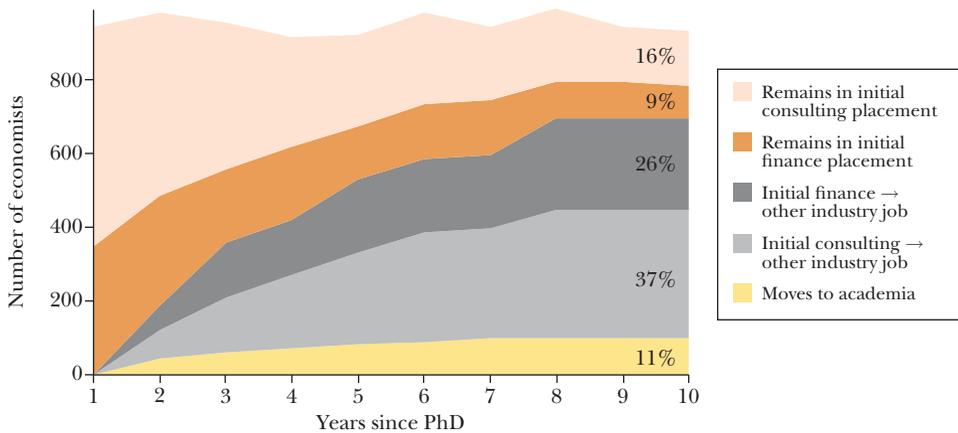
Source: Authors calculations from matched SED/LEHD microdata, 2001–2010 graduates with ten years of continuous earnings only.

Note: Figure shows flows from initial placement to next employer type by industry and time. Top-70 departments as ranked in RePEc, movements up indicate new department is at least 15 ranks higher, movement down means new department at least 15 ranks lower, other moves are lateral moves. For industry moves in panel B, the very few moves into government are pooled with other industry jobs.

their career in a research university remain in their initial placement ten years later. This is a much lower share than academic placements overall (64 percent). This

disclosure reasons that when we categorize job mobility we expand our analysis to include those who place in a department ranked between 1 and 70 instead of using the more restrictive top 5 and top 6–20 groupings in the previous section.

difference appears to be driven in part by a greater likelihood of a negative tenure decision at research universities; there is an uptick in job change at year seven, as can be seen in panel A of Figure 1. But it also reflects greater early career mobility for faculty at research institutions pre-tenure. Among those placing in research departments, one-third have moved elsewhere by the fifth year of their career, compared to one-fifth among academic placements overall.

Are job-changers at research departments generally moving up or down the job ladder? We find evidence of upward mobility for a few, with 5 percent moving up at least 15 ranks within ten years. However, the modal job-changer is moving down, with 23 percent working in a department at least 15 ranks lower than their initial placement in year ten. There are also many lateral moves (14 percent have moved laterally by year ten), disproportionately between the highest-ranked departments. Interestingly, candidates who place in research institutions are more likely to leave academia later than those who did not; 19 percent are working in industry or government in year ten, compared to 9 percent among academic placements overall.

A potential draw for economists leaving academia is higher pay: academics at research institutions who move to industry or government experience higher earnings growth over their early career than those who remain in academia (7–8 percent annualized earnings growth, compared to 4–6 percent). But even within academia, those who change jobs generally experience greater early career earnings growth than those who remain with their initial placement, including those who move down in department rank (5–6 percent annualized earnings growth, compared to 4 percent).[9]

We now turn to early career paths outside of academia. Economists whose initial placement is in a consulting or finance firm change jobs earlier and more frequently than those in academia. Just one-quarter remain in their initial placement ten years later, as shown in panel B of Figure 1. Most job changers remain in the private sector, and job changers in industry appear to be moving up the job ladder, as evidenced by higher earnings growth among job switchers. However, a sizable minority (11 percent) move to academia. Those who move back to academia tend to do so fairly quickly: over half of moves to academia occur in the first three years after finishing the PhD.

In additional analyses, we find that economists with initial placements in consulting are much more likely to move to academia in the first ten years of their career than those in finance (15 percent and 5 percent, respectively). Those who initially place in finance and banking are less likely to remain in their initial job, but generally remain in the same sector. Finance and banking jobs have very high salaries and high rates of wage growth compared to other sectors in which economists work, which likely deters leavers. Those who leave consulting for academia do

---

[9] One reason we may see slightly higher earnings growth among those stepping down the job ladder than those that remain with initial placements is that some of these moves may be resolving joint location problems. It seems plausible that such moves could relocate academics to larger labor markets and more expensive regions of the country.

experience very low rates of early career wage growth (3 percent annualized earnings growth in first ten years), as these moves are often associated with an initial earnings loss and are presumably motivated by preferences for the nonpecuniary aspects of academic jobs.

Because most technology firm placements are quite recent, we can only look at the first five years post-PhD for the cohort of 2001–2015 graduates. After five years, 43 percent remain with their initial employer, another 17 percent remain in the tech sector with a different employer, and 11 percent have taken an academic job. Despite talk of high salaries in technology firms, the average economist who remains with their initial tech placement earned about $140,000 annually during the fifth year of their career. This income is certainly a good one, but not astronomical compensation for a PhD economist working in the private sector. However, while our earnings data include bonuses, they do not include the value of stock options that may be part of the compensation package for these economists.

Result 3: Industry pays better than academia for PhD economists, but the distribution of earnings in each sector is broad and overlapping.

Academia has long been viewed by many as the most prestigious career path for economists, but industry jobs are widely understood to be more lucrative. These generalities mask the fact that some economics professors are very highly compensated, and that not every private sector job is a path to wealth. In other words, the distribution of compensation within sectors is vast, and earnings distributions in academia and industry overlap. An advantage of our data is that we can show the entire earnings distribution for our panel of economists and track its evolution over time. We find that earnings and earnings growth for early-career economists varies widely across and within sectors, leading to large dispersions in pay in mid-career.

Panel A of Figure 2 shows the annual earnings distributions for economists employed in academia, government, and industry for initial placements.[10] We collapse the industries in which economists work into three broad sectors: academia, government, and "industry" (which combines consulting, finance, other industry, and tech). Economists working in academia generally have the lowest initial earnings, followed by those in government, then industry. The median real (2015) annual earnings are $83,270 for academia, $91,600 for government, and $108,600 for industry.[11] Industry has the highest earnings, but also the most dispersion; the

---

[10] Some differences between this analysis and the early career panel in the last section should be noted. The previous section looked at balanced panels of economists who worked continuously in the United States for all ten years after receiving their PhD (or five years, in the case of technology placements). Here we look at earnings for all economists working in the United States in a particular year.

[11] The Survey of Earned Doctorates also provides information on median annual salaries for doctorate recipients with definite postgraduation plans for employment in the United States: $75,000 for academia, $85,000 for non-profits, $95,000 for government, and $105,000 for industry (NCSES 2010, Table 45).

Figure 2
Economist Earnings by Sector
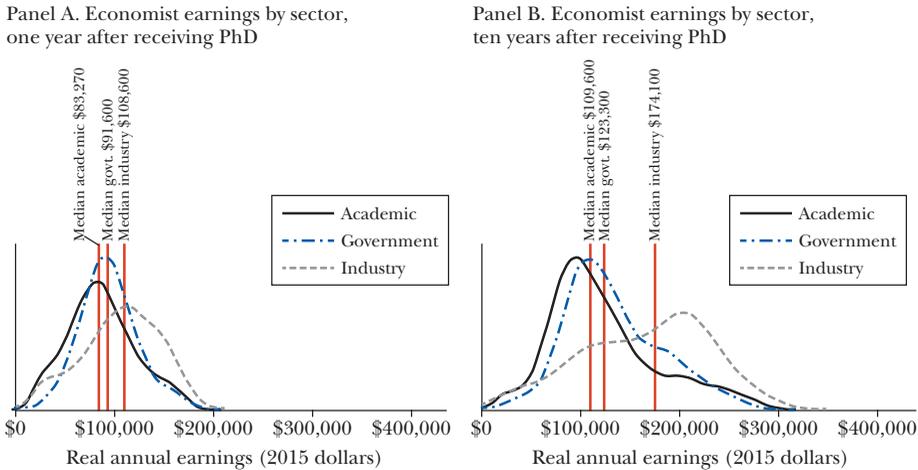
Panel A. Economist earnings by sector,
one year after receiving PhD

Panel B. Economist earnings by sector,
ten years after receiving PhD

Source: Authors' calculations from matched SED/LEHD microdata, 2001–2007 SED graduates only.
Note: Figure shows earnings distribution of economists by sector at initial placement and ten years after PhD. Academic sector is defined as all economists working in NAICS 61 (almost all of whom work in colleges and universities), government is those employed in NAICS 92 (which includes the Federal Reserve Board and other federal agencies) and industry is all other sectors. As the y-axis in a kernel density plot represents an infinitesimal quantity of probability space, we removed the y-axis labels to improve legibility of the figure.

difference between the 75th and 25th percentile is about $54,000 a year, compared to $46,000 in academia, and $36,000 in government. In Table 2, we summarize the 25th percentile, median, 75th percentile, the interquartile range, and the percent difference for these earnings distributions.

By mid-career, dispersion in earnings among economists has become quite large. Panel B of Figure 2 depicts the earnings distributions for economists who have been in their careers for ten years. One striking observation is the comparatively more rapid early-career wage growth for economists in industry. Specifically, median earnings for industry economists rise from $108,600 in their initial placement to $174,100 a decade later. Thus, while the median industry economist earns 30 percent more than the median academic economist in their initial placement, this difference increases to 60 percent after ten years.

We should note that our measure of earnings does not include self-employment. As noted by Fourcade, Ollion, and Algan (2015, p. 90), "[M]any prominent economists have the opportunity to obtain income from consulting fees, private investments and partnerships, and memberships on corporate boards." These opportunities probably differ across sector, fields, and personal characteristics.

Converting these to 2015 dollars as in Figure 2, these medians are very similar to what we show in panel A: $81,522 in academia, $103,261 in government, and $114,130 in industry.

Table 2
Earnings Distribution Moments for Figures 2–3
(in 2015 dollars)

	P25	P50	P75	Dispersion P75–P25	Pct. diff P75–P25
Initial year after PhD (Fig. 2a)					
Academia	62,210	83,270	107,800	45,590	73 percent
Government	73,880	91,600	110,100	36,220	49 percent
Industry	80,060	108,600	134,300	54,240	68 percent
Ten years after PhD (Fig. 2b)					
Academia	84,430	109,600	150,000	65,570	78 percent
Government	98,170	123,300	166,700	68,530	70 percent
Industry	116,600	174,100	213,800	97,200	83 percent
Academia: initial year after PhD (Fig. 3a)					
Top-5 PhD	94,430	119,800	146,800	52,370	55 percent
Top 6-20 PhD	74,100	104,800	127,800	53,700	72 percent
Non-top 20	54,140	75,920	92,850	35,710	62 percent
Academia: ten years after PhD (Fig. 3b)					
Top-5 PhD	127,600	189,000	230,200	102,600	80 percent
Top 6-20 PhD	99,920	138,800	189,800	89,880	90 percent
Non-top 20	79,000	99,170	125,300	46,300	59 percent
Industry: initial year after PhD (Fig. 3c)					
Top-5 PhD	107,000	135,600	152,000	45,000	42 percent
Top 6-20 PhD	95,720	124,900	147,200	51,480	54 percent
Non-top 20	74,900	99,080	122,600	47,700	64 percent
Industry: ten years after PhD (Fig. 3d)					
Top-5 PhD	175,900	208,500	237,000	61,100	35 percent
Top 6-20 PhD	141,100	198,700	230,100	89,000	63 percent
Non-top 20	107,000	157,000	203,000	96,000	90 percent

Source: Authors calculations from matched SED/LEHD microdata, 2001–2007 SED graduates only.
Notes: Industry is defined as all economists who work outside of NAICS 61 or 92, who are concentrated in consulting and finance. Dispersion P75–P25 is the absolute difference in dollars between the 75th and 25th percentiles. Pct. diff is the percentage difference ("P75 is X percent higher than P25") calculated as (P75–P25)/P25.

Weyl (2017) estimates that academic economists receive 20 percent of their income from outside activities, including consulting, and his rough estimates suggest outside income may be as high as 40 percent in fields such as finance and industrial organization. However, according to Del Rossi and Hersch (2020), while consulting is relatively common among economists (more than 40 percent reported legal consulting in the last five years), it contributes less than 10 percent of income, based on 1,205 responses to their survey sent to 8,157 faculty in economics departments. We do not capture this secondary income in our current analysis but acknowledge it could be important.

Comparing panels A and B of Figure 2 displays a significant increase in earnings dispersion, reflecting within-sector differences in early-career earnings growth. Mid-career academic and government economists have right-skewed earnings distributions, indicating that wage gains have been uneven and concentrated among economists in the right tail. Mid-career industry economists, on the other hand, have a much flatter and unusually left-skewed earnings distribution, indicating larger dispersion in earnings gains, but favoring a larger share of industry economists.[12]

Result 4: Initial earnings and earnings growth in academia is higher among those from top-ranked PhD-granting institutions. However, there is enormous dispersion in mid-career earnings among graduates of top programs.

Graduating from a highly ranked economics PhD program could lead to higher earnings for three overlapping reasons: (1) selection effects imply that top programs are more likely to select highly talented candidates; (2) more prestigious programs might provide better training for job candidates; and (3) attending an elite program may give students access to powerful networks within the profession. Selection and signaling effects are likely to have a stronger impact on initial placements outside of academia where interviewers may not be as skilled in assessing candidate quality. Powerful networks, on the other hand, are most likely to benefit graduates who remain in academia.

As we are unsure whether selection, training, or network effects dominate, we examine the role of rank for academic and nonacademic jobs separately. Panel A of Figure 3 shows earnings distributions for academic economists one year after receiving their PhD from top 5, top 6–20, and other programs. Median earnings for economists are $119,800 for those who received their PhD from a top 5 institution, $104,800 for those from a top 6–20 institution, and $75,920 for those from a non-top-20 institution. In short, the distributions show evidence of higher initial earnings for graduates from top departments in academia, although these distributions overlap considerably.

Ten years later, differences in median earnings between top program graduates and others have grown and dispersion among classmates from similarly ranked programs has increased considerably. In panel B, the median earnings of a graduate from a top 5 PhD program working in academia is almost double that of a graduate from a school outside the top 20, and considerably higher than the median graduate from a top 6–20 program. But the differences in medians mask enormous dispersion in earnings for academics graduating from top programs.

---

[12] Our findings on the pattern of dispersion in academic earnings for PhD economists are consistent with the main alternative source of data from the Universal Academic Questionnaire (UAQ) which shows there is considerably higher pay and more dispersion in earnings at the top of the profession.

Figure 3

Economist Earnings in Academia and Industry by PhD Program Rank

Panel A. Economist earnings in academia by program rank, one year after receiving PhD
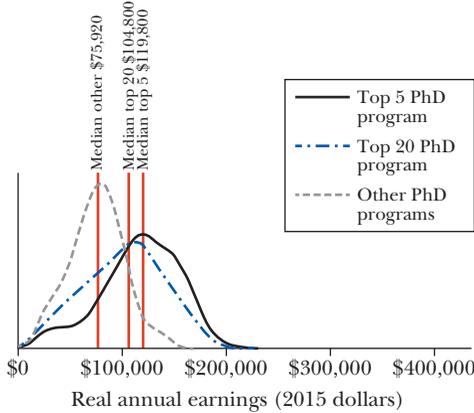
Panel B. Economist earnings in academia by program rank, ten years after receiving PhD
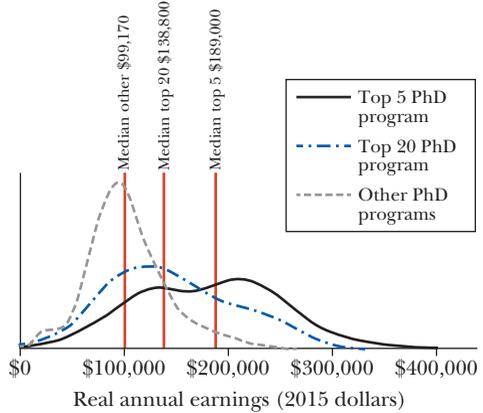
Panel C. Economist earnings in industry by granting-program rank, one year after receiving PhD
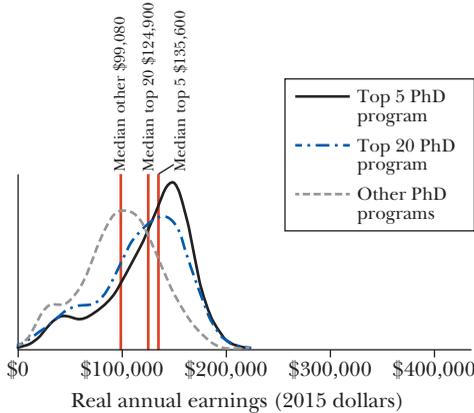
Panel D. Economist earnings in industry by granting-program rank, ten years after receiving PhD
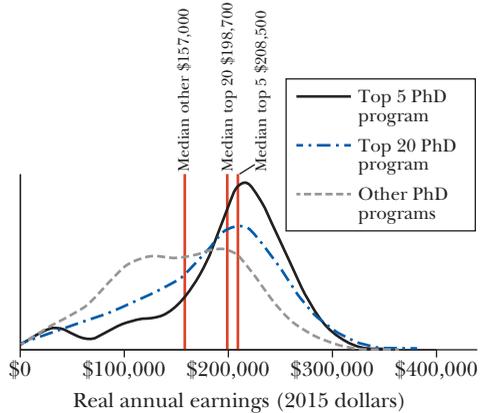
Source: Authors' calculations from matched SED/LEHD microdata, 2001–2007 SED graduates only.
Notes: Figure shows earnings distribution by program rank of PhD-granting institution for economists working in academia and industry at initial placement and ten years after PhD. Academia is defined as all economists working in NAICS 61, almost all of whom work in colleges and universities. Industry is defined as all economists who work outside of NAICS 61 or 92, who are concentrated in consulting and finance. As the y-axis in a kernel density plot represents an infinitesimal quantity of probability space, we removed the y-axis labels to improve legibility of the figure.

As shown in Table 2, the interquartile range for top 5 and top 6–20 graduates is now 80 percent and 90 percent, respectively. Phrased differently, among graduates from top PhD programs, academics in the 75th percentile earn almost twice what their classmates make at the 25th percentile.

Result 5: For PhD economists in industry, wage disparities by rank of PhD-granting institution are lower than in academia.

Panel C of Figure 3 shows initial earnings distributions in industry by program rank. One year after receiving a PhD, economists in industry have median earnings of $135,600 if they graduated from a top 5 institution, $124,900 if they graduated from a top 6-20 institution, and $99,080 if they graduated from a non-top-20 institution. Compared to academia, the initial earnings distributions for graduates of top 5 and top 20 programs overlap a great deal more in industry, a pattern that holds throughout the early career. Panel D shows the distributions for industry ten years after receiving PhD. Median earnings for an economist working in industry ten years after receiving their PhD is $208,500 if from a top 5 program, $198,700 if from a top 6–20 program, and $157,000 if from a non-top-20 program.

For graduates of top programs, the earnings risk in academia is much greater than in industry (as can be seen by comparing panels B and D). In academia at year ten, the 75th percentile graduate from top 5 program makes 80 percent more than their classmate in the 25th percentile, but in industry the 75th percentile earns only 35 percent more than their classmate in the 25th percentile (Table 2).

In academia, the earnings gap between top graduates and everyone else grows markedly in the first ten years of their careers. For economists in industry, earnings gaps at the median are smaller for initial placements and shrink slightly during the first ten years of their careers. While PhD program rank does matter for earnings in industry, wage disparities are smaller.

In both academia and industry, PhD program rank impacts initial placement earnings, with top program graduates earning more. This is consistent with program rank providing a quality signal for initial placements that fades as true productivity is revealed. The fact that graduates of top programs develop professional networks in graduate school that are more advantageous for promotion and wage growth within academia than outside of it may explain why program rank matters for earnings growth in academia but matters considerably less for promotion outside of academia.

These dynamics mean that earnings penalties for remaining in academia after the PhD are much larger for graduates of non-top PhD programs. Ten years after their PhD, earnings for the median academic are only 66 percent of those for the median industry economist. However, these overall profession medians disguise the large dispersion in earnings gaps by program rank. Due to the different growth dynamics described above, the median graduate from a top-5 program in academia earns 90 percent of what a graduate from a top-5 program in industry earns ten years after graduation, while a graduate from a top-20 program earns only 70 percent, and a graduate outside the top-20 earns only 63 percent of what they would earn in industry. This is because earnings gaps in industry are far less pronounced, especially between top 5 and top 6–20 graduates, and do not grow larger over the course of the early career of PhD economists.

Result 6: Initial earnings and earnings growth are both lower for women with an economics PhD.

Women remain underrepresented in the economics profession, accounting for roughly one-third of PhDs granted each year between 2001 and 2017.[13] Although the share of women in the field has risen slightly over time, from 28 percent in 2001 to 34 percent in 2017, the share has essentially remained steady since 2008 (NCSES 2020, Table 15; 2012 Table 17). In this journal, Lundberg and Stearns (2019) offer a comprehensive overview of the "stalled progress" for women in the profession and research into the root causes of the issue.

According to the Survey of Earned Doctorates for academic year 2017, greater shares of women say they plan to work in academia relative to men (55.4 percent compared to 51.9 percent) and smaller shares plan to work in government (9.2 percent versus 14.8 percent) (NCSES 2018, Table 63). Despite this, when we look at placement outcomes, we find that gender diversity is slightly better in government, with 38 percent of economists being women, and in industry, with 36 percent being women, than in academia where only 33 percent of economists are women. This difference may be driven by gender differences in flows between the sectors as well as US jobs versus jobs in other countries.[14] An interesting question concerns whether there are differences in labor force participation of PhD economists by gender. Unfortunately, we cannot provide an answer to labor force status, because we are only able to see the outcomes for those who work in the United States and whose employment is captured by the Longitudinal Employer-Household Dynamics (LEHD) data. An individual who does not appear to be employed in our data could be in the United States and not employed, but could also be employed elsewhere.

---

[13] Underrepresentation by race and ethnicity in economics is also an issue, but we are unable to examine these dimensions for our exercises due to small sample sizes. See Foster, McEntarfer, and Sandler (2023) for related exercises by gender, race, and ethnicity of economists. Numerous organizations aim to improve diversity in the profession, and many of them monitor the placements and earnings of economists by demographic characteristics including gender, race, and ethnicity (for example, Hoover and Washington 2021).

[14] Our results are broadly similar to the survey conducted by the American Economic Association's Committee on the Status of Women in Economics Profession (CSWEP), which is distributed to over 100 doctoral and over 100 nondoctoral programs each year. The CSWEP annual report shows the job placements of new PhDs from the top 10 and top 20 economics departments. In 2017, 31.2 percent of women from top 10 departments were placed in academia in doctoral departments, 46.2 percent were placed in the public sector, and 25.0 percent were placed in the private sector. For women from top-20 departments the corresponding percentages were 26.4 percent in academia in doctoral departments, 31.8 percent in public sector, and 29.5 percent in private sector (CSWEP 2018, Table 3). Thus, women economists from top 10 departments are more likely to be in academia and the government and less likely to be in industry than women from top 20 departments. However, the same report shows that 28.8 percent of assistant professors, 26.1 percent of untenured associate professors, 23.0 percent of tenured associate professors, and 13.9 percent of full professors in economics at doctoral-granting institutions were women (CSWEP 2018, Table 1). Our finding that 33 percent of our sample of 2001–2017 PhD recipients working in academia are women indicates both that the more recent sample of recipients have more female representation and that there is higher female representation at the lower-ranked schools not included in the report's placement tables.

We also do not find that women are disproportionately employed in particular sectors when examining a single-year snapshot of the employment of all economists in our sample: there is a roughly one-third share in academia, government, and in industry jobs (Foster, McEntarfer, and Sandler 2023).

Even at initial placement, the earnings for women economists are lower than that of men economists. The Survey of Earned Doctorates provides median expected annual salary for doctorate recipients with definite post-graduation plans. In 2017, these medians for new PhD economists were $109,000 for men and $100,000 for women (NCSES 2018, Table 48). These gaps differ by sector, with the largest gender earnings gaps for economists entering academia from top-20 PhD programs and lower gaps for those from lower-ranked programs and those entering government.

To understand the career progression for economists by gender both within and outside of academia, we present descriptive evidence on mid-career earnings, ten years after PhD receipt, by sector. Figure 4 displays the earning distributions for men and women in government (panel A), industry (panel B), academia for graduates from non-top-20 programs (panel C), and academia for graduates from top-20 programs (panel D). We also provide the median and 75th percentile earnings for men and women in each panel. We split the academic earnings distribution by program rank.

Starting with government, the level of median earnings for mid-career women economists is $117,900 while for men it is $126,400—that is, median earnings for women is 7 percent less than median earnings for men. The gender gap in median earnings in government is relatively small, but it grows larger at the upper tail. At the 75th percentile, women economists earn $152,700 and men economists earn $173,800, a 14 percent gap. In industry, the earnings distributions are left-skewed, which means that the gender gap narrows at higher points in the distribution and the gender gap manifests as a fatter left tail instead of a thinner right tail. The median earnings are $156,000 for women and $182,300 for men, a 17 percent gap. At the 75th percentile, women economists earn $205,100 while men economists earn $216,500, so women earn 6 percent of the earnings for men at the 75th percentile. In industry, mid-career men and women have very different earnings distributions, but gaps narrow as one moves further up the distribution. Although consulting, tech, and finance are not sectors known for being particularly friendly to women, we find gender earnings gaps are more pronounced in academia, both in initial placement earnings and earnings growth over time.

Given that most of the dispersion in earnings in academia is among graduates from top programs, it is not surprising that this is where we find the largest gaps between men and women. For academics from programs outside the top 20, median earnings for mid-career women is $93,350 while for men it is $101,400, or median earnings for women is 9 percent less than median earnings for men. For graduates from programs in the top 20, the median earnings for women is $132,200 and median earnings for men is $166,600, or median earnings for women is 26 percent less than median earnings for men. The general takeaway from Figure 4 is that earnings gaps are largest for women in academia, especially among graduates of top programs.

Figure 4
Economist Earnings by Sex, Ten Years after Receiving PhD
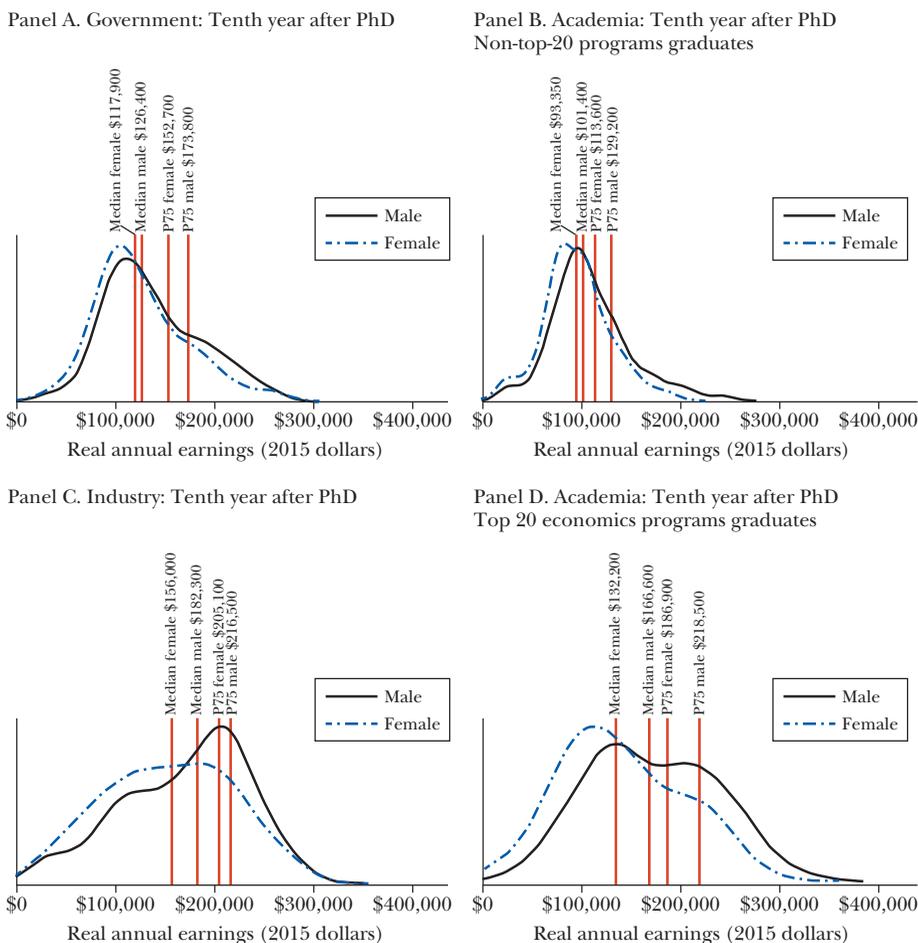
Panel A. Government: Tenth year after PhD

Panel B. Academia: Tenth year after PhD
Non-top-20 programs graduates

Panel C. Industry: Tenth year after PhD

Panel D. Academia: Tenth year after PhD
Top 20 economics programs graduates

Source: Authors' calculations from matched SED/LEHD microdata, 2001–2007 SED graduates only.
Notes: Figure shows earnings distribution by gender across sector for mid-career economists (ten years after PhD). Academic sector is defined as all economists working in NAICS 61 (almost all of whom work in colleges and universities), government is those employed in NAICS 92 (which includes the Federal Reserve Board and other federal agencies), and industry is all other sectors. As the y-axis in a kernel density plot represents an infinitesimal quantity of probability space, we removed the y-axis labels to improve legibility of the figure.

Result 7: The gender gap in growth rates of early career earnings for women PhDs stems from a combination of initial earnings gaps and lower earnings growth.

The gaps in earnings growth by gender could either come from smaller and less frequent promotions, changes in earnings within an employer, or differences in

movement between employers. Bedard, Lee, and Royer (2021) conducted a study of 1,356 academic economists at the top 50 institutions as ranked by *US News & World Report*. They find that gender gaps in salary are very small at the start of careers, but become visible within ten years of PhD attainment. Using panel data, they delve deeper into this pattern and find that women's promotions occur more slowly than men's. This timing difference could also explain the gaps we see. Over a ten-year period, if men are promoted in year seven and women are promoted in year nine, women have not been in their tenured position as long, so have not accumulated the same earnings growth as men.

We use our data on newly minted US-based PhD economists to look at these patterns by gender for the sectors where we see the biggest differences in earnings gaps and largest employment shares: academia and industry. Looking at career ladder earnings does not provide a clear story on why the gender gap grows so much in academia relative to industry. Women are more likely to stay with their initial employer, which is associated with lower earnings, and they receive even lower earnings than men when they stay with their initial employer, but these differences are not large. For example, the annualized earnings change in the first ten years of a woman economist's career when staying with the initial academic employer is 4 percent, while men's is 5 percent.

Both men and women economists benefit from moving from their initial employer, both within academia and to other sectors, but men benefit slightly more for every type of move. For example, men who move out of academia earn 10 percent more per year than their initial earnings, whereas women only earn 6 percent more. As a result, women moving from academia to industry have earnings at year ten that are only 75 percent of the earnings level of similar men ($146,000 versus $195,800). Moreover, women who move back into academia from industry experience less wage growth than men who make that transition. Male economists who move from industry have earnings growth of around 3 percent based on their previous earnings in industry, whereas women only experience 1 percent earnings growth over ten years if they move from industry to academia.

Ginther and Kahn (2021) use data from the research firm Academic Analytics and also find the gender promotion gap is small in research-intensive universities, but larger in less research-intensive universities. However, they note that their measure of promotion includes promotions that occur when switching academic institutions, so this unusual finding might reflect that fact. In our data, when we disaggregate between movements between academic departments, focusing on top-70 ranked departments, we find small gender gaps in earnings for both those who remain in their initial placement employer and those that change employers. The exception is for women who move to higher-ranked departments, where women earn a small premium of 7 percent annualized growth versus men's 5 percent.

Overall, the gender gap in women's earnings in academic economics seems to come from a combination of initial earnings gaps and lower earnings growth. Although the earnings growth differences between men and women are not large, they persist across paths. Women who stay with their initial employer, move to a

similarly ranked employer, or move to a lower ranked employer all have lower earnings growth than men following similar paths. Only women who move to a higher-ranked employer have higher earnings growth than men, and movements up academic rankings are relatively rare.

References

Athey, Susan, and Michael Luca. 2019. "Economists (and Economics) in Tech Companies." *Journal of Economic Perspectives* 33 (1): 209–30.

Bedard, Kelly, Maxine Lee, and Heather Royer. 2021. "Using Longitudinal Data to Explore the Gender Gap for Academic Economists." *AEA Papers and Proceedings* 111: 69–73.

Committee on the Status of Women in the Economics Profession (**CSWEP**). 2018. "2017 CSWEP Annual Report." *CSWEP Reports* 1: 1–16.

Del Rossi, Alison F., and Joni Hersch. 2020. "Gender and the Consulting Academic Economist." *Economic Inquiry* 58 (3): 1200–1216.

Foster, Lucia, Erika McEntarfer, and Danielle H. Sandler. 2023. "Diversity and Equity in Labor Market Outcomes for Economists." *Journal of Economics, Race, and Policy* https://doi.org/10.1007/s41996-023-00117-z.

Foster, Lucia, Erika McEntarfer, and Danielle H. Sandler. 2023. "Replication data for: Early Career Paths of Economists Inside and Outside of Academia." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E193216V1.

Fourcade, Marion, Etienne Ollion, and Yann Algan. 2015. "The Superiority of Economists." *Journal of Economic Perspectives* 29 (1): 89–114.

Freeman, Richard B. 1999. "It's Better Being an Economist (But Don't Tell Anyone)." *Journal of Economic Perspectives* 13 (3): 139–45.

Ginther, Donna K., and Shulamit Kahn. 2004. "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" *Journal of Economic Perspectives* 18 (3): 193–214.

Ginther, Donna K., and Shulamit Kahn. 2021. "Women in Academic Economics: Have We Made Progress?" *AEA Papers and Proceedings* 111: 138–42.

Hoover, Gary A., and Ebonya Washington. 2021. "Report: Committee on the Status of Minority Groups in the Economics Profession (CSMGEP)." *AEA Papers and Proceedings* 111: 764–79.

Internal Revenue Service (IRS). 2021. "W2 record data." https://www.irs.gov/forms-pubs/about-form-w-2.

Jones, Todd, and Arielle Sloan. 2021. "The Academic Origins of Economics Faculty." IZA Discussion Paper 14965.

Kleemans, Marieke, and Rebecca L. Thornton. 2021. "Who Belongs? The Determinants of Selective Membership into the National Bureau of Economic Research." *American Economic Association Papers and Proceedings* 111: 117–22.

Lundberg, Shelly, and Jenna Stearns. 2019. "Women in Economics: Stalled Progress." *Journal of Economic Perspectives* 33 (1): 3–22.

National Center for Science and Engineering Statistics (NCSES). 2017. "Survey of Earned Doctorates—microdata." https://www.nsf.gov/statistics/srvydoctorates/.

National Center for Science and Engineering Statistics (NCSES). 2020, 2018, 2012, 2010, 2008. "Survey of Earned Doctorates." https://www.nsf.gov/statistics/srvydoctorates/ (accessed on December 1, 2020; December 4, 2018; April 15, 2013; November 29, 2010; and November 18, 2009).

Research Papers in Economics (RePEc). "Top US Economics Departments." https://ideas.repec.org/top/top.usecondept.html (accessed on September 15, 2021).

Scott, Charles E., and John J. Siegfried. 2008. "American Economic Association Universal Academic Questionnaire Summary Statistics." *American Economic Review* 98 (2): 630–33.

Scott, Charles E., and John J. Siegfried. 2018. "American Economic Association Universal Academic Questionnaire Summary Statistics." *AEA Papers and Proceedings* 108: 616–18.

Siegfried, John J. 2002. "Report of the Director: Job Openings for Economists." *American Economic Review* 92 (2): 511–13.

Siegfried, John J. 2016. "Report of the Director: Job Openings for Economists." *American Economic Review* 106 (5): 740–42.

Siegfried, John J. 2018. "Report: Job Openings for Economists." *AEA Papers and Proceedings* 108: 694–96.

Siegfried, John J. 2022. "Report: Job Openings for Economists." *AEA Papers and Proceedings* 112: 737–39.

Stock, Wendy A., and Richard M. Alston. 2000. "Effect of Graduate-Program Rank on Success in the Job Market." *Journal of Economic Education* 31 (4): 389–401.

US Census Bureau. 2021. "Longitudinal-Employer Household Dynamics (LEHD) microdata." https://lehd.ces.census.gov/data/lehd-snapshot-doc/latest/.

Weyl, E. Glen. 2017. "Finance and the Common Good." In *After the Flood: How the Great Recession Changed Economic Thought,* edited by Edward L. Glaeser, Tano Santos, and E. Glen Weyl, 277–301. Chicago: University of Chicago Press.

Retrospectives: Margaret Reid, Chicago, and Permanent Income

Evelyn L. Forget

When Margaret Gilpin Reid (1896–1990) was named a Distinguished Fellow by the American Economic Association, the citation published in the September 1980 issue of the *American Economic Review* recognized the impact of her work on "generations of students and colleagues."

Margaret Reid was one of the pioneers in several areas of research on consumer and household behavior, each of which has now burgeoned into a major field of study on its own. For example, she did some of the earliest work on the concept and measurement of permanent income. Again, she was one of the first to see that one could systematically study the economics of the household use of time. And, of course, she has been a major contributor to the statistical analysis of the demand for housing. The empirical tradition at the University of Chicago owes much to Margaret Reid's example and teaching. She was a famous taskmaster in the art of applying critical thinking to data.

■ *Evelyn L. Forget is Professor of Economics and Community Health Sciences, University of Manitoba, Winnipeg, Manitoba. Her email address is Evelyn.forget@umanitoba.ca.*

Reid worked at a key point in the history of economics in the mid-twentieth century, when opportunities were beginning to open for American women in academia and government research, but social roles and norms had not yet caught up with the changes underway. Like many of her contemporaries, Reid would never have entered university at all, let alone have developed an expertise in consumer economics, if agricultural colleges had not been spurred to offer training in home economics to mostly female high school teachers. That opportunity allowed her to transition from working as a rural school teacher to attending the Manitoba Agricultural College and then completing a PhD at the University of Chicago.

Reid's professional career can be roughly divided into three phases, all linked by the theoretical concept of income, the importance of applied statistical work, and the related significance of historical and institutional context in economics. From 1930 until 1943, she was employed at Iowa State College, primarily teaching home economics. In 1943, she began the first of two leaves in Washington, DC, that brought her into contact with Dorothy Brady and, through Brady, Rose and Milton Friedman. In 1951, with the active support of Friedman and Theodore Schultz (who had been a colleague at Iowa State), she joined the University of Chicago as Professor of Economics and Home Economics, where she remained until her mandatory retirement in 1961. After her retirement, she continued as an active researcher, participating in, among other things, the workshop run by Gary Becker on "New Home Economics." She also spent years struggling to write a never-finished book on income and health outcomes, a task that foreshadowed some of the social epidemiology that would pre-occupy other social scientists, beginning especially in the 1970s, under the label of the "social determinants of health."

From Home Economics to Consumer Economics

Margaret Reid was born on a small homestead near Cardale, Manitoba, in 1896. She was a young woman teaching at small rural schoolhouses when policy decisions by the provincial government created a unique opportunity. Large-scale immigration from Central and Eastern Europe was transforming the North American Midwest in the early years of the twentieth century, and immigrants brought with them diets and customs of household management very different from those of the dominant culture. In response, local governments began to mandate home economics classes in high schools and to press agricultural colleges to offer a specialization in home economics for schoolteachers in order to train them to teach the next generation of homemakers.[1]

Home economics was seen by government funders and agricultural college administrators as one of the helping professions, similar to social work and nursing, and an appropriate field of study for the sisters of the young men who made their way to these same colleges to study agriculture. Home economics was, by nature

---

[1] For biographical data, see Forget (2023, 2010, 1996) and Yi (1996). A full bibliography of Reid's contributions is in Forget (2023).

and intent, an applied and interdisciplinary field. As taught in the school system, it focused on household management, childcare, and consumer safety, especially in food preparation and household production. Much of the research tended to focus on the ways that social institutions, including legislation, might help protect families from unscrupulous merchants and the risks associated with agricultural production. However, the field was also beginning to rely on the masses of data emerging from agricultural field stations and surveys about household time use, diets, housing, family health, incomes, and expenditure. For someone with an empirical bent, this data was creating new opportunities to examine how consumer expenditure might respond to changing prices and incomes.

Reid entered the Manitoba Agricultural College in 1916 as a home economics student and, instead of returning to teach high school when she graduated in 1921, she went on to the University of Chicago. Chicago had a tradition of hiring a woman to jointly teach economics and home economics, and Reid studied consumer economics with Hazel Kyrk, who was then Professor of Economics and Home Economics. Reid submitted a dissertation in 1931 that was expanded and published three years later as *Economics of Household Production.*

Reid's path into economics by way of home economics influenced her later work. It was driven by a personal and intimate understanding of rural poverty and its effects on the aspirations and achievements of people over a lifetime. She believed that economics can, and should, encourage efforts to allow families to improve their circumstances, that education should focus on increasing the capacities of household decision-makers, that there should be institutions to protect workers and consumers from fraud, and that public institutions had a role to play in mitigating risk. She also believed that detailed empirical analysis of consumer behavior was a central concern.

In 1930, Reid began to work alongside Elizabeth Hoyt at Iowa State College, lecturing in both economics and home economics. In her research, Reid operated primarily in the area of home economics focusing on decision-making and production within the household, a topic on which Gary Becker would focus attention in the economics profession several decades later. Even at this early stage of her work, she was grappling with the concept of "permanent income," although she had neither the statistical nor theoretical framing for the concept. Instead, she presented her analysis as measuring and enhancing the financial well-being of farm families. Central to this research was a recognition that the consumption patterns of farm families aligned poorly with their reported incomes, both because of the significance of household production among rural families and because reported market income varied dramatically from year to year due to factors outside the control of households.

In the book growing from her dissertation, Reid (1934, v) argued that both the economic contributions made by women and the decisions that women made about whether to enter the paid labor market could only be appreciated if household production and its associated costs were well understood. In particular, researchers who gathered data on labor market participation and income from farm families were liable to underestimate significantly both family well-being and the work effort of women if they neglected household production.

Reid acknowledged the difficulty of distinguishing between "work" and "leisure" in any attempt to measure household production, noting that the same activity often had elements of both. Therefore, she identified household production as any activity that could, in principle, be replaced by a market alternative. She developed four different ways of measuring household production, each with its strengths and weaknesses: opportunity cost, hired worker cost, retail price, and boarding service cost (Reid 1934: 160–69).

Opportunity cost measured the value of household production by the market wage foregone. While it was a useful way to approach decisions an individual woman might make about whether to enter the paid labor market, it suggested that the value of household production varied depending on the labor market opportunities of the producer, rather than the value of the good or service itself.

The retail price method was used in *Food for People* (Reid 1943a: 134–136). She estimated the value-added by household production by deducting the cost of purchased inputs from the price of market substitutes for goods and services consumed. For example, the value of the labor involved in sewing clothing at home could be estimated by subtracting the cost of the fabric from the price that would have been paid had the finished garments been purchased at the general store. Similarly, the labor involved in cooking a meal could be valued by subtracting the cost of purchased groceries from the value of the same meal at a restaurant. In those cases where goods and services were available through the market, this method worked reasonably well. However, in cases where markets were imperfect or altogether nonexistent, which happened often enough in the rural midwest of that period, the method failed. Goods and services were often produced at home because there was no retail alternative.

She noted that hired workers and boarding houses can often provide substitutes for goods and services produced within a household, and these methods have subsequently been used in court cases for wrongful deaths. For example, one could hire a cook or a childminder. However, both approaches are limited by an inability to guarantee that the goods and services produced by hired workers or boarding houses are of the same quality as those that would be produced by a household member.

Reid (1934, 93–117) ultimately concluded that household production depends on market income, geography, education, tastes, race, and stage of the life cycle. Market income and tastes primarily influenced the extent to which purchased goods and services might be seen as reasonable substitutes for those produced within the household, and the other variables influenced the opportunities available for paid employment outside the household. She also anticipated how new technologies and changing labor markets might affect household production. This style of analysis, focused so heavily on demographic and sociological factors, plants it firmly within the realm of home economics, as does her concern with the education of household decision-makers, a preoccupation that governed much of her joint work with Elizabeth Hoyt and Hazel Kyrk. The roles of advertising, labelling, credit, legal protection, and the responsibility of the state for ensuring consumer protection were examined in *Consumers and the Market* (Reid 1938). She examined

"Consumer-Business Relations" (Reid 1942) and imagined how marketing might be different in "Marketing Rewritten from the Consumers' Point of View: Discussion" (Reid 1940).

However, the nature and style of Reid's research was changing. She began to focus more of her time and energy on the empirical relationship between income and consumption, rather than the institutional aspects of consumer well-being. While some early publications and presentations revealed this growing interest, it was not until her sojourn in Washington, DC, where she began working on cost-of-living indices, that the transition was complete. But while the nature of her research changed, her orientation towards consumer education and family well-being persisted throughout her career.

From Ames to Chicago

Reid was promoted to full professor at Iowa State in 1940. However, her growing interest in consumer economics and the statistical relationship between consumption and income made her something of a poor fit among the home economists with whom she had been working.[2] At the same time, the college itself was wracked by the "oleomargarine scandal" that saw 19 researchers, including Theodore Schultz, who taught agricultural economics at Iowa State from 1930 to 1943, resign to protest pressure from the dairy industry to change their research findings (Burnett 2011).

The case involved the production of a pamphlet entitled "Putting Dairying on a War Footing," which recommended that households substitute margarine for butter so that more dairy products would be available for American troops posted overseas. Written by Oswald Brownlee (1943), it was approved by a four-person committee that included Theodore Schultz and Margaret Reid. Local dairy producers objected, and Schultz resigned just as the President of Iowa State announced a "Committee to Reorganize the Department of Economics and Sociology" that included no one from the department. Sixteen of the 26 members of the Department in 1943 were gone by 1945. Six, including Schultz, went to the University of Chicago.

The political situation in the department played a significant part in Reid's decision to take two leaves, first to work as an economist in the Division of Statistical Standards in the Executive Office of the President (1943–1944) and subsequently to head the Family Economics Division of the Department of Agriculture (1945–1948). However, Reid was also attracted by the availability of unique data, the opportunity to reform policy, and a chance to collaborate with other researchers who shared her interests.

Reid was initially attracted to Washington, DC, by the specific challenge of examining nutritional and food-expenditure standards that might inform levels of government support for low-income families, but other interests soon took

---

[2] See, for example, Reid (2010), Box 53, Folder 1, Journal circa 1943.

precedence. She met Dorothy Brady and Rose Friedman. Brady, who held a PhD in mathematics, was a home economist at the Bureau of Human Nutrition and Home Economics in the US Department of Agriculture and was raising her son as a divorced single mother (Forget 2000). Dorothy Brady was an acquaintance of Milton Friedman, who was then spending his days at the Treasury Department, when she first began to work with Rose. Rose Friedman was midway through a Chicago PhD degree with Frank Knight, which made her something of an oddity in a department where most of the female students worked with Hazel Kyrk. Brady initially encouraged Rose Friedman to work with her on income and savings in different communities, in part hoping to distract her from a devastating pregnancy loss. Burns (2022) offers a detailed and intriguing picture of the relationships that emerged.

The research undertaken by Rose Friedman and Brady made use of the data collected as part of the Consumer Purchases Study and contributed to the NBER Conference on Research in Income and Wealth, where it was first presented by Brady in 1945 and published two years later (Brady and Friedman 1947). They documented the relationship between income and expenditure for a variety of families and argued that savings and consumption depended not on absolute income, but rather on relative income. While this finding was consistent with consumer studies in home economics that were dominated by sociological factors, it took on a new importance in the context of debates over Keynesian economics. James Tobin (1951, 135-6, 152) recognized contemporaneously that their study was a key contribution to the relative income hypothesis, which could pose an important critique of the Keynesian position.

Rose Friedman's formal collaboration with Dorothy Brady lasted only six months, but the friendship persisted. Meanwhile collaboration between Brady and Reid grew. When Reid left government service and moved to the University of Illinois in 1948, she found a position for Brady to help analyze the massive amount of consumption data with which she was working.[3] At the same time, Brady brought Reid into closer collaboration with the Friedmans, and the friendship grew until she became a regular guest at their summer home in New Hampshire. In his Nobel autobiography, Friedman (n.d.) writes: "The catalyst in combining my earlier consumption work with the income analysis in professional incomes into the permanent income hypothesis was a series of fireside conversations at our summer cottage in New Hampshire with my wife and two of our friends, Dorothy S. Brady and Margaret Reid, all of whom were at the time working on consumption."

Shortly after she arrived at the University of Illinois, Brady sent Rose Friedman a paper based on their earlier collaborative work on food expenditures. Milton Friedman found the paper intriguing, but doubted that it could support her conclusions. He responded with a spirited critique and a reference to a chapter in his thesis with Simon Kuznets, in which he used the incomes of doctors and dentists to

---

[3] Margaret Reid to Kathryn V. Burns, Acting Head, Home Economics Department, January 19, 1949, Reid (2010), Folder 15 Box 1. Cited by Burns (2022).

distinguish between permanent and transitory income.[4] Brady drew his criticism to the attention of Reid who was exploring similar issues in a very different population.

Reid was particularly interested in the consumption patterns of farm families. To an extent even greater than the professionals considered by Friedman, farm families had incomes that fluctuated widely due to variations in weather, crop prices, and other factors outside their control. Therefore, if incomes and consumption levels are arbitrarily divided by year, there would be little correlation between current income and consumption. By the time Reid moved to the University of Illinois, she was ready to begin a project to investigate the relationship between various forms of income, other economic resources, and expenditures among farm families, acknowledging explicitly that "a great many income concepts exist."[5]

As this work unfolded, Reid's PhD advisor Hazel Kyrk was about to retire from the University of Chicago. The retirement of the sole woman professor, who was jointly appointed with Home Economics, created an opportunity to hire a replacement. By this time, Reid had developed a strong, independent research program based on the statistical analysis of newly available consumption data, including the work she had undertaken at the Family Economics Division of the federal Department of Agriculture in Washington, DC, on food expenditure patterns of low-income households. As a consequence, when Milton Friedman and Theodore Schultz recruited her to come to the University of Chicago in 1951, she was not hired as a technician or assistant, but as a full professor—an economist with her own significant knowledge base that was both very different from, yet highly complementary to, the backgrounds of her new colleagues.

For Friedman, one motivation for recruiting Reid was his eagerness to move the department away from an emphasis on macroeconomic modelling based on Keynesian-style simultaneous equations.[6] Brady and Reid offered an alternative focus in their analysis of the consumption function. In March 1951, at Friedman's urging, Reid and Brady sent a memo to Theodore Schultz, who had become department head, outlining a proposed program of consumption research.[7] It noted that consumption had moved to center stage in economic research because of the Keynesian revolution, but "the early efforts to discover the consumption function through the use of family data were carried on almost as if nothing had previously

---

[4] Friedman to Brady, November 8, 1948, Box 21, Folder 18, Friedman papers, "Brady, Dorothy S"; cited by Burns (2022).

[5] "The Consumption Function in Farm Families", Reid (2010), Box 1, Folder 15.

[6] This shift also involved the move of the Cowles Commission from the University of Chicago to Yale University in 1955. Dimand and Rivot (2021) document the methodological contrast between the Cowles Commission, which favored macroeconomic modelling based on maximum-likelihood methods to estimate Keynesian simultaneous-equations models, and the distinct approach associated with the National Bureau of Economic Research and associated with Arthur Burns and Wesley Mitchell. Friedman, a student of Burns and Mitchell, was quite critical of the Cowles approach. The confrontations and exchanges between the two approaches no doubt contributed to the departure of Cowles from Chicago for Yale.

[7] To T. W. Schultz from Margaret G. Reid and Dorothy S. Brady, memorandum, March 12, 1951, subject: proposed research project on the consumption and savings of families, 4. Box 21, Folder 18, Friedman papers, "Brady, Dorothy S". Cited by Jennifer Burns (2022), who notes that "Friedman later rewrote this memo, so we know this version was written without his input."

been done in this field" and, as a consequence, much of it was "spotty" and "shoddy." They did not mention that most of the ignored research had been done in the discipline of home economics by women.

Theodore Schultz was well-acquainted with Margaret Reid from their time at the University of Iowa and was very supportive (Forget 2023). Burns (2022) cites the minutes of a department meeting in May 1951, in which Schultz announced that the retirement of Kyrk had opened the possibility of a new appointment, joint with Home Economics, in consumption economics along with the introduction of a full-scale workshop to be directed by Reid and Brady. While Reid was hired with unanimous support, the department was less receptive to the expansion of consumer economics as a field. In a subsequent meeting, Schultz pressed for a second motion recognizing that the appointment of Reid should be seen as a first move in the establishment of a research enterprise that would include hiring Brady and a full-scale workshop in consumption.[8] The workshop was never established, and Dorothy Brady returned to government.

The Genesis of the Permanent Income Hypothesis Theory

As these events unfolded, the permanent income hypothesis was emerging in the context of the friendship between Reid, Brady, and the Friedmans. Over the summers of 1949 and 1950, Friedman, Brady, and Reid explored the relationships between windfall and permanent incomes, savings, and consumption at the Friedmans' summer home. In 1951, Milton Friedman sent Brady, Schultz, and Reid a document outlining "our tentative hypothesis," and Brady encouraged him to write it up more formally for publication (Burns 2022).[9] Friedman was distracted by other tasks, but returned to the project in 1953 when he created an extensive manuscript that explored the idea more fully. In a letter to Margaret Reid, he acknowledged that he had drawn heavily upon her (unspecified) paper and was quite excited by the outcome, which he promised to send her after Brady had it typed.[10]

It is unclear which specific paper by Reid that Friedman had in mind. Her early studies on income and consumption appear in the NBER publication *Studies in Income and Wealth*. A 1952 paper "Effect of Income Concept upon Expenditure Curves of Farm Families" examines the role of transitory income on farm expenditure (Forget 2023). Burns (2022) suggests that the Reid paper in question might have been "The Relation of the Within Group Transitory Component of Incomes to the Income Elasticity of Family Expenditures" which was never published, although it was presented at the 1953 AEA meetings and also referred to the hypothesis.[11] The

---

[8]Minutes, Meeting of the Department, May 23, May 28, 1951, 41.2 Shultz Department of Economics, Chicago. Cited by Burns (2022).

[9]Milton Friedman to Brady, June 14, 1951, Friedman Papers, Box 21, Folder 18. Brady to Milton Friedman, July 17, 1951, Friedman Papers, Box 21, Folder 18. Cited by Burns (2022).

[10]Milton Friedman to Margaret Reid, August 7, 1953, Reid (2010), Box 6, Folder 1.

[11]Margaret Reid, "The Relation of the Within Group Transitory Component of Incomes to the Income Elasticity of Family Expenditures," unpublished paper, Reid (2010), Box 10, Folders 8–9.

importance of Reid's contributions in clarifying the permanent income hypothesis in this paper was highlighted by Modigliani (1986, 299) in his Nobel lecture:

> [A] fundamental contribution was the highly imaginative analysis of Margaret Reid ["Relation of the Within-Group Transitory Component of Income to the Income Elasticity of Family Expenditures" (Reid n.d.)] which pointed to a totally different explanation for the association between the saving ratio and relative income, namely that consumption was controlled by normal or "permanent," rather than current, income. This contribution was an important source of inspiration. . . .

This particular paper may have been expanded and modified, or at least strongly influenced subsequent work Reid published with Marilyn Dunsing, first as "Effect of Variability of Incomes on Levels of Income-Expenditure Curves for Farm Families" (Reid and Dunsing 1956) in *Review of Economics and Statistics,* then in *Journal of the American Statistical Association* as "Effect of Varying Degrees of Transitory Income on Income Elasticity of Expenditures" (Dunsing and Reid 1958).

Throughout her tenure at Chicago, Reid continued to examine the impact of transitory and "normal" income, drawing inspiration from the work of Friedman and Modigliani and offering inspiration in turn. She had been working on these ideas from her earliest examination of the budgets of farm families drawn from the data collected by the Agriculture Experiment Stations and, later, from her time in Washington, DC. The theoretical framing seems to have been a collaborative production, acknowledged informally, although the formal credit was less generously allocated.

The jointly conceived work eventually appeared in *A Theory of the Consumption Function,* in which Friedman (1957, p. ix) credited both Brady and Reid, recognizing Reid's "characteristic enthusiasm, persistence and ingenuity" and acknowledged the key role she played in the development of his own work on permanent income and the consumption function. Specifically, he claimed that Reid had done the initial testing of the hypothesis and had encouraged him to write it up so that she could refer to it in a paper presenting her conclusions.

Burns (2022) notes that over the gestation of the book, Friedman alternately referred to the theory in question as "our work," "our idea," and "my hypothesis," ultimately settling for the latter formulation. She wonders why both Brady and Reid would press him to write up the idea rather than taking ownership themselves; for example, she wonders if Reid might have doubted whether she possessed the theoretical training to do the idea justice, or if this may have been an instance of a lesser-known scholar believing that the idea would go further, faster, if it came from the pen of someone better known.

Retirement: Health, Income, and the "New Home Economics"

Reid took mandatory retirement from the University of Chicago in 1961, but continued to haunt the libraries, conferences, and journals related to income,

poverty, and health outcomes. According to Claudia Goldin (2021: 26), who was a graduate student at the University of Chicago in the late 1960s and early 1970s, Reid was "the only senior female economist I encountered during my graduate career." Reid was, in Goldin's words, "one of the ancients." She "was amazed by Margaret Reid" but she had had no sense at that time that Reid's "life would help inform [her] understanding of the evolution of women's economic roles" (46).

In *Housing and Income*, Reid (1962) identified the roles played by permanent and transitory income, income effects associated with position in the life cycle, and other sociodemographic factors associated with income to demonstrate that the value of housing increases relatively more quickly than income—thereby challenging the claim that the proportion of income spent on housing declines as income increases. This work was highlighted by the AEA Distinguished Fellow citation and demonstrates the ways in which Reid's empirical work on consumption and income was animated throughout her life by her early education in home economics.

Her background in home economics also made her an active and vocal participant in Gary Becker's workshop on "New Home Economics." When Becker returned to Chicago in 1970, he continued the seminars that he had started in the early 1960s at Columbia, exploring many of the same topics and themes that preoccupied traditional home economists through the lens of rational choice theory. For example, Reid's (1934) discussions of how to value home production foreshadowed the method Becker (1965) would later use in "A Theory of the Allocation of Time," although Becker chose to treat all unpaid activities as alternatives to paid labor rather than distinguishing between consumption and production along the same lines as Reid. Indeed, Becker's work on time allocation, household production, and family often covered topics that had been discussed earlier by Reid and Kyrk, but typically did not cite that earlier work.

The irony of reconsidering many of the ideas she had pioneered in the context of Becker's workshop, after she had been denied the opportunity to conduct her own workshop in consumption economics, echoed the earlier tension Reid had experienced at Iowa State when she had been the young scholar challenging the established field of home economics with new ideas and new methods. In both cases, the topics that preoccupied home economists were being examined in ways that brought home economics into closer alignment with microeconomics, and in both cases established home economists reminded the newcomers that institutions were important and that rational choice had its limits (Reid 1977; Grossbard-Shechtman 2001).

However, Reid's primary retirement project was a detailed examination of the relationships between income and health. Her notes collect details of related sociodemographic variables, including education, income security, position in the life cycle, race, changes in healthcare technology, and changes in permanent income. In some ways, her work on health mirrored the transition of her earlier career, which took her from a deep contemplation of the contextual and sociological details of consumption that dominated home economics to the more abstract presentations of her Chicago years. She never published her planned book, but

amidst the detailed charts, datasets, graphs, and notes in Reid's papers are several drafts, including one that appears to be an outline of the larger project: "Health, Age and Income of Populations: An Introduction and Preview" (1985).[12] Reid planned to examine the roles of economic resources, conflicting evidence on death rates, disability, health and labor force participation, race, education, and medical expenditures, hearkening back to her earliest work on consumption in which the sociodemographic factors were at least as important as economic variables including income in all its forms. But in this outline, permanent and transitory income still emerged as key factors.

The correlation between income and health is a theme that would come to dominate social epidemiology in subsequent decades. On the one hand, the relationship between poverty and contagious diseases had been recognized for centuries. After all, people who live in crowded housing, who take public transit to work, and who cannot afford to stay home when they are ill are more likely to get infected and to spread disease to one another and to their families. If people cannot afford decent diets, good housing, and medical care, they are less likely to survive infections over which the more fortunate can prevail. By the 1960s, it had also become apparent that the prevalence and severity of chronic conditions were strongly associated with income, and that people from families with the lowest incomes were more likely to suffer comorbidities, to be hospitalized for conditions that adequate primary care might have prevented, and to die earlier than their neighbors with more resources.

But at the time Reid was writing, the relationships between income, health, and wealth were still poorly understood; in particular, the direction of causation was ambiguous. People who have multiple chronic conditions are likely to take more time off work and to be less productive than their healthier counterparts. Therefore, poor health is one factor responsible for low incomes. But low incomes might also be responsible for poor health.

A number of specific questions were being formulated. If poverty is the only aspect of income influencing health, then one might expect that health outcomes would be unrelated to income above a fairly low-income threshold. And yet, health and income are correlated at levels far above any poverty line. Moreover, people receiving government income assistance and low-income working people might have very similar incomes, and yet the health of the former group is much worse than the latter. In retirement, there is far less variation in reported incomes than among working-age populations, yet health disparities are greater among the aged than among those who are younger.

All of this was ripe territory for Margaret Reid, with her deep understanding of how income varies across the life cycle and how little correlation there often is between transitory and permanent income. Once again, she faced the situation that she grappled with in the 1950s: if income is arbitrarily chopped into annual reports—that is, if the focus is on transitory income—then there may well be little

---

[12] Several drafts are saved in her papers, including Reid (2010), Box 37, Folder 10–11; Box 38, Folder 1–7. Boxes 30 through 44 include drafts, data, computations and bibliographies from the early 1970s through to 1985 related to income and health.

correlation with outcomes influenced more strongly by permanent income. The subject matter Reid tackled in retirement was vast and, while she worked on it well into the 1980s, the planned book was never completed.

Narrating an Academic Career

There are many ways to tell a story. It is unclear what perspective Reid herself would have emphasized in describing her own career.

Without question, Margaret Reid entered the profession at a time when systemic issues within the discipline, higher education, and society more generally nudged her into home economics rather than economics. Whatever her innate interests and ability, she entered university because of political decisions to expand home economics. She studied with Hazel Kyrk, as did most of the women students at Chicago. Consequently, she specialized in applied consumer economics. The expansion of home economics across the nation created job opportunities for women graduates in female-dominated home economics departments and, simultaneously, made it less likely that women would be hired as professors of economics. Did the gender separation supported by the expansion of home economics allow economics to remain male-dominated longer than other social sciences (Bettinger and Long 2005; Lundberg and Stearns 2019)? Was Reid guided into empirical work that ended up supporting the ambitions and career honors of male colleagues (for discussion, see Miles and Gibson 2022)? Despite her stated desire, Reid was never given the opportunity to offer a workshop at the University of Chicago, which was where most of the real work of research and graduate education took place (as discussed in Cherrier and Saïdi 2021).

Here is an alternate perspective. Margaret Reid seized the opportunity to leave the life of a rural schoolteacher, to attend university and then to earn a PhD at the University of Chicago. As a woman in a female-dominated field of home economics, Reid had the opportunity to develop her research program without falling prey to the nonacademic expectations and male preeminence women often met in other academic departments. She was able to develop a significant and independent body of research that could never be attributed to a male coauthor and consequently when she was hired by the University of Chicago, her rank was full professor. As it turned out, Reid faced relatively few of the barriers that other women navigating the economics profession reported (Chassonery-Zaïgouche, Forget, and Singleton 2022; Forget 2000).

Did Reid receive due credit from the economics profession for her contributions? It is challenging to look at the roster of economic titans at the University of Chicago in the mid-twentieth century without asking how much of the well-deserved credit each received was nonetheless supported by the often-unacknowledged work of teammates, assistants, secretaries, and wives, among whom many women found their place and made their careers. But the economics discipline of the mid-twentieth century had not yet developed a broader understanding of how one might acknowledge team-based research (Hengel and Pythian-Adams 2022; Kumar

and Ratnavelu 2016; Hengel and Moon 2020). And yet, even in this context, the essential contributions of Margaret Reid were acknowledged. The synergy created by diverse individuals helped to bring about the efflorescence of creativity related to income at the University of Chicago during the 1950s.

References

Becker, Gary S. 1965. "A Theory of the Allocation of Time." *Economic Journal* 75 (299): 493–517.

Bettinger, Eric P., and Bridget Terry Long. 2005. "Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students." *American Economic Review* 95 (2): 152–57.

Boumans, Marcel. 2016. "Friedman and the Cowles Commission." In *Milton Friedman: Contributions to Economics and Public Policy*, edited by Robert A. Cord and J. Daniel Hammond, 585–604. Oxford, UK: Oxford University Press.

Brady, Dorothy S., and Rose D. Friedman. 1947. "Savings and the Income Distribution." In *Studies in Income and Wealth*, 247–65. New York: National Bureau of Economic Research.

Brownlee, Oswald Harvey. 1944. *Putting Dairying on a War Footing*. Ames, IA: Iowa State College Press.

Burnett, Paul. 2011. "Academic Freedom or Political Maneuvers: Theodore W. Schultz and the Oleomargarine Controversy Revisited." *Agricultural History* 85 (3): 373–97.

Burns, Jennifer. 2022. "Hidden Figures: A New History of the Permanent Income Hypothesis." *History of Political Economy* 54 (S1): 43–68.

Chassonnery-Zaïgouche, Cléo, Evelyn L. Forget, and John D. Singleton. 2022. "Women and Economics: New Historical Perspectives." *History of Political Economy* 54 (S1): 1–15.

Cherrier, Béatrice. 2011. "The Lucky Consistency of Milton Friedman's Science and Politics, 1933–1963." In *Building Chicago Economics: New Perspectives on the History of America's Most Powerful Economics Program*, edited by Robert Van Horn, Philip Mirowski, and Thomas A. Stapleford, 335–67. Cambridge, UK: Cambridge University Press.

Cherrier, Béatrice, and Aurélien Saïdi. 2021. "Back to Front: The Role of Seminars, Conferences and Workshops in the History of Economics." *Revue d'économie politique* 131 (4): 609–35.

Dimand, Robert W., and Sylvie Rivot. 2021. "From 'Science as Measurement' to 'Measurement and Theory': The Cowles Commission and Contrasting Empirical Methodologies at the University of Chicago, 1943 to 1955." *European Journal of the History of Economic Thought* 28 (6): 940–64.

Dunsing, Marilyn, and Margaret G. Reid. 1958. "Effect of Varying Degrees of Transitory Income on Income Elasticity of Expenditures." *Journal of the American Statistical Association* 53 (282): 348–59.

Forget, Evelyn L. 2023. "Margaret Gilpin Reid (1896–1991)." In *The Palgrave Companion to Chicago Economics*, edited by Robert A. Cord, 343–56. Cham, Switzerland: Springer Nature.

Forget, Evelyn. 2010. "Margaret Gilpin Reid." In *The Elgar Companion to the Chicago School of Economics*, edited by Ross B. Emmett, 315–17. Cheltenham, UK: Edward Elgar Publishing.

Forget, Evelyn L. 2000. "Dorothy Stahl Brady." In *A Biographical Dictionary of Women Economists*, edited by Robert W. Dimand, Mary Ann Dimand, and Evelyn L. Forget, 80. Cheltenham, UK: Edward Elgar Publishing.

Forget, Evelyn. 1996. "Margaret Gilpin Reid: A Manitoba Home Economist Goes to Chicago." *Feminist Economics* 2 (3): 1–16.

Friedman, Milton. 1957 *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press.

Friedman, Milton. n.d. "Milton Friedman—Biographical." Nobel Prize Outreach AB. https://www.nobelprize.org/prizes/economic-sciences/1976/friedman/biographical/ (accessed on September 5, 2023)

Friedman, Milton, and Rose D. Friedman. 1998. *Two Lucky People: Memoirs*. Chicago: University of Chicago Press.

Goldin, Claudia. 2021. *Career and Family: Women's Century-Long Journey toward Equity*. Princeton, NJ:

Princeton University Press.

Grossbard-Shechtman, Shoshana. 2001. "The New Home Economics at Colombia and Chicago." *Feminist Economics* 7 (3): 103–30.

Hengel, Erin, and Eunyoung Moon. 2020. "Gender and Equality at Top Economics Journals." Unpublished.

Hengel, Erin, and Sarah Louisa Phythian-Adams. 2022. "A Historical Portrait of Female Economists' Coauthorship Networks." *History of Political Economy* 54 (S1): 17–41.

Koopmans, Tjalling C. 1947. "Measurement without Theory." *Review of Economics and Statistics* 29 (3): 161–72.

Kumar, Sameer, and Kuru Ratnavelu. 2016. "Perceptions of Scholars in the Field of Economics on Co-authorship Associations: Evidence from an International Survey." *PLoS One* 11 (6): e0157633.

Laskaridis, Christina. 2022. "'Writing History as a Way of Life': The Life and Work of Margaret Marie Garritsen de Vries." *History of Political Economy* 54 (S1):127–57.

Lundberg, Shelly, and Jenna Stearns. 2019. "Women in Economics: Stalled Progress." *Journal of Economic Perspectives* 33 (1): 3–22.

Miles, James, and Lindsay Gibson. 2022. "Rethinking Presentism in History Education." *Theory and Research in Social Education* 50 (4): 509–29.

Modigliani, Franco. 1986. "Life Cycle, Individual Thrift, and the Wealth of Nations." *American Economic Review* 76 (3): 297–313.

Reid, Margaret G. 1934. *Economics of Household Production.* New York: John Wiley and Sons, Inc.

Reid, Margaret G. 1938. *Consumers and the Market.* New York: F. S. Crofts and Co.

Reid, Margaret G. 1940. "Marketing Rewritten from the Consumer's Point of View: Discussion." *Journal of Marketing* 4 (4): 134–37.

Reid, Margaret G. 1942. "Consumer-Business Relations." *Journal of Home Economics* 34 (9): 655–58.

Reid, Margaret G. 1943. *Food for People.* New York: John Wiley and Sons, Inc.

Reid, Margaret G. 1952. "Effect of Income Concept upon Expenditure Curves of Farm Families." In *Studies in Income and Wealth*, Vol. 15, 131–74. New York: National Bureau of Economic Research.

Reid, Margaret G. 1962. *Housing and Income.* Chicago: University of Chicago Press.

Reid, Margaret G. 1966. Review of *The Economics of Health* by Herbert E. Klarman. *Journal of Political Economy* 74 (3): 301–03.

Reid, Margaret G. 1977. "How New Is the 'New Home Economics?'" *Journal of Consumer Research* 4 (3): 181–83.

Reid, Margaret G. 2010. "Margaret G. Reid Papers 1940–1990." Hanna Holborn Gray Special Collections Research Center, University of Chicago Library. https://www.lib.uchicago.edu/e/scrc/findingaids/view.php?eadid=ICU.SPCL.REIDMG#idp148633800.

Reid, Margaret G., and Marilyn Dunsing. 1956. "Effect of Variability of Incomes on Levels of Income-Expenditure Curves for Farm Families." *Review of Economics and Statistics* 38 (1): 90–95.

Tobin, James. 1951. "Relative Income, Absolute Income, and Saving." In *Money, Trade, and Economic Growth: In Honor of John Henry Williams*, 135–36, 152. New York: MacMillan.

Yi, Yun-Ae. 1996. "Margaret G. Reid: Life and Achievements." *Feminist Economics* 2 (3): 17–36.

Recommendations for Further Reading

Timothy Taylor

This section will list readings that may be especially useful to teachers of undergraduate economics, as well as other articles that are of broader cultural interest. In general, with occasional exceptions, the articles chosen will be expository or integrative and not focus on original research. If you write or read an appropriate article, please send a copy of the article (and possibly a few sentences describing it) to Timothy Taylor, preferably by e-mail at <taylort@macalester.edu>, or c/o Journal of Economic Perspectives, Macalester College, 1600 Grand Ave., Saint Paul, MN 55105.

Smorgasbord

The *World Development Report 2023* discusses the issues and evidence about "Migrants, Refugees, and Societies" (World Bank, https://www.worldbank.org/en/publication/wdr2023). "[T]here are globally about 184 million migrants (about 2.3 percent of the world's population)—37 million of them refugees . . . The share of migrants in the global population has remained relatively stable since 1960. However, this apparent stability is misleading because demographic growth has been uneven across the world. Global migration increased more than three times faster than population growth in high-income countries and only half as fast as population growth in low-income countries. . . . There is no "pre-migration" harmony to return to. In every society, tensions, competition, and cooperation have

■ *Timothy Taylor is Managing Editor, Journal of Economic Perspectives, based at Macalester College, Saint Paul, Minnesota.*

always existed across a variety of groups that are partly overlapping and constantly changing. Some of these tensions reflect socioeconomic divides: they are not about migration but about poverty and economic opportunity—and large numbers of migrants happen to be poor. Because many of those who moved or their descendants have been naturalized, some of the cultural issues attributed to migration are, in fact, about the inclusion of national minorities. Migration is also just one of many forces transforming societies in an age of rapid change, alongside modernization, secularization, technological progress, shifts in gender roles and family structures, and the emergence of new norms and values, among other trends."

Nick Chater and George Loewenstein discuss "The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray" (*Behavioral and Brain Sciences*, published online September 5, 2022, https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/iframe-and-the-sframe-how-focusing-on-individuallevel-solutions-has-led-behavioral-public-policy-astray/A799C9C57F388A712BE5A8D34D5229A1). "An influential line of thinking in behavioral science, to which the two authors have long subscribed, is that many of society's most pressing problems can be addressed cheaply and effectively at the level of the individual, without modifying the system in which the individual operates. We now believe this was a mistake, along with, we suspect, many colleagues in both the academic and policy communities. Results from such interventions have been disappointingly modest. But more importantly, they have guided many (though by no means all) behavioral scientists to frame policy problems in individual, not systemic, terms: To adopt what we call the 'i-frame,' rather than the 's-frame.' The difference may be more consequential than i-frame advocates have realized, by deflecting attention and support away from s-frame policies. Indeed, highlighting the i-frame is a long-established objective of corporate opponents of concerted systemic action such as regulation and taxation. We illustrate our argument briefly for six policy problems, and in depth with the examples of climate change, obesity, retirement savings, and pollution from plastic waste."

At this year's annual meeting conducted at Jackson Hole by the Federal Reserve Bank of Kansas City, with the theme of "Structural Shifts in the Global Economy," Charles I. Jones presented a paper on "The Outlook for Long-Term Economic Growth" (August 2023, for full proceedings of the three-day symposium, see https://www.kansascityfed.org/research/jackson-hole-economic-symposium/jackson-hole-economic-policy-symposium-structural-shifts-in-the-global-economy/). Jones writes: "What are the prospects for economic growth in the United States and other advanced countries over the next several decades? U.S. growth for the past 150 years has been surprisingly stable at 2 percent per year. Growth theory reveals that in the long run, growth in living standards is determined by growth in the worldwide number of people searching for ideas. At the same time, a growth accounting exercise for the United States since the 1950s suggests that many other factors have temporarily contributed to growth, including rising educational attainment and a rising investment rate in ideas. But these forces are inherently temporary, implying that growth rates could slow in the future. This prediction is

reinforced by declining population growth rates throughout the world. In contrast, other forces could potentially sustain or even increase growth rates. The emergence of countries such as China and India provides large numbers of people who could search for ideas. Improvements in the allocation of talent—for example, the rise of women inventors—and increased automation through artificial intelligence are other potential tailwinds."

Viral V. Acharya, Mathew P. Richardson, Kermit L. Schoenholtz, and Bruce Tuckman have edited a collection of ten short and readable essays on *SVB and Beyond: The Banking Stress of 2023* (CEPR Press, August 2023, https://cepr.org/publications/books-and-reports/svb-and-beyond-banking-stress-2023). As one example, Bruce Tuckman contributes "Silicon Valley Bank: Failures in 'Detective' and 'Punitive' Supervision Far Outweighed The 2019 Tailoring of Preventive Supervision." Tuckman writes: "Supervision clearly failed to avert the failure of Silicon Valley Bank (SVB). 'Supervision' includes a broad range of regulatory actions. . . . [I]t is useful to divide supervision into 'preventive,' 'detective,' and 'punitive.' 'Preventive' supervision refers to imposing specific rules across all banks or across particular subcategories of banks, e.g., capital and liquidity ratios, supervisory stress tests, and standards of governance, controls, and risk management. Preventive supervision aims to remove from managerial discretion swaths of behaviour that are deemed inconsistent with the safety and soundness of individual banks and with the safety of the financial system. 'Detective' supervision refers to scrutinising individual banks not only for compliance with the rules of preventive supervision, but also for behavior that—while not explicitly violating preventive rules—is inconsistent with safety and soundness or with systemic stability. . . . [D]etective supervision is a crucial part of the supervisory toolkit because fixed, preventive rules cannot foresee and anticipate all sources of risk and bank business plans. In fact, systemic risks have often arisen precisely from business plans that have migrated toward high-return and high-risk activities that are not adequately addressed by preventive supervision. Finally, 'punitive' supervision refers to compelling banks to alter behavior in response to the findings of detective supervision . . . e.g., forcing a bank to arrange for contingent credit lines, forcing a bank to sell assets, etc. . . . All in all, then, in explaining the failure of supervision in the case of SVB, failures of detective and punitive supervision are far more significant than changes to preventive supervision . . ."

Peter J. Boettke and Rosolino A. Candela offer some thoughts "On the feasibility of technosocialism" (*Journal of Economic Behavior and Organization*, 2023 (205), pp. 44–54, https://www.sciencedirect.com/science/article/abs/pii/S0167268122004048). To launch the discussion, they quote Jack Ma, founder of Alibaba: "Over the past 100 years, we have come to believe that the market economy is the best system, but in my opinion, there will be a significant change in the next three decades, and the planned economy will become increasingly big. Why? Because with access to all kinds of data, we may be able to find the invisible hand of the market. The planned economy I am talking about is not the same as the one used by the Soviet Union or at the beginning of the founding of the People's Republic of

China. The biggest difference between the market economy and planned economy is that the former has the invisible hand of market forces. In the era of big data, the abilities of human beings in obtaining and processing data are greater than you can imagine. With the help of artificial intelligence or multiple intelligence, our perception of the world will be elevated to a new level. As such, big data will make the market smarter and make it possible to plan and predict market forces so as to allow us to finally achieve a planned economy." Boettke and Candela respond: "However, we argue that the proposal provided by technosocialism is analogous to putting old wine into an irrelevant new bottle. What seems to be a novel proposal to deliver the age-old aspiration of socialism is not fundamentally different from the market-socialist model which had been proposed by Oskar Lange and Abba Lerner in the 1930s in response to Ludwig von Mises and F.A. Hayek, both of whom had argued that economic calculation under socialism was impossible."

Jason P. Robey, Michael Massoglia, and Michael T. Light describe "A Generational Shift: Race and the Declining Lifetime Risk of Imprisonment" (*Demography*, 2023, 60:4, 977–1003, https://read.dukeupress.edu/demography/article/60/4/977/380376/A-Generational-Shift-Race-and-the-Declining). From their abstract: "This study makes three primary contributions to a fuller understanding of the contemporary landscape of incarceration in the United States. First, we assess the scope of decarceration. Between 1999 and 2019, the Black male incarceration rate dropped by 44 percent, and notable declines in Black male imprisonment were evident in all 50 states. Second, our life table analysis demonstrates marked declines in the lifetime risks of incarceration. For Black men, the lifetime risk of incarceration declined by nearly half from 1999 to 2019. We estimate that less than 1 in 5 Black men born in 2001 will be imprisoned, compared with 1 in 3 for the 1981 birth cohort. Third, decarceration has shifted the institutional experiences of young adulthood. In 2009, young Black men were much more likely to experience imprisonment than college graduation. Ten years later, this trend had reversed, with Black men more likely to graduate college than go to prison."

The National Academy of Sciences has published *Toward a 21st Century National Data Infrastructure: Mobilizing Information for the Common Good* (2023, https://nap.nationalacademies.org/catalog/26688/toward-a-21st-century-national-data-infrastructure-mobilizing-information-for-the-common-good) and *Toward a 21st Century National Data Infrastructure: Enhancing Survey Programs by Using Multiple Data Sources* (2023, https://nap.nationalacademies.org/catalog/26804/toward-a-21st-century-national-data-infrastructure-enhancing-survey-programs-by-using-multiple-data-sources). From the second report: "Much of the statistical information produced by federal statistical agencies since the 1950s—information about economic, social, and physical well-being that is essential for the functioning of modern society—has come from sample surveys. . . . At the time they were established, many sample surveys represented the only way to obtain reliable, accurate, and regularly updated information about the population and businesses of the United States. But surveys have faced a number of challenges in recent years, including decreasing response rates, increasing costs, and user demand for more

timely and more granular data and statistics. At the same time, there has been a proliferation of data from other sources, including data collected by government agencies while administering programs (administrative records), satellite and sensor data, private-sector data such as electronic health records and credit card transaction data, and massive amounts of data available on the internet. How can these new data sources be used to supplement or replace some of the information currently collected on surveys, and to provide new frontiers for producing information and statistics to benefit American society?"

More on After-Effects of the Pandemic

This issue includes an article by Rebecca Jack and Emily Oster on "COVID-19, Schools Closures, and Outcomes." Santiago Pinto provides a complementary discussion in "The Pandemic's Effects on Children's Education" (*Economic Brief*: Federal Reserve Bank of Richmond, August 2023, #23–29, https://www.richmondfed.org/publications/research/economic_brief/2023/eb_23-29). "School closures and switches to hybrid/virtual learning due to the pandemic adversely affected student achievement through several channels, including a decline in skill accumulation and a disruption of peer effects and peer-group formation. Preliminary evidence suggests that losses took place early in the pandemic and that there has not been an apparent recovery. Also, the impact on students has been far from uniform, as economic losses tend to fall more deeply on younger students and students from disadvantaged backgrounds. Simply returning schools and instructional practices to where they were prior to 2019 will not avoid such losses. A wide range of remediation policies has been suggested, and evidence suggests that instruction practices—such as tutoring and individualized/small group instruction—appear to be effective."

In this issue, José María Barrero, Nicholas Bloom, and Steven J. Davis discuss "The Evolution of Work from Home." Tom Doolittle and Arthur Fliegelman of the Office of Financial Research discuss some implications in "Work-from-Home and the Future Consolidation of the U.S. Commercial Real Estate Office Sector: The Decline of Regional Malls May Provide Insight" (OFR Brief Series 23-03, August 24, 2023, https://www.financialresearch.gov/briefs/2023/08/24/work-from-home-and-commercial-real-estate/). "The work-from-home phenomenon is setting conditions for the consolidation of the U.S. commercial real estate (CRE) office sector. . . . In addition, indications that actual office occupancy by workers remains at or below 50 percent signal that employers lease significantly more office space than they currently need. Should firms reduce their office space requirements to reflect the reality of employees' work-from-home preference, the CRE office sector could suffer a contraction, posing a risk to (1) financial institutions with exposure to the sector and (2) municipalities reliant on CRE tax revenue. In fact, a diminished CRE office sector, recently valued at $3.2 trillion, could suffer a significant devaluation over time. That would generate significant financial instability through loan defaults, foreclosures, and equity value depletion. To assess the likelihood and

potential extent of a CRE office sector consolidation, this brief examines another CRE sector that has suffered decline and restructuring due to changes in user preferences: regional malls. Once a ubiquitous fixture of American life, regional malls in the U.S. have declined in number by one-quarter, and no new regional malls have been built in nine years. Additionally, this brief analyzes the timing and recognition of financial losses in the office sector that would likely occur should work-from-home become permanent and marginal office properties become structurally vacant and require repurposing."

The US Government Accountability Office has published "Federal Real Property: Preliminary Results Show Federal Buildings Remain Underutilized Due to Longstanding Challenges and Increased Telework" (July 13, 2023, https://www.gao.gov/products/gao-23-106200). "Our review of three selected weeks during January, February, and March 2023 found that 17 of the 24 federal agencies used on average an estimated 25 percent or less of the capacity of their headquarters buildings. On the higher range, agencies used an estimated 39 to 49 percent of the capacity of their headquarters on average. . . . All 24 agencies said that their in-office workforce has not returned to pre-pandemic levels due to increased use of telework and remote work."

Interviews with Economists

Steven Levitt interviews Robert Solow in in "Ninety-Eight Years of Economic Wisdom" (Freakonomics website, June 23, 2023, audio and transcript, https://freakonomics.com/podcast/ninety-eight-years-of-economic-wisdom/). "[T]here are, however, a lot of people, in the profession and outside the profession, who think that a modern, industrial, capitalist economy cannot exist without growing. . . . So, I want to imagine an economy like ours and think about what it would be like if it were stationary, if it were not growing and not shrinking, but just fixed at whatever size we're talking about. The first thing that would have to be true is that the population is constant. Now, I want to make another assumption, imagine that there's no innovation going on. There are no new products, no new industries, nothing like that. The economy is just stationary. It just repeats itself. . . . I think the important thing to realize is that there is no law of economics, no principles of economics, that say that such an economy could not exist and be healthy. It's not written anywhere that for a capitalist economy, it's grow or die. That's just not true. . . .Now I come to the rub that I don't think most people think about: this non-growing economy has, as I said, no new industries, no new products, nothing like that. That can't be good for social mobility. What I'm afraid of is that in such an economy, the same good jobs and high status occupations would repeat themselves year after year. And the people who have those jobs would groom their children to follow in their footsteps. And that kind of society would tend to be a hereditary oligarchy. And that's not good. So if I were trying to bring about—for the sake of warding off climate change, for the sake of

preserving the environment—a non-growth economy, what I would be thinking about is how you provide for social mobility, how you provide for the children of relatively poor parents to become relatively better off while some of the children of relatively well-off parents fall in the income distribution. That's the hard part. There's nothing in my background to make me a specialist in how to do that, but I can see that it would be a really serious problem."

William Kearney interviews Ben Bernanke in "Real Policymaking Involves a Lot of Other Things Besides Pure Technical Analysis" (*Issues in Science and Technology*, Summer 2023, https://issues.org/ben-bernanke-interview/). "[T]he really big steps that are needed to avert climate change—such as developing new energy technologies and retrofitting old buildings and creating new infrastructure for electric vehicles—all those things are the province of the private sector or more likely the government. And by government, I mean broadly, like Congress. I think the Fed properly should focus most of its attention on its mandate, on the objectives given to it by Congress, which are full employment and price stability. I think inequality is a similar issue in its complexity. The Fed is paying more attention to inequality and is monitoring unemployment rates across different groups. . . . But again, the Fed really only has one instrument—namely, financial conditions being tighter or easier and then promoting or slowing economic growth—and it can't use that one instrument to achieve many different objectives at the same time. It can't ease policy for one group and tighten policy for another group. It has to have the same policy for everyone in the country. This is not to deny that inequality and climate change are first-order, very important issues politically and socially, but the Federal Reserve is just one agency, and it should focus primarily on the goals that Congress sets forth for it and the tools it has to achieve those goals."

Chad Bown interviews Lee Branstetter about recent research on the topic "Is China's industrial policy working?" (Trade Talks podcast, April 23, 2023, audio and transcript available, https://tradetalkspodcast.com/podcast/182-is-chinas-industrial-policy-working/). "We find that the Chinese government is not giving subsidies to initially more productive firms. If anything, the statistical association is actually negative. The Chinese government is, on average, giving more subsidies to less productive firms. . . . Chinese firm's annual reports do often include language that describes what particular subsidies were for. But if we focus on that subset of subsidies that are meant to promote research and development, or the subset of subsidies that are meant to support upgrading of equipment, even for these specific subsidies, we find no relationship with productivity. It's not the case that firms that are more productive are more likely to receive these subsidies in the first instance. And it's not the case that firms that receive these subsidies become more productive later. . . . As we dug into this data, it became increasingly clear to us that the subsidies provided to Chinese firms had lots of objectives, many of which were not connected to productivity. We see significant quantities of subsidies going into declining industries like mining. We see significant subsidies that appear to be designed to support employment in large firms . . ."

Discussion Starters

Donald Shoup explores the idea of "Parking Benefit Districts" (*Journal of Planning Education and Research*, published online March 2023, https://journals.sagepub.com/doi/pdf/10.1177/0739456X221141317). "Transportation planners have neglected curb parking because nothing is moving, and land-use planners have neglected it because it is in the roadway. No one seems to know how to solve the curb parking problem, except for followers of Nobel laureate William Vickrey who proposed that cities should set the prices for curb spaces to "keep the amount of parking down sufficiently so there will almost always be space available for those willing to pay the fee" (Vickrey 1954). Prices can vary by place and time of day to leave one or two open curb spaces on every block. Where all but one or two curb spaces on a block are occupied, the parking is both well used and readily available. . . .[S]treets are a city's blood vessels, and overcrowded free curb parking is like plaque on the vessel walls, leading to a stroke. Market prices for curb parking prevent this urban plaque. . . . The goal is not to persuade drivers they should pay for curb parking. The goal is to convince stakeholders they should charge for curb parking. Anyone who does not store a car on the street may begin to see free curb parking the way landlords see rent control. Free curb parking is rent control, for cars. If people want better public services more than they want free curb parking, the curb lane can benefit everyone, not just drivers who store their cars on the street. . . . Because New York does not charge drivers for parking in 97 percent of its three million curb spaces, it offers a titanic subsidy for cars. If the city charged only $5.50 per curb space per day, it would earn $6 billion a year, about the same as the $6.1 billion farebox revenue from all New York City public transit in 2019. . ."

Scott Lincicome, Gabriella Beaumont-Smith, and Alfredo Carrillo Obregon discuss "Formula for a Crisis: Protectionism and Supply Chain Resiliency—the Infant Formula Case Study" (Cato Briefing Paper No. 146, January 11, 2023, https://www.cato.org/briefing-paper/formula-crisis). "Given the pandemic's intense, unpredictable, and heterogenous effects on supply and demand patterns in the U.S. and abroad, supply chain problems (delays, shortages, gluts, etc.) inevitably arose for numerous products. However, baby formula was unique in both its magnitude and duration, as well as its being isolated to the United States and avoided in the rest of the developed world and neighboring countries (save a brief period in Canada when Americans traveled there for emergency supplies). . . . The barriers to competition in the U.S. infant formula market, including tariff and nontariff barriers hindering foreign producers' access, the WIC program's structure, and obscure domestic policies such as FDA regulations, altogether make it harder for new entrants to compete. The crisis illustrates how these policies encourage concentration among a handful of large producers, prevent rapid adjustment to economic shocks, and require fundamental reform."

The American Economic Association

MIX
Paper from responsible sources
FSC™ C132124
FSC
www.fsc.org

The Journal of
Economic Perspectives

Fall 2023, Volume 37, Number 4

Symposia

After the Pandemic

Wealth

The Electricity Grid

Economics Career Paths

Features

AMERICAN
ECONOMIC
ASSOCIATION