The Journal of Economic Perspectives

A journal of the
American Economic Association

Fall 2024

The Journal of
Economic Perspectives

A journal of the American Economic Association

Editorial offices:
Journal of Economic Perspectives
American Economic Association Publications
2403 Sidney St., #260
Pittsburgh, PA 15203
email: jep@aeapubs.org

The *Journal of Economic Perspectives* gratefully acknowledges the support of Macalester College.
Registered in the US Patent and Trademark Office (®).

Composed by American Economic Association Publications, Pittsburgh, Pennsylvania, USA.
Printed by LSC Communications, Owensville, Missouri 65066, USA.

THE JOURNAL OF ECONOMIC PERSPECTIVES (ISSN 0895-3309), Fall 2024, Vol. 38, No. 4. The JEP is published quarterly (February, May, August, November) by the American Economic Association, 2014 Broadway, Suite 305, Nashville, TN 37203-2418. For details and further information on the AEA go to https://www.aeaweb.org/. Periodicals postage paid at Nashville, TN, and at additional mailing offices.

POSTMASTER: Send address changes to the *Journal of Economic Perspectives*, 2014 Broadway, Suite 305, Nashville, TN 37203. Printed in the USA.

The Journal of
Economic Perspectives

Contents

Volume 38 • Number 4 • Fall 2024

Symposia

Articles

Features

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.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 3–26

Export-Led Industrial Policy for Developing Countries: Is There a Way to Pick Winners?

Tristan Reed

Developing countries have benefited from openness and participation in international trade, and export growth to higher-income markets is the best-understood route to sustained poverty reduction at the national level. The twentieth-century growth experiences of Japan and the Republic of Korea coincided with their early membership in the General Agreement on Trade and Tariffs, just as the twenty-first-century experiences of China and Vietnam coincided with new membership in the World Trade Organization. More recently, growth in Africa, the second-fastest growing region after Asia, has been driven by commodity exports to the growing economies of Asia. While much has been written recently about the effects of trade on inequality within countries, the era of hyper-globalization coincided with an extraordinary decline in global inequality. Between 1980 and 2016, the incomes of the bottom 40 percent of the global income distribution more than doubled, twice as much growth as for the middle classes in high-income economies (Alvaredo et al. 2018).

For a model that seeks to explain these patterns, say that economic development occurs when firms pay a fixed cost to upgrade to a more productive technology. The more productive technology raises the wage, as did jobs in the manufacturing export sector in Asia, as did jobs in the service sector adjacent to the commodity export sector in Africa. The fixed cost captures economies of scale in production or the costs of innovation and transferring foreign technology and management expertise to local firms. In the presence of this fixed cost, a large market size is necessary for development. In Goldberg and Reed (2023a), using this model, the

■ *Tristan Reed is an Economist, Development Research Group, World Bank, Washington, DC. His email address is TReed@worldbank.org.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.3.

threshold market size is estimated to be about 325 million people, more than the population of most countries.

Governments guided by this model could—and frequently do—undertake some form of export-led industrial policy, defined as government policy designed to grow industries that sell to the international market. A government guided by this model would not target industrial policy toward import substitution in the absence of a populous middle class whose demand would allow firms to achieve scale locally. The menu of export-led industrial policy instruments is broad in scope, and includes many that could be untargeted and offered to all industries, in principle. Yet, because governments have scarce resources and because policy is tailored to some extent to the needs of beneficiaries, in practice these instruments are targeted where they are expected to have the greatest economic return. For instance, with a fixed amount to spend on road construction, a government may prioritize projects that connect an export industry to the international market, given the large corporate tax yield expected as that industry grows. With a fixed amount to spend on tertiary education, a government may prioritize programs that cultivate skills demanded by specific export industries, because workers trained in those skills are most likely to find remunerative employment. Temporary import tariff protection and other subsidies like grants and tax credits are designed to allow local firms to achieve dynamic economies of scale, but with insufficient domestic demand, those scale economies are only possible in the international market, so subsidies are targeted at industries with the most potential to become internationally competitive. Targeted industries are identified explicitly as strategic sectors in development plans, and implicitly by the special attention given them by government officials who must learn about their details to implement policy.

This essay is divided into two parts, corresponding to the What and the How of export-led industrial policy. The What section looks at five measures that have been used as a justification for targeting a certain industry or sector: international market growth, international competition, comparative advantage, technological relatedness, and competitive advantage. Certain measures, like a lack of international market growth or high competition, can deter a country from targeting an industry at all. Other measures may offer a justification for industry targeting, but receive different weights depending on how a government calibrates its preference for risk. The risk is that industrial policy fails because the targeted industry is not economically viable. Targeting industries where the country has already demonstrated a comparative or competitive advantage is relatively low risk. If a country wishes to build new advantages in areas with growing international markets, low competition, and technological relatedness, it takes on a medium-level risk. High-risk industrial policy would target industries with growing markets and low competition, but without technological relatedness. Ultimately, the discussion suggests that, as in private investment, while it is not possible to pick winners with certainty, governments and development agencies can and do build business cases for projects that benefit specific tradable industries, and reject projects supporting industries where growth appears infeasible. Is there a way to pick winners? My tentative answer is sort

of, yes—or at a minimum government can reduce its chances of picking losers. This is what successful export-led industrial policy has achieved in the past.

The How section begins by looking at how trade rules constrain developing countries from pursuing industrial policies that discriminate against foreign commercial interests. It then describes some instruments governments use for industrial policy, while still operating within the legal constraints of the international trading system; for example, governments can provide targeted industries or sectors with improved access to new customers or input suppliers, training, physical and regulatory infrastructure, or product quality certification. Some interventions along these lines have been evaluated in recent literature on firms, trade, and development (for discussion, see the surveys by Verhoogen 2023; McKenzie 2024). Potentially important institutional details of export promotion agencies are highlighted, drawing on the case of Costa Rica, a small country that has grown rapidly through exports and foreign investment, while reducing poverty and protecting the environment. Though the discussion focuses on export promotion, I expect these institutional details have external validity for industrial policy targeting production for the domestic market, especially in countries with a large middle class.

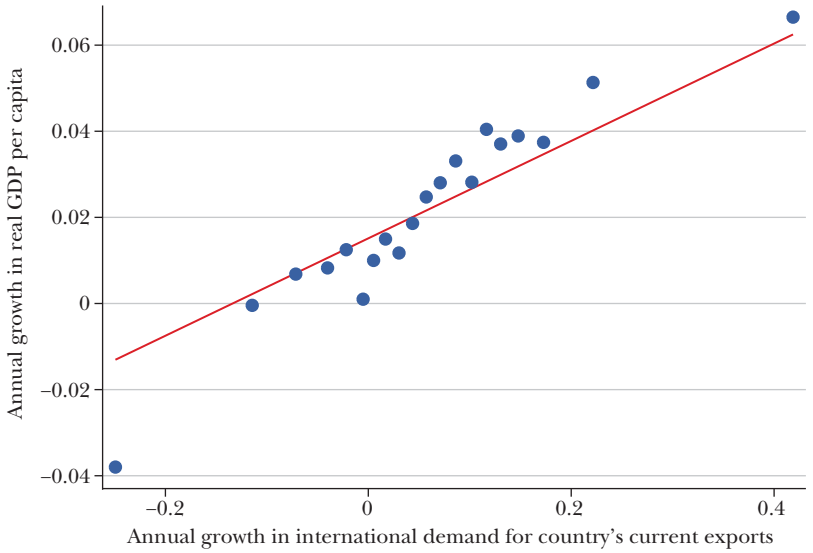
The approach to export-led industrial policy described here—choosing sectors to target based on trade data and choosing instruments that do not violate existing trade rules—is similar in spirit to what countries around the world have already adopted. As one example, it is somewhat like the Enterprise Map Project of John Sutton, which detailed the capabilities and ambitions of leading firms in several African economies, yielding messages about what government inputs could help them grow. Of course, if any industrial policy is to succeed, it requires a high-bandwidth government agency capable of doing the analysis and putting the public good ahead of rent-seeking behavior by large firms.

The What: Growth, Competition, and Productivity Potential as Guides for Industrial Policy

Within the broad category of exports, there is no shortage of theories that certain sectors are more beneficial for development than others. Manufacturing has long received focus, though that has given way to an alternative focus on tradable services. Other theories emphasize products exported by countries with a high GDP per capita (Hausmann, Hwang, and Rodrik 2007), products exported by countries with a similar GDP per capita (Lin 2011), complex products (Hidalgo and Hausmann 2009), or upstream products (Liu 2019). Some products are promoted because they are green, in the sense that they have an end-use related to decarbonization, or are produced without carbon emissions.

The drawback of these product-specific theories is they rank sectors in general, but not the market potential facing a given country. This section explores empirical measures that can target industrial policy in each country context, and illustrates how some of these measures are already being used for this purpose. The value of

Figure 1
Economic Growth and Exports to Expanding Markets



Source: Real GDP per capita at purchasing power parity is from the Penn World Tables 10.1 (Feenstra, Inklaar, and Timmer 2015). Exports and imports are from COMTRADE via the Growth Lab at Harvard University (2019).
Note: Sample is 9,123 observations of country years, including 183 countries, spanning the years 1963 to 2019. Dots shown on scatter plot are averages within 20 bins that partition the axes and include an almost equal number of observations. Growth is approximated by differences in log values between years. The slope of the best fit line is 0.11 (standard error = 0.01) and the R^2 is 0.04.

measures is not that they quantify the social return to industrial policy per se, but rather that they identify the potential of an industry to grow in a specific context, a necessary condition to realize social returns from supporting that sector. I focus on measures with a close linkage to export-led growth: international market growth, international competition, comparative advantage, technological relatedness, and competitive advantage.

International Market Growth

A useful measure for some countries has been the expected growth rate of the international market for particular products. Figure 1 shows the correlation between annual growth of GDP per capita and the annual growth rate of world imports from the country's current export basket, which measures annual growth in international demand for the country's current exports. For each country, demand growth is the weighted average growth rate of world imports of each four-digit Standard International Trade Classification (SITC, Revision 2), where the weights are the share of the country's export value in that classification. The country's own import growth is excluded from the calculation of growth in international demand,

to address some concern about reverse causality. The four-digit SITC classification, developed by the United Nations, groups goods with similar processing or manufacturing processes, and is narrow enough to cover one or several product lines in a firm. For instance, the three-digit SITC 782 classifies "motor vehicles for the transport of goods or materials and special purpose vehicles," while the four-digit SITC 7821 classifies "motor vehicles for the transport of goods or materials" and the four-digit SITC 7822 classifies "special purpose motor lorries and vans such as breakdown lorries, fire-engines, fire-escapes, road sweeper lorries, [. . .]."

Figure 1 shows a clear relationship: growing economies export to growing markets. The extremes in the upper-right-hand corner and lower-left-hand corner include commodity price booms and busts, oil exporters during the 1973/1974 oil embargo and the 2008/2009 financial crisis. The intermediate points include export-led success stories, such as Vietnam in 2010, when real GDP per capita grew at 11.5 percent and international demand for current exports grew at 15 percent. This point is almost tautological, as foreign demand enters directly into GDP as exports, but often neglected. Government support has been given historically to declining sectors (Baldwin and Robert-Nicoud 2007). Hides and skins are exports targeted by some low-income countries that today face declining demand.

The pattern in Figure 1 suggests countries seeking to grow should target entry and expansion of exports in growing markets. This goal can be challenging, of course, as growth forecasts are notoriously difficult. An observed boom in coal exports induced by the loss of natural gas supply elsewhere does not mean that coal exports offer a long-term value proposition. Nonetheless, recent industrial policy initiatives have reflected this approach, targeting for instance electric vehicles. China and Korea, which today have captured this growth through exports, also anticipated it in their industrial strategies over a decade ago. The fastest growing four-digit products between 2018 and 2021, the most recent three-year period with global import data in COMTRADE, are roasted iron pyrites (SITC 2814); hemp, raw or processed but not spun (SITC 2652); and castor oil seeds (SITC 2235), all with global demand growing above 40 percent per annum. A challenge in targeting these industries is that they are commodities. In the electric vehicle example, countries targeted an industry in which it was feasible to develop a differentiated product that is protected to some extent from competition, given heterogeneity in consumer preferences.

International Competition

The presence of tough competition suggests targeting growth or entry in an export sector is higher risk. Simple measures of the toughness of competition are the number of exporters (or the inverse of the Herfindahl index of national export shares). Atkin, Costinot, and Fukui (2021) observe that complex products are exported by a larger number of countries, raising a question about whether industrial policy targeting complex products rather than simple products can be successful. This result is a consequence of the fact that product complexity is defined based on whether the product is exported by countries that export many other products. Industrial policy targeting products exported by countries with high

GDP per capita may face the same issue, because countries with high GDP per capita export a greater variety of products (Hummels and Klenow 2005).

Of course, just because there are few competitors in a market does not mean it is easy to enter. Few competitors could reflect that the market leader has a significant cost or quality advantage, as is apparent for instance with solar panels from China. In such cases, a country considering entering that market can assess whether it can achieve similar cost and quality as the market leader, or else develop a niche with a distinct combination of cost and quality that consumers value. Many developing countries have tried to use the disruptive entry strategy: offering low cost and low quality, when only high cost and high quality is available. Today, middle-income exporters increasingly offer high quality and low cost.

Comparative Advantage

Trade theory provides a method to infer a country's current cost and quality advantage by looking at its current specialization in exports. This cost and quality advantage is often referred to as productivity, but in a way which refers to a broader concept than production efficiency (say, as measured by labor productivity or total factor productivity) that encompasses all factors allowing an exporter to capture market share. One approach countries have taken is to target industrial policy towards sectors in their economy that are relatively more productive, considered in this way.

At first blush, this approach may seem obviously wrong. One might argue that already-exporting sectors are least in need of government support. Nonetheless, market failures may still exist in these sectors, making them smaller than they could be. Moreover, in a large class of models, the welfare benefit of an increase in productivity in a sector is proportional to the size of that sector (using Hulten's theorem). If a government's industrial policy is focused on increasing productivity—some might call this approach productivism or productive development—it makes sense to check whether there are opportunities to increase productivity in the largest sectors.

Sectors that are more productive relative to some benchmark country and a benchmark sector within the economy are known as sectors of comparative advantage. Every country has some comparative advantages. For example, Costinot and Donaldson (2012) show that countries with natural inputs (such as soil characteristics, water availability, and climate) that increase yields of crops produce relatively more of those crops. The opportunity cost of market failure is greater in sectors with comparative advantage. If firms producing crops with abundant natural inputs lack access to finance—small firms, for example, which markets might fail to serve due to imperfect information—there could be returns to subsidizing finance for these firms, as it could allow them to achieve scale in the international market. In contrast, if a country were to subsidize finance in a sector without a relative productivity advantage, which is thus not involved or barely involved in global trade, the benefits seem limited. Such a sector can only grow to serve the domestic market, and perhaps only then if the domestic market is protected from imports by tariffs. The success

of industrial policy in relatively less productive sectors depends on dynamic effects (say, significant improvements in efficiency or quality, or reductions in input costs). For this reason, industrial policy in the sectors with the highest levels of comparative advantage can be considered low risk, whereas interventions in sectors with lower levels of comparative advantage can be considered high risk.

The measurement of comparative advantage from export data, which is widely practiced and rests on the assumptions of the Ricardian trade model, warrants some discussion. The most widely used measure of comparative advantage is the share of national exports in a sector divided by the share of world exports in that sector. This measure is called Balassa-revealed comparative advantage, after Balassa (1965). In 2019 in Costa Rica, for example, 16 percent of exports were business services, while for the global economy, only 2 percent of exports were business services. In this example, the statistic is greater than one ($16/2 = 8$), indicating that Costa Rica is more specialized in exports of business services compared to world exports. One infers from this that Costa Rica is relatively more productive in business services, compared to other sectors in which the country does not specialize.

Despite the widespread use of Balassa-revealed comparative advantage as a proxy for relative productivity, the measure has received significant critical scrutiny in the international trade literature; indeed, French (2017, p. 83) argues that Balassa-revealed comparative advantage is "not generally useful." What are the issues, and what do they imply for how and whether the measure is used to target policy?

Balassa-revealed comparative advantage is frequently used to make statements like "this country is more productive in product k" in an unqualified sense, perhaps implying "relative to the world" and "relative to other products." In a two-country world, where there is a home country and all other countries are grouped together as "rest of world," this statement makes sense. Industries can be ranked within a country by the average Balassa-revealed comparative advantage of their products, and those with a higher ordinal ranking are more productive relative to those with lower values.

It is more complicated to compare values of Balassa-revealed comparative advantage across countries because the definition of "rest of world" changes across countries, and data availability varies. To address these issues, an approach is to double-index revealed comparative advantage to a benchmark country (for instance, the largest trading partner, most often China and second most often the United States) and a benchmark product (for instance, live animals):

$$\textit{Revealed Comparative Advantage} \equiv \left(\frac{x^{k}_{i,World}}{x^{k}_{China,World}}\right) \Big/ \left(\frac{x^{live\ animals}_{i,World}}{x^{live\ animals}_{China,World}}\right),$$

where $x^{k}_{i,World}$ are the total exports by country i of product k.[1] A practical advantage of this measure is that it requires only the exports of country i and the benchmark

[1] One may alternatively define $x^{k}_{i,World}$ as net exports, equal to total exports minus total imports by country i of product k, to ensure that revealed comparative advantage does not identify products with minimal

country, and so can be calculated without data on total exports of product k by the world, as required by Balassa-revealed comparative advantage. Large economies like the United States or China are good benchmarks because they export most products, and report product-level export data almost immediately, whereas data on total exports of product k by the world are only available with a two- or three-year lag in the export series from COMTRADE, CEPII BACI, and the Growth Lab at Harvard University.[2] The relevant benchmark country may vary depending on the question one has. If a country is negotiating a bilateral trade agreement that lowers trade costs, indexing comparative advantage to the other country in that agreement identifies which industries are expected to grow in the context of that agreement. A country may also index to a country viewed as a local competitor, or aspirational peer. Vietnam might benchmark to higher-income Malaysia, for instance. One could conceivably pick a different benchmark depending on the product. If a country has a comparative advantage in a product relative to the leading exporter of that product, it is likely it could succeed in exporting more of that product.

Another approach argues measures of comparative advantage should control for market potential as determined by trade costs and foreign demand, which the Balassa-revealed comparative advantage measure does not. Modern trade models are designed to reflect the empirical regularity that exports follow a gravity pattern: countries export more both to markets that are closer geographically, culturally, or linked by trade agreements and to larger markets. Costinot, Donaldson, and Komunjer (2012) proposed a measure of comparative advantage that uses a regression of bilateral exports on fixed effects to isolate the component of productivity not driven by bilateral trade costs (an importer-exporter fixed effect) or foreign demand (an importer-product fixed effect). The remaining exporter-product fixed effect identifies regression-revealed comparative advantage.

While theoretically consistent, regression-revealed comparative advantage may miss certain opportunities for export success. For instance, for geopolitical reasons, the United States may increase demand for imports from Mexico as it pursues a policy of friend-shoring (on potential consequences of friend-shoring, see Goldberg and Reed 2023b). Ranking sectors within Mexico by regression-revealed comparative advantage could potentially down-rank sectors with exports concentrated in the US market, as US demand is ignored. But the Mexican government, particularly at a moment of increased US demand, might reasonably prefer to prioritize sectors that benefit from that demand. Another challenge is that regression-revealed comparative advantage can be undefined, both in small countries that export a product to only one other country; and in the case of services exports, which in developing

domestic value added. This is a concern especially in entrepôt countries, like Djibouti and Singapore. There are other approaches to identifying export value added that rely on inter-country input-output matrices, though these are available only at a high level of product aggregation.

[2] The latter two databases are most popular because they impute frequently missing export values for developing countries in COMTRADE with so-called mirror data on imports reported by other countries.

countries are often reported as flows to the world rather than as bilateral trade flows, given challenges in quantifying them through surveys.

Any measure of productivity based on realized export values could reflect economic distortions instead of actual productivity in a competitive market. There are two types of distortions: distortions in other economies and in one's own economy. Developing countries complain about subsidies for agricultural production in other countries (for recent estimates of these, see World Bank 2023). Low revealed comparative advantage in rice or cotton, for instance, could reflect not low productivity relative to other countries in a competitive market, but a country's inability to compete with subsidized exports from the United States. Still, given that a country cannot immediately change the subsidy policies of other countries—the Doha round of negotiations covering agricultural subsidies has indefinitely stalled—this quality may be an advantage of revealed comparative advantage. The measure shows what sectors are competitive, holding fixed the policies of competitors.

Distortions in one's own economy can lead to a situation in which low revealed comparative advantage could reflect market failures, but high revealed comparative advantage could reflect own-government subsidy programs without benefits net of costs that nonetheless persist for political reasons. This observation reinforces the point that if revealed comparative advantage is used to target industrial policy, that policy, like any policy, should still be disciplined by cost-benefit analysis. A modified regression-revealed comparative advantage measure that controls only for domestic subsidies, but still does not control for fixed effects outside a government's own control, can potentially identify the extent to which revealed comparative advantage depends on domestic subsidies.

Despite all these critiques, many countries do use revealed comparative advantage to target industrial policy. Looking at countries around the world, Juhász et al. (2023) show that industrial policy has been much more common in country-sectors with Balassa-revealed comparative advantage greater than 1. For instance, Senegal's national development plan argues that mining of phosphate, an input into fertilizer, is "growth and employment-generating" and sets specific targets of "strengthening of exports and positioning of Senegal as one of the leaders with, in the short term, a production of more than 3.5 million tons of lime phosphates and 800,000 tons of acid; in the medium term . . . doubling of phosphate production by 2020" (République du Sénégal 2018, note 65), while a program of domestic fertilizer subsidies has sought to increase demand for local production. For diphosphorus pentoxide, Senegal has Balassa-revealed comparative advantage equal to 343.

A benefit of targeting industries with already high productivity is that it can avoid the capture of policy by less productive firms. Despite this benefit, one might argue that industrial policy targeted towards revealed comparative advantages can reduce dynamism, as government support for one sector might divert entrepreneurs from new opportunities. This argument is less persuasive given that comparative advantage is slow-moving. Hanson, Lind, and Muendler (2015, p. 32) estimate the time path of regression-revealed comparative advantage and find "it will take 5.5 years

for half of the initial shock to log comparative advantage to dissipate and 18.4 years for 90 percent of the initial shock to dissipate." To reduce dynamism, a program of support would have to persist for longer than four or five years, the typical length of an executive's term in a democracy.

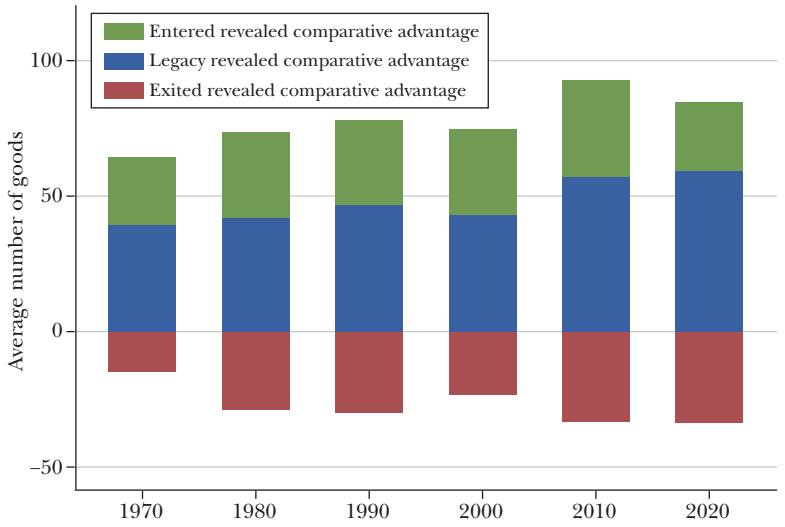
Technological Relatedness

When many people think of industrial policy, they have in mind a policy that favors industries or firms that are not yet exporting. In addition to constraining growth of currently productive sectors, market failures may stop industries from emerging at all. An industry that could be productive in the absence of market failures, which Harrison and Rodríguez-Clare (2010) call "latent comparative advantage," will not be observed in current export data. Might countries be able to predict sectors in which they are likely to gain comparative advantage in the future, based on the experience of other countries?

Hausmann and Klinger (2006) developed a country-sector level measure of product space that is used for this purpose. The product space is a matrix with rows indicating countries and columns indicating products, where the elements are either 1 if the country has a revealed comparative advantage in that product, and 0 otherwise. The authors collapse a country's position in the product space into a scalar called product space density. For a given country-product, this measure is equal to the probability that a country with similar comparative advantages as the country in question has a comparative advantage in the sector. The measure captures technological relatedness. It is a probabilistic description of what advantages a country might have, given its existing advantages. For instance, a country with a revealed comparative advantage in piston engine parts (SITC 7139) has 78 percent chance of also having a revealed comparative advantage in other parts and accessories for passenger motor vehicles (SITC 7849). In contrast, the same country has only a 4 percent chance of having a revealed comparative advantage in computers (SITC 7522). Over time, the measure is a robust predictor of which new products countries move into, controlling for other factors (Bahar et al. 2019). From this perspective, industrial policy targeted at sectors that currently do not have high productivity, but do have high technological relatedness, can be considered medium risk. An example of this approach would be Costa Rica's successful promotion of avocado and pineapple exports, which rely on similar agricultural technology as bananas, the country's long-term comparative advantage. The potential to reduce risk using product space density can explain its enduring popularity.

In comparison, high-risk industrial policy would target sectors that both have currently low productivity and low technological relatedness, as measured by product space density. Not all new advantages are predicted by product space density. In fact, GDP per capita growth is faster in countries with less path dependence, in the sense that they develop more advantages that are not predicted by product space density (Coniglio et al. 2021). An example of this pattern would be Costa Rica's jump from exporting bananas and pineapples to exporting semiconductors, medical devices, and business services. High-risk industrial policy that

Figure 2
Dynamics of Comparative Advantage in Developing Countries



Source: COMTRADE via the Growth Lab at Harvard University (2019).
Note: Sample includes all countries with less than $10,000 in GDP per capita measured in 2017 dollars at purchasing parity in a year. A good is classified as having a revealed comparative advantage if Balassa-revealed comparative advantage is greater than 1. Goods with legacy revealed comparative advantage had revealed comparative advantage in the previous decade as well as the current decade. Goods that entered revealed comparative advantage have revealed comparative advantage in the current decade but not the previous decade. Goods that exited revealed comparative advantage had revealed comparative advantage in the previous decade but not the current decade.

reaches beyond existing technological relatedness, if successful, could also have returns. However, such success is harder to predict.

The concept of product space may be especially useful for developing economies. As economies shift comparative advantage over time, higher-income countries have a better chance of producing any given product, because they have a denser preexisting product space and already know how to produce many products. Larger developing economies, like India and Indonesia, also have a preexisting pattern of producing many products. The product space is most useful when it is least dense, and it can help countries distinguishing medium-risk target industries where they have a higher degree of technological relatedness from high-risk target industries where such relatedness does not exist.

To give a sense of the balance a government could strike between targeting current and latent comparative advantage, Figure 2 shows the dynamics of Balassa-revealed comparative advantage over past decades in developing countries, defined here as those with less than $10,000 in GDP per capita measured in 2017 dollars at purchasing power parity in a year. Products are defined by the four-digit SITC, so there are about 700 possible goods in each year (service exports

are excluded, as the series are not available in early decades). In the first year of each decade, Figure 2 reports for the average country the count of goods within three categories: (1) legacy revealed comparative advantage, or goods that had an advantage in that year and also did ten years ago; (2) entered revealed comparative advantage, or goods that had an advantage in that year but did not ten years ago; and (3) exited revealed comparative advantage, or goods that do not have an advantage in that year but did so ten years ago. In any decade, roughly half of revealed comparative advantages are legacy, and the other half are recent entrants, reflecting the dynamic yet slow-moving nature of industry productivity.

Competitive Advantage

Some policymakers have sought to describe the targets for export-led industrial policy in terms of competitive advantage, rather than comparative advantage. In business strategy, competitive advantage refers to a firm's ability to grow profits sustainably, ideally by growing market share or retaining market share in a growing market. In *The Competitive Advantage of Nations*, Porter (1990) measured a country's competitive advantage in a sector by its share of world exports in that sector.

The transparency of this measure explains its enduring popularity. Market share proxies for absolute advantage in trade theory, or the rank of a country's productivity compared to other countries within a sector.[3] A gain in market share indicates that a country has improved its ability to provide a combination of cost and quality preferred by consumers.

A potentially beneficial approach is to target industries that are relatively more productive or potentially more productive compared to others (that is, an industry with high comparative advantage or high technological relatedness), but that has a low market share compared to other countries (that is, an industry with low competitive advantage). Such industries may have room to grow. In contrast, an approach of targeting sectors with the greatest competitive advantage, as measured by large international market share, will target sectors that are already very large.

Gain in competitive advantage measured by world market share is neither necessary nor sufficient to measure the net benefit of industrial policy. If a country succeeds at lowering cost or raising quality, this could increase global demand, allowing exports to rise while market share remains constant. Alternatively, market share could rise as exports hold constant in a declining global market. There are also international considerations. An industrial policy with positive net economic benefit targeting market share expansion may nonetheless cause conflict with other countries if the policy takes market share from those countries.

[3] Neary (2003) suggests an alternative definition of competitive advantage, measured by the number of firms in the domestic product market. Lower entry costs or antitrust enforcement could increase the number of firms, reducing markups and allowing the country to offer lower prices, holding efficiency, costs, and quality constant. In practice, the business strategy definition of competitive advantage based on market shares has dominated.

The How: Implementing Export-led Development Strategy

Once a country has targeted sectors, what should it do to encourage their export-led growth? I begin here by describing how many industrial policies, including those that have been used by developing countries, are prohibited under the rules of international trade agreements. I then discuss industrial policy instruments that are not prohibited and institutional structures like export promotion agencies. Many of the policies described here may appear to be untargeted towards any sector, and so one might argue they are not industrial policy per se. Yet, even if these policies can be applied to the whole economy in principle, governments with limited resources make choices about who will be the main beneficiaries, and tailor these policies to address specific constraints in those beneficiaries' industries. In practice, selected beneficiaries of these policies are often those in sectors with international market growth, weak international competition, technological relatedness, and/or comparative or competitive advantage.

Different Rules for Developing Countries

An export-led development strategy requires access to global markets, and market access requires good relations with trading partners and continued engagement in trade agreements. But while developing countries have benefited from participation in the world trading system, there is an irony. The rules of the trading system are biased against them, both de jure and de facto.

De jure, developing countries still face higher tariffs compared to wealthy countries. Calculations of actual tariff rates include preferences under bilateral and regional trade agreements (such as a free trade area or customs union) and provisions like most-favored nation status. Using the 2017 tariff data from the World Integrated Trade Solution (WITS), which are complete for almost all countries, the pattern that has emerged over time is that for nonagricultural goods (HS Chapters 25–89), value-weighted applied tariffs are 1.4 percent for low-income countries and are 2.1 percent for high-income countries, using the World Bank income classification. They are largest for lower-middle income countries at 2.5 percent. Low-income countries retain some market access privileges given their income status, but these privileges are not available to middle-income countries. Trade agreements have allowed countries to retain higher tariffs especially in agriculture, which is the comparative advantage of many developing countries. Guatemala's national development strategy is direct about the issue (República de Guatemala 2023, p. 472):

> Despite the positive results shown by exports in the last twelve years (with some important ups and downs in some periods), the country has had to put up with the high levels of tariff protection in sectors of special interest—particularly the primary sector—that persist in the main trading partners (U.S. and Europe), despite the intensification of trade relations.

De facto, developing countries have less leeway to discriminate against trading partners and to flout the rules of trade agreements. One example is that because litigation of trade rules requires significant resources, wealthy countries are more able to undertake litigation compared to developing countries. For instance, a common trade dispute involves dumping, which describes a situation a situation in which a company is alleged to have exported a product at a price lower than the price it normally charges on its domestic market. Trade rules allow countries to charge tariffs to offset dumping, which can be large, on average 10–20 times higher than tariffs imposed under most-favored-nation status. Nunn (2019) shows these are much more frequently initiated by wealthier countries against developing countries, rather than the other way around.

In other cases, high-income countries enact policies that openly conflict with international trade rules. Global Trade Alert is a nonprofit organization which compiles perhaps the most comprehensive lists of trade-limiting policies. Juhász et al. (2023) use these data to catalog "the who, what, when, and how of industrial policy," with a focus on policies that "discriminate against foreign commercial interests" and find these are most implemented by high-income countries. An example is the provision in the US Inflation Reduction Act of 2022 that, if a consumer is to receive a tax credit after purchasing an electric vehicle, a percentage of the value of minerals in the battery must be extracted or processed in the United States, or a country with which the United States has a free trade agreement. The intention of this requirement is to stimulate battery manufacturing in the United States and the 20 countries with which it has free trade agreements. Such a requirement appears to conflict with the World Trade Organization Agreement on Subsidies and Countervailing Measures, which prohibits any subsidies that are available contingent on the use of domestic inputs or "local content." WTO rules also prohibit industrial subsidies contingent on export performance, which were used historically as an industrial policy tool—for example, in Korea.

An underlying principle behind international trade agreements is to avoid subsidy races, in which dumping or domestic content rules spread from country to country. This fear has some foundation. Using data from the Global Trade Alert and other sources, Evenett et al. (2024) show that implemented industrial policy "measures are correlated with the past use of measures by other governments in the same sector, pointing to the tit-for-tat nature of industrial policy."

Industrial Policy within the Rules

Is there a way for developing countries to pursue industrial policy, while still adhering to their commitments under trade agreements? Here, I offer some examples.

1) Tariff negotiations. A major factor in the decision of multinational firms to produce in and export from a country is whether they will have preferential access to the largest markets, typically the United States and Europe. For example, after the US-Vietnam Bilateral Trade Agreement, employment in Vietnam grew faster in industries most exposed to US tariff reductions, driven by foreign affiliates of

multinationals (McCaig, Pavcnik, and Wong 2023). This growth was not necessarily the result of industrial policy. Vietnam has many natural advantages, including proximity to East Asian value chains and a relatively young and educated population for its level of development. Still, the example suggests that tariff reductions through preferential trade agreements could complement policies that effectively develop a comparative advantage. Negotiating partners' tariff reductions in specific sectors can be considered industrial policy in the sense that the policy is intended to accelerate growth in these sectors relative to others.

2) Access to new customers and suppliers. Trade is characterized by significant information frictions; in particular, many domestic firms do not know the preferences of, or how to contact, all potential buyers. Many countries use export promotion agencies to address these frictions, with a variety of services. Martincus and Carballo (2008, p. 90) describe PROMPEX, the export promotion agency of Peru:

> [The agency] trains inexperienced exporters on the export process, marketing, and business negotiations; performs and disseminates analyses on country and product market trends; provides specific information on trade opportunities abroad as well as specialized counseling and technical assistance on how to take advantage of these opportunities; coordinates and supports (and in some cases co-finances) firms' participation in international trade missions and trade shows, and arranges meetings with potential foreign buyers in particular; organizes these kinds of trade events; and sponsors the creation of consortia of firms aiming at strengthening their competitive position in external markets.

This work can be considered industrial policy because it requires collecting information about opportunities and constraints in specific sectors, and targeting interventions towards those sectors. An older generation of studies suggested that export promotion agencies had been ineffective (for example, Hogan, Keesing, and Singer 1991), but recent research has called this blanket conclusion into question. The Martincus and Carballo (2008) study of Peru's PROMPEX shows services from the agency helped exporters grow, using a matched difference-in-differences strategy with firm fixed effects. The effects were greater on the extensive margin, where firms increased the number of products exported and the number of countries served, rather than the intensive margin, where firms increased sales of existing products to their current markets. In a cross-country analysis, Lederman, Olarreaga, and Payton (2010) use instrumental variables to show that export promotion agency budgets cause increases in exports.[4] Randomized experiments have evaluated the effects of specific interventions that can be implemented by these agencies. In a study of a nongovernmental organization providing new export orders to rug manufacturers in the Arab

[4] Contemporary research has found positive effects of export promotion agencies in specific high-income contexts: Belgium (Broocks and Van Biesebroeck 2017), Canada (Van Biesebroeck, Yu, and Chen 2015), Denmark (Munch and Schaur 2018), and the United States (Matray et al. 2024).

Republic of Egypt, Atkin, Khandelwal, and Osman (2017) find that the new orders led to increased exports and productivity after three years, with some evidence of learning-by-doing. In a study of providing training and consulting in digital marketing to firms in the Balkans, including search engine optimization and improved Facebook content, Cusolito, Darova, and McKenzie (2023) find a significant increase in the intensive margin of export sales in a destination-product market. Interventions to help firms identify new international input suppliers could also have benefits, though more research is needed.

3) Quality certification and standards. As tariffs have fallen, quality regulation has become a prominent constraint on trade. Countries set minimum quality standards for a good to be imported, which is permitted under trade rules if standards are applied both to foreign and domestic firms. Such standards often apply in health-related sectors like food and pharmaceuticals, but also apply more broadly. For example, units of measure reported by machinery and equipment must be verified for accuracy before sale and labeling requirements can require the verified energy consumption of a product to be displayed on packaging. Most recently, the European Union's Carbon Border Adjustment Mechanism (CBAM) requires firms to declare the emissions embedded in imports of carbon-intensive products (like cement, iron and steel, aluminum, fertilizers, electricity, and hydrogen) and to pay duties on that carbon equivalent to those paid by firms within the European Union. Rather than regulating the quality of a product, the CBAM regulates the quality of the production process. Governments in exporting countries can play a role in helping firms conform to these standards, for instance by coordinating local inspections by authorities from importing countries.

Understanding how quality regulation affects exporters is an area of active research. When a set of middle-income Latin American countries imposed quality standards, Zavala et al. (2023) show a result of increased imports—which is consistent with a belief that regulations can give greater confidence to consumers, shifting demand out. Similarly, Zavala et al. (2023) show that certain standards, like labelling requirements, where compliance is relatively easy, can help developing countries gain market share relative to wealthy countries. But other standards, in particular food safety requirements, have caused developing countries to lose market share to exporters in high-income countries that have more experience complying with these standards. An issue here is that many developing countries do not regulate quality in the local market, so that firms in those countries retain an option to sell low-quality goods domestically and have a reduced incentive to upgrade quality for the export market. Macedoni and Weinberger (2022) argue that applying quality regulation in the domestic market can enhance domestic welfare, even if it leads some firms to exit, but their model does not include the spinoff effect that enforcing quality regulation in the domestic market could provide greater market access for exporters. For example, a recent effort in Myanmar to ban the use of harmful herbicides and pesticides had health benefits for domestic consumers, but also made it easier for Myanmar's exporters to guarantee their products did not contain chemicals that are banned in other markets.

4) Sector-specific physical and regulatory infrastructure. Many sectors rely on very specific infrastructure. In India, Asher et al. (2022) show that irrigation canals increased agricultural productivity and set off a process of structural transformation. Export processing zones are often used to concentrate specific inputs needed by multinationals and exporting firms, like large volumes of electricity, natural gas, or purified water. Wang (2013) shows these zones can increase productivity in China, though this research does not distinguish between the effects of infrastructure and liberalized regulation within the zone with respect to property rights and tax incentives. Subsidies for on-the-job training or the development of new degree programs to suit the needs of industry, or international schools for the children of expatriate workers are other examples.

Constructing sector-specific infrastructure is not restricted under trade rules. The Trade Facilitation Agreement under the World Trade Organization also lays out regulatory provisions countries can take to accelerate the movement, release, and clearance of exports, such as the creation of a digital "single window" through which all firms submit required documents to the customs authority. Single-window processes may vary across sector, for instance if the product in question is perishable, or faces quality regulation abroad. So far, it has been difficult to identify cross-country variation in time to export or logistics performance that is explained by these provisions rather than other factors like GDP per capita and geography (Hillberry and Zhang 2018). Nonetheless, physical and regulatory infrastructure may be complementary to other export promotion efforts. Looking at export promotion in Tunisia, Cadot et al. (2015) find that it helped firms grow and diversify, but that the effects dissipated after three years. The authors argue that the short-lived nature of these effects reflects the program's focus on accessing new customers, rather than on long term investments in infrastructure.

5) Sector-specific public–private dialogue. Many sectors must interact with multiple regulatory agencies that fail to coordinate with one another, leading to redundant regulatory requirements. Further, governments may fail to elicit information from private firms about the constraints they face, leading to low-return public investment. Peru and several other countries have sought to resolve these challenges by establishing temporary fora for public–private dialogue known as *mesas ejecutivas*, which focus on improving productivity in a specific sector (Ministerio de la Producción del Perú 2016). Rather having a generic conference on the broad issue of national competitiveness, *mesas ejecutivas* brought together key regulators and executives in promising sectors for weekly meetings to identify and solve problems in those sectors. For instance, an aquaculture working group worked to remove regulation it deemed unnecessary, like wastewater discharge permits, and established innovation and technology transfer centers in key regions to distribute competitive research and development grants. Projects funded included the development of genetically improved fish eggs, so that local firms were no longer reliant on foreign eggs for breeding.

6) Subsidies without export or local content contingencies. Developing countries often cannot afford substantial subsidies, but it is worth noting that subsidies without local

content requirements or export contingencies are not prohibited under trade rules. This leaves scope for temporary production or investment subsidies to nascent sectors, and even conditioning these subsides on performance targets like growth in employment or value added. Under WTO rules, subsidies are "actionable," meaning that other countries can levy countervailing tariffs if they demonstrate harm to their producers. But for smaller developing economies, such subsidies could potentially benefit their domestic economy without causing measurable harm to producers in other countries.

Institutional Design and Export-Promotion Agencies

Industrial policy can be considered what Hausmann (2008) calls "high bandwidth development policy," in which the challenge is to identify "the right mix of public inputs for each sector, and more importantly, what is a valuable change from the current provision." The instruments just described are examples of public inputs. Doing this work well requires a high-capacity agency, with skilled staff. What form should such an agency take?

Given the focus of this essay on export-led industrial policy, a government's export-promotion agencies appear uniquely positioned to respond to the idiosyncratic needs of specific sectors and to coordinate the provision of public inputs by the rest of the government. With dedicated domestic and foreign liaison offices, they have dedicated staff that interact with firms and their international customers and suppliers. Yet in developing countries, these agencies have smaller budgets and fewer employees compared to high-income countries: the median agency in a low-income country has 30 employees, in a middle-income country, 50 employees, but in a high- income country, 300 employees (Choi et al. 2023).

A productive line of research would be to identify commonalities among the best- (and worst-) performing export-promotion agencies, to guide capacity building in these institutions. The export promotion agency in Costa Rica provides one example that has succeeded in medium-risk and high-risk industrial policy, all while the country steadily reduced the share of the population living in extreme poverty. Building on a comparative advantage in plantation agriculture (Méndez and Van Patten 2022), in the 1980s the country successfully diversified into other food products. In the 1990s, the country attracted Intel, the semiconductor manufacturer, to locate an export production facility there, building on a well-educated population and proximity to the US market but in the absence of export industries with related technology. By 2006, semiconductors would comprise about 20 percent of exports, and today the country has diversified into other high-value exports like medical instruments and offshore back-office functions for multinational corporations. The Intel deal was notable because "the absence of firm-specific concessions for Intel, side-deals, or large government grants stands in contrast to the tactics many countries use to land large investments" (Spar 1998, p. v). Though Costa Rica does offer tax-exemptions to foreign investors generally, such subsidies are also offered by other countries and were not decisive for Intel. The government's success was more likely due to its ability to provide a rapid and coordinated response, led

by President José María Figueres, to Intel's requirements. For instance, the Ministry of Transportation agreed to grant additional licenses to foreign air carriers if there were not sufficient flights, and accelerated development of a new cargo port; the Ministries of Education and of Science and Technology in collaboration with Intel's human resources staff and local academics developed a new associate degree program that would prepare locals for work at Intel.

Several institutional design features stand out at PROCOMER, Costa Rica's export promotion agency:

1) Cabinet-level leadership able to coordinate across agencies. The President of the agency's board is the Minister of Foreign Trade, who reports to the President of the Republic, and who can coordinate directly with other Ministers whose agencies are responsible for specific inputs. This type of leadership aligns with the principle of Rodrik (2004) that industrial policy should be "monitored closely by principal with a clear stake in outcome and who has the highest level of political authority."

2) Joint governance by the public and private sectors. The agency's board has nine directors: five from the private sector (typically chief executives) and four from the public sector (including the Minister of Foreign Trade). Participation of the private sector provides continuity between elections, and ensures that needs articulated by the private sector have a voice. Participation of the public sector ensures that policy is directed towards increasing productivity rather than only profitability. Sector working groups along the lines of *mesas ejecutivas* have been employed.

3) Close collaboration between export and foreign investment promotion agencies. Costa Rica is a country of five million people, so almost all foreign investment is export-oriented rather than intended to serve the local market, and many exports are done by multinational affiliates. As a result, the export promotion agency and the foreign investment promotion agency, CINDE, jointly coordinate the government response to foreign companies that are potential exporters. In fact, CINDE attracted Intel, while PROCOMER was established the year Intel arrived. There is evidence that foreign direct investment promotion agencies are themselves effective. Harding and Javorcik (2011) show using a difference-in-differences approach that when sectors are targeted by investment promotion agencies, they receive more foreign investment from the US economy compared to other sectors. This effect is concentrated in countries with the most cultural distance from the United States, and those countries with less effective governments and higher corruption—indicating that effects are biggest where red tape and information asymmetries were previously most severe.[5] A fruitful line of future research is to measure the complementary between these two agencies, with a focus on the potential interactions between domestic and multinational firms. Some larger countries have investment promotion agencies with mandates to promote both domestic and foreign investment. An open question is whether the absence of a dedicated foreign investment

[5] Evidence from Europe suggests foreign investment promotion agencies can also be effective in high-income countries, especially in attracting investment to less-developed regions of those economies (Crescenzi, Di Cataldo, and Giua 2021).

promotion agency in these countries diverts government attention from attracting export-oriented multinationals.

4) Dedicated revenue with regular evaluation. PROCOMER was originally funded by a levy on exports from firms operating in Costa Rica's tree trade zones, but the revenue base was later expanded to a levy on all exports. This dedicated revenue gives firms who pay the tax an interest in ensuring the agency provides value. The budget is also evaluated annually. Most recently, tax exemptions for investors were compared to benefits in terms of additional tax revenue, incremental wage increases, and pension contributions, and it was found the agency generates about $2.50 of those benefits for every $1 in tax incentives offered to foreign investors (República de Costa Rica 2022). Ideally, this analysis should be completed by an independent government agency, like the Ministry of Finance or Treasury.

5) Unique online portal for potential exporters and buyers. Effective export promotion agencies have client relationship management tools that help them respond to the idiosyncratic needs of firms at scale. In Costa Rica, this system is built on web-based surveys available to potential exporters and investors that evaluate each group's readiness to start exporting. Some measures of readiness are generic, including whether the firm has identified a foreign distribution agent, but others are sector-specific, having to do with conformity to quality regulation in the foreign market. The tool helps the agency to focus its efforts on where the social payoff is likely to be greatest. Another tool used is a database for international buyers that allows them to browse the offerings of exporters.

6) Use of private sector competition to deliver hard and soft infrastructure. In Costa Rica, over 30 privately-owned export processing zones compete to cater to the needs of exporters and investors. Some of the zones provide plug-and-play office and production space, as well as schooling for the children of employees, and all offer renewable energy. By relying on the competitive market to deliver infrastructure, the investment and export promotion agencies have enriched their offering, without additional expenditure.

Conclusion

Most discussion of industrial policy these days has focused on actions by the largest economies, like China and the United States, and a narrow set of industrial policy instruments, specifically tariffs and investment subsidies. This discussion is not relevant to governments in countries with small domestic markets pursuing export-led development strategies that require them to abide by international trade rules, and who lack fiscal space for large subsidy programs. Yet, these countries nonetheless need an approach to target policy instruments that resolve market failures specific to certain export sectors, and exploit latent opportunities for productive diversification. Such policy instruments can be called industrial policy because they require sector-specific information to implement, and are targeted at specific sectors. I have provided an outline of the approach many countries are taking and

examples from a rich toolkit of policy instruments available, some of which have been validated by research in specific contexts.

The phrase "developing countries" used to be a euphemism for poor governance, some of which was linked to support of the wrong sectors or people. Even so, because lower-income countries are productive in fewer sectors and have less technology, targeting sectors can be more valuable for them. With recent improvements in governance and democratic accountability, there is cause for some optimism that the industrial policy of the future—with developing countries now licensed by economists to use it—can succeed. Even so, like all development policy, I expect this will be very hard to get right.

■ *For insightful questions and discussion, the author is grateful to Timothy Taylor and the editors, Laura Alfaro, Dave Donaldson, Nike Lawrence, Rachel Glennerster, Penny Goldberg, Ricardo Hausmann, Russell Hillberry, Asim Khwaja, Daniel Lederman, Gaurav Nayyar, Mike Nyawo, Andres Valenciano, and Deborah Winkler, as well as seminar participants at the Annual World Bank Conference on Development Economics, the Association for Comparative Economic Studies meeting in San Antonio, the New Thinking in Industrial Policy conference at Columbia University, and the Growth Lab at the Harvard Kennedy School. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

References

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman, eds. 2018. *World Inequality Report 2018.* Cambridge, MA: Belknap Press.

Asher, Sam, Alison Campion, Douglas Gollin, and Paul Novosad. 2022. "The Long-Run Development Impacts of Agricultural Productivity Gains: Evidence from Irrigation Canals in India." CEPR Discussion Paper 17414.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2017. "Exporting and Firm Performance: Evidence from a Randomized Experiment." *Quarterly Journal of Economics* 132 (2): 551–615.

Atkin, David, Arnaud Costinot, and Masao Fukui. 2021. "Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom?" NBER Working Paper 29500.

Bahar, Dany, Samuel Rosenow, Ernesto Stein, and Rodrigo Wagner. 2019. "Export Take-Offs and Acceleration: Unpacking Cross-Sector Linkages in the Evolution of Comparative Advantage." *World Development* 117: 48–60.

Balassa, Bela. 1965. "Trade Liberalisation and 'Revealed' Comparative Advantage." *Manchester School* 33 (2): 99–123.

Baldwin, Richard E., and Frédéric Robert-Nicoud. 2007. "Entry and Asymmetric Lobbying: Why Governments Pick Losers." *Journal of the European Economic Association* 5 (5): 1064–93.

Broocks, Annette, and Johannes Van Biesebroeck. 2017. "The Impact of Export Promotion on Export Market Entry." *Journal of International Economics* 107: 19–33.

Cadot, Olivier, Ana M. Fernandes, Julien Gourdon, and Aaditya Mattoo. 2015. "Are the Benefits of Export Support Durable? Evidence from Tunisia." *Journal of International Economics* 97 (2): 310–24.

Choi, Yewon, Ana Fernandes, Arti Grover, Leonardo Iacovone, and Marcelo Olarreaga. 2023. "Trade Promotion Organizations in the Pandemic World." World Bank Policy Research Working Paper 10374.

Coniglio, Nicola D., Davide Vurchio, Nicola Cantore, and Michele Clara. 2021. "On the Evolution of Comparative Advantage: Path-Dependent versus Path-Defying Changes." J*ournal of International Economics* 133: 103522.

Costinot, Arnaud, and Dave Donaldson. 2012. "Ricardo's Theory of Comparative Advantage: Old Idea, New Evidence." *American Economic Review* 102 (3): 453–58.

Costinot, Arnaud, and Dave Donaldson, and Ivana Komunjer. 2012. "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas." *Review of Economic Studies* 79 (2): 581–608.

Crescenzi, Riccardo, Marco Di Cataldo, and Mara Giua. 2021. "FDI Inflows in Europe: Does Investment Promotion Work?" *Journal of International Economics* 132: 103497.

Cusolito, Ana P., Ornella Darova, and David McKenzie. 2023. "Capacity Building as a Route to Export Market Expansion: A Six-Country Experiment in the Western Balkans." *Journal of International Economics* 144: 103794.

Evenett, Simon, Adam Jakubik, Fernando Martín, and Michele Ruta. 2024. "The Return of Industrial Policy in Data." *World Economy* 47 (7): 2762–88.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The next generation of the Penn World Table." *American Economic Review* 105 (10): 3150–82. https://www.rug.nl/ggdc/productivity/pwt/ (accessed on August 10, 2023).

French, Scott. 2017. "Revealed Comparative Advantage: What Is It Good For?" *Journal of International Economics* 106: 83–103.

Goldberg, Pinelopi K., and Tristan Reed. 2023a. "Presidential Address: Demand-Side Constraints in Development. The Role of Market Size, Trade, and (In)equality." *Econometrica* 91 (6): 1915–50.

Goldberg, Pinelopi K., and Tristan Reed. 2023b. "Is the Global Economy Deglobalizing? And If So, Why? And What Is Next?" *Brookings Papers on Economic Activity* 54 (1): 347–423.

Growth Lab at Harvard University. 2019. "International Trade Data (SITC, Rev. 2)." https://doi.org/10.7910/DVN/H8SFD2 (accessed on August 10, 2023).

Hanson, Gordon H., Nelson Lind, and Marc-Andreas Muendler. 2015. "The Dynamics of Comparative Advantage." NBER Working Paper 21753.

Harding, Torfinn, and Beata S. Javorcik. 2011. "Roll out the Red Carpet and They Will Come: Investment Promotion and FDI Inflows." *Economic Journal* 121 (557): 1445–76.

Harrison, Ann, and Andrés Rodríguez-Clare. 2010. "Trade, Foreign Investment, and Industrial Policy for Developing Countries." In *Handbook of Development Economics*, Vol. 5, edited by Dani Rodrik and Mark Rosenzweig, 4039–4214. Amsterdam: North-Holland.

Hausmann, Ricardo. 2008. "The Other Hand: High Bandwidth Development Policy." Harvard Kennedy School Working Paper RWP08-060.

Hausmann, Ricardo, and Bailey Klinger. 2006. "Structural Transformation and Patterns of Comparative Advantage in the Product Space." Harvard Kennedy School Working Paper RWP06-041.

Hausmann, Ricardo , Jason Hwang, and Dani Rodrik. 2007. "What You Export Matters." *Journal of Economic Growth* 12: 1–25.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences* 106 (26): 10570–75.

Hillberry, Russell, and Xiaohui Zhang. 2018. "Policy and Performance in Customs: Evaluating the Trade Facilitation Agreement." *Review of International Economics* 26 (2): 438–80.

Hogan, Paul, Donald B. Keesing, Andrew Singer. 1991. "The Role of Support Services in Expanding Manufactured Exports in Developing Countries." Washington, DC: World Bank.

Hummels, David, and Peter J. Klenow. 2005. "The Variety and Quality of a Nation's Exports." *American Economic Review* 95 (3): 704–23.

Juhász, Réka, Nathan Lane, Emily Oehlsen, and Verónica C. Pérez. 2023. "The Who, What, When, and How of Industrial Policy: A Text-Based Approach," What, When, and How of Industrial Policy: A Text-Based Approach," STEG Working Paper 050.

Matray, Adrien, Karsten Müller, Chenzi Xu, and Porrya Kabir. 2024. "EXIM's Exit: The Real Effects of Trade Financing by Export Credit Agencies," NBER Working Paper 32019.

Lederman, Daniel, Marcelo Olarreaga, and Lucy Payton. 2010. "Export Promotion Agencies: Do They Work?" *Journal of Development Economics* 91 (2): 257–65.

Lin, Justin Yifu. 2011. "New Structural Economics: A Framework for Rethinking Development." *World Bank Research Observer* 26 (2): 193–221.

Liu, Ernest. 2019. "Industrial Policies in Production Networks." *Quarterly Journal of Economics* 134 (4): 1883–1948.

Macedoni, Luca, and Ariel Weinberger. 2022. "Quality Heterogeneity and Misallocation: The Welfare Benefits of Raising Your Standards." *Journal of International Economics* 134: 103544.

Martincus, Christian Volpe, and Jerónimo Carballo. 2008. "Is Export Promotion Effective in Developing Countries? Firm-Level Evidence on the Intensive and the Extensive Margins of Exports." *Journal of International Economics* 76 (1): 89–106.

McCaig, Brian, Nina Pavcnik, and Woan Foong Wong. 2023. "Foreign and Domestic Firms: Long Run Employment Effects of Export Opportunities." CESifo Working Paper 10168.

McKenzie, David. 2024. "Is There Still a Role for Direct Government Support to Firms in Developing Countries?" New Zealand Economic Papers. https://doi.org/10.108‰7799 54.2023.2290484.

Méndez, Esteban, and Diana Van Patten. 2022. "Multinationals, Monopsony, and Local Development: Evidence from the United Fruit Company." *Econometrica* 90 (6): 2685–2721.

Ministerio de la Producción del Perú. 2016. Mesas Ejecutivas: A New Tool for Productive Diversification. Lima, Peru: Ministerio de la Producción del Peru.

Munch, Jakob, and Georg Schaur. 2018. "The Effect of Export Promotion on Firm-Level Performance." *American Economic Journal: Economic Policy* 10 (1): 357–87.

Neary, J. Peter. 2003. "Competitive versus Comparative Advantage." *World Economy* 26 (4): 457–70.

Nunn, Nathan. 2019. "Rethinking Economic Development." *Canadian Journal of Economics/ Revue canadienne d'économique* 52 (4): 1349–73.

Porter, Michael E. 1990. The Competitive Advantage of Nations. New York: Free Press.

República de Costa Rica. 2022. *Balance de Zona Franca 2017–2021.* San Jose, Costa Rica: República de Costa Rica.

República de Guatemala. 2023. *Plan Nacional de Desarrollo K'atun: Nuestra Guatemala 2032.* Guatemala City: República de Guatemala.

Rodrik, Dani . 2004. "Industrial Policy for the Twenty-First Century." Harvard Kennedy School Working Paper RW04-047.

République du Sénégal. 2018. *Plan d'actions prioritiares 2019–2023.* Dakar: République du Sénégal.

Spar, Debora. 1998. "Attracting High Technology Investment: Intel's Costa Rican Plant." Foreign Investment Advisory Service Occasional Paper 11.

Van Biesebroeck, Johannes, Emily Yu, and Shenjie Chen. 2015. "The Impact of Trade Promotion Services on Canadian Exporter Performance." *Canadian Journal of Economics/Revue canadienne d'économique* 48 (4): 1481–1512.

Verhoogen, Eric. 2023. "Firm-Level Upgrading in Developing Countries." J*ournal of Economic Literature* 61 (4): 1410–64.

Wang, Jin. 2013. "The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities." *Journal of Development Economics* 101: 133–47.

World Bank. 2023. *Unfair Advantage: Distortive Subsidies and Their Effects on Global Trade.* Washington, DC: World Bank.

Zavala, Lucas, Ana Fernandes, Ryan Haygood, Tristan Reed, and Jose-Daniel Reyes. 2023. "Quality Regulation Creates and Reallocates Trade." World Bank Policy Research Working Paper 10601.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 27–54

The Political Economy of Industrial Policy

Réka Juhász and Nathan Lane

Industrial policy, or state action meant to shift the composition of economic activity, has been controversial since the birth of economics. Economists have long studied, dissected, and taxonomized the market failures that might justify an industrial policy intervention. Juhász, Lane, and Rodrik (2023) group the theories into three main categories. The category of "externalities" includes negative externalities like environmental harms, as well as positive externalities like learning by doing, national security, or good jobs. The category of coordination or agglomeration activities involves situations where an industry can be affected by the availability of complementary goods, services, or related downstream and upstream activities. Public goods can also be industrial policies when they target specific types of economic activity. The category of public good provision arises when private production depends on adequate regulation, education, infrastructure, or law enforcement.

When it comes to pursuing industrial policy, however, political and economic forces are inevitably intertwined. In fact, economists' apprehension about industrial policy is actually less about the economic rationales per se than it is about the political economy of industrial policy (Krueger 1990). In reality, politics means that

■ *Réka Juhász is Assistant Professor of Economics, University of British Columbia, Vancouver, British Columbia. She is also a Faculty Research Fellow, National Bureau of Economic Research, Cambridge, Massachusetts, and a Research Affiliate, Centre for Economic Policy Research, London, United Kingdom. Nathan Lane is Associate Professor of Economics, University of Oxford, Oxford, United Kingdom. He is also a Research Affiliate of the Center for Economic Studies at the Ifo Institute for Economic Research, Munich, Germany. Their email addresses are reka.juhasz@ubc.ca and nathaniel.lane@economics.ox.ac.uk.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.27.

the practice of industrial policy will look very different from what simple theory prescribes.

For example, consider industrial policies that promote infant industries. These arise when a market failure (for instance, a learning-by-doing externality) prevents the industry from emerging at market prices. In theory, a temporary policy, such as a limited period of import tariff protection, will raise the domestic price, incentivizing producers to enter the market. Importantly, the tariff only needs to be in place temporarily while the industry moves down its long-run cost curve through its experience in learning by doing. At this point, the domestic industry becomes competitive and the policy is discontinued (for example, Bardhan 1971).

The infant industry idea has been known for a long time. It appears in Alexander Hamilton's (1791) Report on Manufactures (as discussed by Sylla in this symposium), but it breaks through into the canonical economics literature in the writings of John Stuart Mill (1848). However, Mill's later correspondence shows that he had been receiving letters about how his careful discussion of short-term and temporary import protection for an infant industry was being used as justification for permanent protectionism. Mill (1868) responded to one correspondent in these terms:

> [I]ndustries artificially fortified, even though it be professedly for a time only, raise up private interests which combine, as they have done in the United States, but too effectually, to convert what was intended as a temporary expedient into a permanent institution . . . These considerations have greatly shaken the consideration I expressed in my book; and though I think that the introduction of a foreign industry is often worth a sacrifice, and that a temporary protecting duty, if it was sure to remain temporary, would probably be the best shape in which that sacrifice can be made, I am inclined to believe that it is safer to make it by an annual grant from the public treasury, which is not so likely to be continued indefinitely . . .

Other prominent economists of the time followed a similar intellectual trajectory. Alfred Marshall was an early advocate of industrial policy but pivoted after observing the policy play out in the United States: "[P]rotective policy in fact was a very different thing from a protective policy as painted by sanguine economists" (Irwin 1991; Marshall and Whitaker 1975, p. 93). Frank Taussig (1914), in his early empirical study of nineteenth-century US infant industry tariffs, noted that professedly short-term industrial policies had a way of enduring.

In modern terms, industrial policies can suffer from time inconsistency. Infant industry policy may be counterproductive in the absence of government commitment: after all, if firms believe the government will extend the policy indefinitely, they may underinvest in the cost reductions required to become competitive in international markets (Matsuyama 1990; Tornell 1991). However, infant industry programs have sometimes overcome this challenge; for example, Taiwan's Industrial Development Bureau withdrew temporary protection for local producers of

VCR players when the industry failed to become internationally competitive (Wade 1990).

More broadly, the infant industry example illustrates that industrial policy is a deeply political phenomenon. Industrial policies are chosen by policymakers operating in political institutions. These politicians belong to coalitions, are swayed by constituents, wield power (formal and informal), and care about retaining it. Industrial policies have distributional consequences and impact firms, sectors, and regions, as well as workers and owners of capital. Their benefits and beneficiaries are often specific and identifiable, while their costs are often diffuse, making them a potent way to target political constituents. Thinking in practical terms about industrial policy immediately and unavoidably dunks us into the world of political economy.

In short, understanding the patterns of what industrial policy is enacted, how policies evolve, their palatability, and why policies succeed or fail requires a political economy of industrial policy. In this paper, we consider the political forces shaping how industrial policy is chosen and the dimensions of state capacity shaping how industrial policy is implemented.

Strangely, modern political economic analysis of industrial policy is scant, even amid the new body of economic research on industrial policy. Our goal is to make modest inroads toward a more robust political economy of industrial policy. We demonstrate the utility of this framework using case studies and data on industrial policy practice, drawing on data from Juhász et al. (2022). Our paper combines insights from across the social sciences with the language of the economic field of political economy. Indeed, the research literature on comparative social science and comparative politics has long considered how politics and noneconomic forces shaped the use of industrial policies (notable examples include Wade 1990; Haggard 1990; Evans 1995; Chibber 2002). The crescendo of recent industrial policies across the United States, the European Union, China, India, Brazil, and elsewhere has made understanding the political economy of these policies all the more pressing.

In the next two sections, we structure our discussion of the political economy of industrial policy around (1) political constraints for what choices are made and (2) capacity constraints that affect implementing these choices. In our discussion of political constraints, we consider the case of climate change policies. On economic principles alone, carbon pricing is part of a first-best policy response to the market failures associated with carbon emissions. Yet, carbon pricing policies have often proven difficult to implement, including, famously, in the United States. In contrast, industrial policies targeting green activity have proliferated. This case illustrates key ways in which political realities affect which climate policies are adopted. More generally, we examine the ways in which political constraints such as time inconsistency and political credibility impact and, at times, undermine industrial policy, as well as how thoughtful industrial policies may work to overcome these constraints.

The second section illustrates capacity constraints using the experience of export-promotion industrial policy in Thailand, where domestic politics first precluded and then supported the adoption of East Asian–style industrial policy

in the 1970s and 1980s. The state's inability to effectively implement its desired industrial policy contributed to the failure to shift the economy towards export promotion in the 1970s. In contrast, investments in the Thai state's ability to design, deploy, and monitor industrial policy supported a more successful export boom in the 1980s. Indeed, we argue that virtually every successful industrial policy episode has involved substantial new investments in state capacity. More generally, we consider the tension between the need for capable, autonomous bureaucracies and the reality that industrial policy is designed and deployed by political and, at times, politicized entities. We also explore how industrial policy agencies navigate the informational challenges posed by doing policy well.

In our view, a political economy of industrial policy is neither naïve nor fatalist about the challenges of conducting industrial policy. Instead, we offer a constructive confrontation with the dilemmas facing policymakers. We depart from an earlier generation of political economy work in that we do not view government failure as a necessary feature of industrial policy. Rather, it is endogenous—more likely to emerge when industrial policies are chosen beyond a country's political and capacity constraints. One implication is that any economy should be wary of brashly mimicking the policies pursued in other places: after all, successful industrial policies need to work within their political economy environment, and these particulars may vary and shift.

Political Constraints

In this section, we consider how political realities impose constraints on choices about industrial policy (Drazen 2000; Persson and Tabellini 1990). We begin with a case study discussion of the role of political constraints in choices about climate policy. We then build on this example to identify two particularly salient types of political constraints that influence industrial policy choices: (1) how the policy-making process introduces issues of political credibility and time inconsistency and (2) the constraints posed by the reality that politicians wish to hold power.

Case Study: Political Constraints Facing Green Industrial Policy and Carbon Pricing

The experience of green industrial policy is a tale of political constraints. We refer to green industrial policies as those that aim to change the domestic structure of economic activity towards lower-carbon technologies. The primary justification for such policies, of course, is the environmental damage from carbon and carbon-equivalent emissions. But alongside the negative externality arising from carbon emissions, there is a second set of potential market failures associated with the innovation needed to invent and diffuse low-carbon or carbon-neutral technologies. As a result, the first-best policy may be a combination of carbon taxes, which we would not call green industrial policy, and direct subsidies for clean-energy innovation and adoption, which is clearly a form of green industrial policy (Acemoglu et al. 2016).

Theoretically, the market failures around carbon emissions present a glaring case for carbon, or Pigouvian, pricing: that is, using tax or emissions trading schemes to equate the private and social costs of carbon. In terms of economic efficiency, pricing carbon is seen as low-cost, simple to administer, market-based, and less invasive than the alternatives. Moreover, there is evidence that carbon pricing also incentivizes firms to invest in energy-efficient technologies (Colmer et al. 2023). Consider the case of Norway's principal energy firm, Equinor (previously Statoil), which invested in carbon capture and storage technology in response to the substantial carbon tax introduced by the country in 1991 (Rathi 2024).

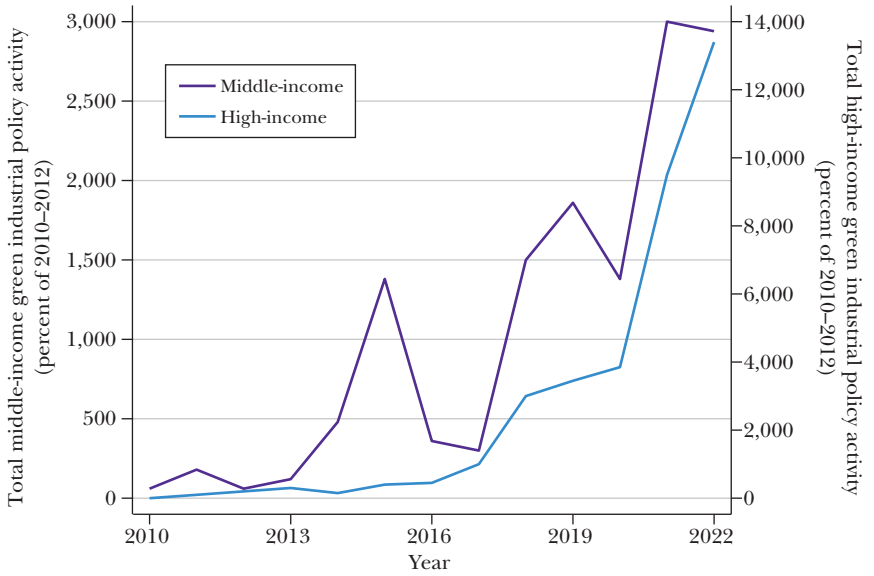
Among economists, green industrial policies have been far more controversial. These policies are seen as inferior to carbon pricing, especially in terms of efficiency. Unlike a single carbon price, green industrial policies target specific domestic activities and do so using a wide array of different instruments—some more inefficient than others. Examples of such policies are subsidies for research and development in new green technologies (likely part of a first-best policy response), guaranteed premium prices for electricity generation from renewable sources ("feed-in tariffs"), or targeted support for the domestic production of green products (like electric vehicles, wind, or solar). Some green industrial policy instruments may even, perversely, *slow* the transition to net zero emissions, at least in the short run. For example, many countries have placed stringent local content requirements on the production of wind turbines, solar panels, and electric vehicles. While these policies potentially boost the domestic supply chain, they may also risk raising the domestic price of renewables, lowering demand, and impeding their deployment. Such dilemmas have only underscored the economic appeal of carbon pricing.

For these reasons, price-based policies have become the preferred vehicle for addressing climate change among economists. The "Economists' Statement on Carbon Dividends," which advocates for a US carbon price, has been signed by 28 Nobel laureates in economics, four former chairs of the Federal Reserve, and 15 former chairs of the US Council of Economic Advisors (at https://www.econstatement.org/). In the words of the former managing director of the International Monetary Fund (IMF), Christine Lagarde, "[p]rice it right, tax it smart, do it now" (quoted in Ball 2018, p. 134).

However, green industrial policies have proliferated while the adoption of carbon pricing policy has been more problematic. In fact, after decades of efforts to implement carbon prices, the World Bank (2024) estimates that less than 1 percent of global emissions are covered by a carbon price above the recommended level, and only 24 percent of global emissions face any carbon price at all. The United States, the world's second-largest emitter of carbon behind China, has no federal price on carbon.

As a starting point for discussing the recent diffusion of green industrial policy, we use the first comprehensive dataset on global industrial policy practice from Juhász et al. (2022). This dataset uses natural language processing and a supervised machine learning model to identify industrial policies from the Global Trade Alert dataset, an independent organization set up in 2009 to track international

Figure 1
Green Industrial Policy Activity in G20 Countries, 2010–2022 (**Annual Count of Policies Relative to 2010–2012 Average**)



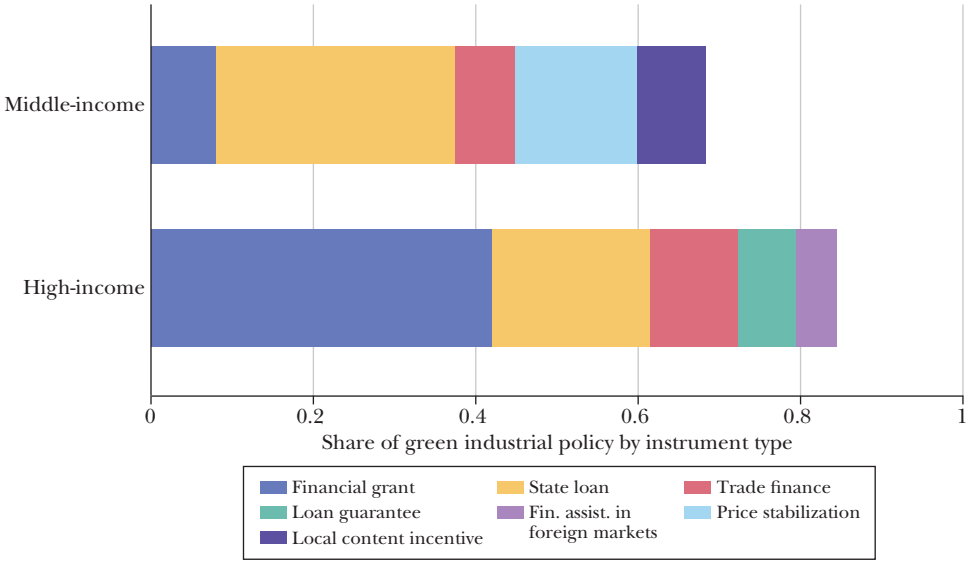
Source: Green industrial policies are classified based on the industrial policies identified in Juhász et al. (2022), who use data from the Global Trade Alert. High- and middle-income status is classified using data from the World Bank.
Note: An industrial policy is classified as being "green" if the text of the policy description contains keywords associated with green policies. G20 countries are listed in online Appendix A.

commercial policy activity. Importantly, Global Trade Alert attempts to capture all types of policies (for example, producer subsidies, tariffs, and consumer subsidies) that discriminate (positively or negatively) against foreign producers over domestic ones. We use the Juhász et al. (2022) database to extract green industrial policies, which we identify using a dictionary of green policy keywords (for a description, see online Appendix A.1).

Figure 1 illustrates the dramatic expansion of green industrial policies over the past decade. Specifically, we plot trends in green industrial policy activity across the Group of Twenty (G20) countries, which includes the European Union and 19 of the other largest world economies. Together, the G20 countries account for about 85 percent of global GDP and about two-thirds of the global population. In this analysis, we divide countries into high- and middle-income status using World Bank definitions. Figure 1 shows that green industrial policy activity has risen sharply across both middle and high-income economies. While there was little growth in

Figure 2

Top Five Green Industrial Policy Instruments across G20 Economies by Income Group, 2010–2022



Source: This figure represents green industrial policies by instrument type (top five instruments only). An industrial policy is classified as being "green" if the text of the policy description contains keywords associated with green policies.
Note: Green industrial policies are classified based on the industrial policies identified by Juhász et al. (2022) using the Global Trade Alert database. High- and middle-income status is classified using data from the World Bank.

green industrial policy activity until about 2015, we see a "hockey-stick" type time trend across both middle- and high-income economies starting later in the 2010s.

Which green industrial policies, specifically, are behind these trends? In Figure 2, we plot the share of industrial policy activity by the type of instrument, focusing on the most popular measures. Across both high- and middle-income countries, the most common instruments tend to be targeted financial grants and fiscal instruments, such as loans, loan guarantees, and trade financing for green activities. Among these are less controversial policies, such as green research and development subsidies.

However, Figure 2 also shows the relative popularity of more controversial local content incentives among middle-income economies. Examples include Chinese consumer subsidy programs from the early 2010s for the purchase of energy-efficient household goods (like refrigerators, air conditioner units, and television sets) that met local content requirements. Similarly, loans for wind generation from the Brazilian National Bank for Economic and Social Development (BNDES) have been

tied to local content requirements. Despite the limitations discussed above, local content incentives may benefit domestic producers in countries behind the technology frontier; absent such conditions, the subsidies and other incentives of these green policies may benefit producers abroad. Local content incentives also make an appearance in high-income countries, as in the US Inflation Reduction Act of 2022, in which tax credits for electric vehicles or renewable electricity generation are subject to local content requirements.

The global diffusion of various green industrial policies in Figures 1 and 2 stands in contrast to the difficulties of implementing carbon pricing. What explains this difference?

Despite its economic attractiveness, the political constraints for carbon pricing have been formidable (Furceri, Ganslmeier, and Ostry 2023). The costs of carbon taxes are immediate, while the benefits accrue in the future. These costs cut across traditional economic and political coalitions: consumers and producers, labor and capital, and political left and right constituencies. Across political systems, these dynamics have worked to the advantage of opponents who are able to veto climate policies (Mildenberger 2020). Hence, carbon pricing has been politically contentious, inspiring researchers to focus on their political constraints (for example, Jenkins 2014; Karapin 2016; Cullenward and Victor 2020). Leading carbon pricing researchers have declared that political acceptability is a first-order concern; for example, Klenert et al. (2018, p. 669) argue that "[t]raditional economic lessons on efficiency and equity are subsidiary to the primary challenge of garnering greater political acceptability."

In numerous settings, carbon pricing has threatened industry incumbents, who have become pivotal antagonists in the politics of carbon pricing (Brulle and Downie 2022; Basseches et al. 2022). The most prominent US emissions trading legislation, the American Clean Energy and Security Act of 2009 (commonly known as the Waxman-Markey bill) was sunk by lobbying from expectant losers, including nonemitting industries indirectly exposed to potential losses (Meng and Rode 2019; Cory, Lerner, and Osgood 2021). The year before, Canada's Liberal Party imploded, a loss driven in part by a controversial national carbon tax scheme that earned the ire of carbon-intensive provinces and constituents. Carbon pricing wins have also generated political blowback (Pahle et al. 2022), sometimes with dramatic reversals as has occurred in Australia, France, Switzerland, and the state of Washington. In contrast, the carbon-pricing experience has been smoother for early adopters, like Sweden and Finland, with more amiable political climates and weaker incumbents (Meckling, Sterner, and Wagner 2017; Harrison 2010).

Politics has also constrained ratified carbon-pricing schemes in ways that have rendered them less effective. Carbon pricing wins—carbon taxes and trading systems alike—have required political bargains with varying degrees of concessions, exemptions, and rebates. The result is that the price of carbon is often too low. This was true for the world's largest carbon pricing scheme, the EU Emissions Trading System, which passed as a politically feasible alternative to failed European carbon taxes. In the European Union, tax policy would have required unanimous support

from member states, whereas the Emissions Trading System was packaged as an environmental policy and faced lower political hurdles. To garner early buy-in, the European Union gambled by providing firms with carbon emissions allowances. Although the move cultivated industry support and constituencies for the Emissions Trading System, it also inspired intense lobbying efforts over these allowances. Subsequently, the Emissions Trading System experienced a significant period of over-allocation that depressed the price of tradeable permits to emit carbon (Sato et al. 2022). These issues are by no means unique to the EU effort, and trading schemes grapple with over-allocation and low prices due to both technical and political constraints (Quirion 2021; Jenkins and Karplus 2017).

Political realities have also meant that the administrative burden of carbon pricing can be quite large. In theory, a carbon pricing system is parsimonious: it simply administers a single global price on carbon. In practice, a patchwork of different carbon pricing policies has emerged across localities. Geographic differences in carbon prices introduce the challenge of "carbon leakage," where instead of reducing carbon emissions, they simply shift them to locations with a low carbon price. Closing such loopholes is difficult and administratively burdensome. Indeed, the European Union is currently grappling with implementing a Carbon Border Adjustment Mechanism (CBAM), effectively a tariff on carbon-intensive imported products, to level the playing field between domestic and foreign emitters. Hence, politics complicates the theoretical appeal of administering a single carbon price—with the additional wrinkle of coordinating global policy.

Compared with the political tumult of carbon pricing, green industrial policies carry manifest political advantages. Where a carbon price is a "stick," green industrial policies provide "carrots" to certain sectors or activities, making them powerful policy tools with which to build political constituencies for decarbonization efforts. With immediate concentrated benefits and diffuse costs, green industrial policies have thus been supported by both voters and firms (Meckling and Karplus 2023). Where a uniform carbon price cuts across sectors, the targeted nature of industrial policies means that they can be tailored to individual sectors and bridge constituencies across political environments (Cullenward and Victor 2020). Take, for example, a popular green industrial policy in electricity generation, the feed-in tariff, which guarantees a premium price for electricity generated from renewable sources such as wind and solar. One explanation for their widespread use across rich and poor countries alike is that they benefit politically valuable rural constituents. Bayer and Urpelainen (2016) argue this advantage explains the proliferation of this tool across democracies, specifically, its ability to simultaneously woo renewable energy producers and influential rural constituents.

Furthermore, political scientists and policy designers have noted that green industrial policies seem to provide a means of *shifting* the political environment in favor of carbon taxes. By creating the sectors and jobs that eventually benefit from carbon pricing, green industrial policies, in fact, tilt the future political landscape to one more favorable of Pigouvian policy. This pattern has been documented across varying political settings globally. Green industrial policies preceded pricing policies

in nearly two-thirds of the cases studied by Meckling et al. (2015). In California, which has become a well-studied case in these feedback dynamics, aggressive public support for renewable development dates back to the earlier policies in the 1970s (Biber 2013). Renewable industry constituencies have underpinned continual policy expansion, staved off reversals, and helped split traditional anti-climate policy coalitions. For instance, policies have promoted green energy production by utilities, who, in turn, became advocates of larger, subsequent reforms (Vormedal and Meckling 2023; Kim, Urpelainen, and Yang 2016).

The experience of climate policies shows us that whatever the arguments for economic efficiency, political feasibility is a binding constraint. Although carbon pricing and green industrial policy are often discussed as substitutes, the political economy of industrial policy provides a more dynamic view. Political economy also hints at the potential of a portfolio approach to climate policy. Increasingly, economists (Blanchard, Gollier, and Tirole 2023) argue for the advantages of multipronged approaches to addressing climate change, ones where green industrial policies play a role. They do so by complementing current carbon pricing schemes and through their potential to shift the politics of larger-scale reforms to confront climate change.

Yet, green industrial policies are not immune to their own political constraints, where less efficient interventions, such as feed-in tariffs, may be more politically feasible than measures like green research and development, which directly target technological constraints to clean energy (Harrison, Martin, and Nataraj 2017). Similarly, much like the risks facing infant industry policy (discussed above), the beneficiaries of green industrial policy may well lobby for them to be kept in place long after the rationale for them has expired. We now turn to such political challenges in more detail.

Political Credibility and Time Inconsistency

Industrial policies often have long time horizons and require politicians to commit to a sequence of policies over time. In the infant industry promotion policies discussed earlier, dynamic learning-by-doing externalities take time to be realized. It is also true for green industrial policy, which may require firms to invest in risky new technologies and products. The dynamic nature of these policies introduces issues of political credibility and time inconsistency, which can act as real-world constraints to effective industrial policies.

In practice, this setting presents a variety of challenges for policymaking. If a policy is not credible, firms will underinvest in ways that undermine the policy itself. For example, in the case of infant industry promotion, if the policymaker cannot credibly commit to removing a protective import tariff after a certain period, firms are not incentivized to undertake the investments that would make them competitive. Conversely, the threat of policy reversals can complicate their political credibility. Green industrial policies pursued across North America and the European Union face the threat of such policy reversals (Vihma, Reischl, and Anderson 2021; Marquardt, Oliveira, and Lederer 2022). As Stern (2022, p. 1271)

makes clear, "[g]overnment-induced policy risk is one of the major deterrents to [green] investment."

The political constraints posed by political credibility and time inconsistency are not unique to industrial policy, in fact, such issues permeate monetary and fiscal policy. A large body of work in normative political economy is dedicated to thinking about how to design institutions and policies that overcome such challenges. Most prominently, political economists have emphasized the power of delegation (Persson and Tabellini 1999), in which aspects of policy are devolved to independent organizations insulated from political forces. Indeed, some instances of industrial policy, such as those used in post–World War II Japan, have featured institutional delegation—an issue which we return to in our discussion of "state capacity" below.

However, institutional design is itself a political choice. Although monetary policy has been delegated to independent authorities, the distributive and particularist nature of fiscal policy has made delegation less common. Nevertheless, the world has seen a proliferation of fiscal councils and fiscal rules meant to overcome issues of political credibility (End 2023; Larch and Braendle 2018). Independent industrial policy institutions (discussed below) also exist. In the realm of trade reform, supranational authorities and multilateral agreements have lent outside credibility to trade policy reform (Rodrik 1995; Staiger 1995). As international organizations find their footing during the current industrial policy renaissance, one wonders if supranational bodies can play a similar institutional role in credible industrial policy. This consideration may be especially relevant with the return of industrial policy in the European Union, where EU-wide institutions govern the forms of state aid pursued by member states.

Although delegation may be possible for industrial policy in certain contexts, for better or worse, much of industrial policymaking is likely to remain firmly in the domain of politics. In these contexts, the question of designing policy in the face of these constraints is essential. In her guide to what makes industrial policy work, Harrison (2024) provides a framework for better design: industrial policy should correct market failures, consult the private sector, promote competition, and—importantly—*conclude*. Specifically, she considers the ways in which legislation can be written to expire, sunset, and terminate. This means designing industrial policies that "self-destruct," mitigating risks that policies become entrenched.

Policy Choice, Political Equilibria, and Political Power

When are the appropriate industrial policies chosen, if at all? Industrial policy is the outcome of a political equilibrium, which is shaped by the desire of policymakers to hold power.[1] When industrial policy shifts the structure of economic activity, it often creates winners and losers. For example, policies that promote green energy production may threaten coal-belt politicians (Hess 2014). When economic policy

[1] This section adopts the framework of Acemoglu and Robinson (2013) and Robinson (2010), and draws on Persson and Tabellini (1990), Drazen (2000), and Bueno de Mesquita (2016).

choices carry political consequences, they may work against policymakers' incentives (Acemoglu and Robinson 2013). For example, our case study suggests that carbon pricing policies are less appealing in practice than in theory because policymakers perceive them as politically risky. Hence, the political environment shapes which industrial policies are chosen. Taking political economy seriously, Robinson (2010) considers the two ways in which industrial policies are adopted: (1) working within constraints posed by the current political equilibrium or (2) shifting the political equilibrium itself. It is worth unpacking each.

First, policymakers can propose an industrial policy that works within the current parameters of the policy environment: extant coalitions, key players, current administrative capacity, and so on. Such policies may be more politically pragmatic and employ the existing pockets of state competencies. The multipronged nature of the Inflation Reduction Act of 2022 in the United States has been criticized for its array of objectives, which include decarbonization, the creation of "good" jobs, and reshoring supply chains from geopolitical adversaries. Seen through the lens of political economy, the wide range of objectives may make it more feasible to pass such legislation by appealing to multiple constituencies. Indeed, the local content requirements tied to different policies in the legislation ensures that the tax credits provided for politically contentious decarbonization goals benefit local producers, making the goals more politically palatable.

Thus, working within current political constraints can lead to outcomes that are "second best, at best," in the face of the political rules and dilemmas facing policymakers (Dixit 2009; Rodrik 2008). Our case study in the next section describes how Thailand was initially unsuccessful at mimicking the export-led policy of East Asian neighbors but was eventually able to adopt a version that worked within its distinct political constraints.

Second, rather than working within the constraints of the current political equilibrium, the political equilibrium can shift; for example, a policy can empower beneficiaries whose incentives are aligned with the industrial policy. This may sound abstract, but this is exactly the logic seen in the case study above, where green industrial policy creates constituencies—and thus shifts the political equilibrium—for future carbon pricing.[2]

The main message of this section is that while the choice of industrial policy needs to account for economic principles, it also needs to account for the local political environment, which requires attention to the political institutions and the political incentives they promote, the key players, the distribution of power, and how policy may alter it. The experience of climate policy illustrates how political realities shape which policies are chosen and how the policy choices today may modify the future political equilibrium. Accounting for the political environment also implies that policies that emerge in one political context are not guaranteed to

[2]This observation opens up questions as to the parameters of policy advice and the degree to which economists ought to internalize the political incentives of policymakers. For a discussion, see Dixit (1997) and Zingales (2020).

work within another. Our case study of Thai industrial policy in the 1970s and 1980s below illustrates this point.

State Capacity Constraints

State capacity—and its role in deploying policy—has become an essential ingredient in explaining long-run development and the divergent experiences of post–World War II industrialization (Evans 1995). Just as many East Asian economies demonstrated an ability to pursue policies that supported development and industrialization, the post–World War II period also produced a rogue's gallery of predatory states, such as those in the Philippines, Ghana, and Zaire, that became case studies in botched policymaking (Killick 2010; Boyce 1993). History is littered with five-year economic plans that were ill-conceived or vastly outstripped the ability of states to implement them. Historically, moves toward industrial policy have required thinking about state capacity.

It would be wrong to think of state capacity as static and exogenous, especially in the context of industrial policy. Positive and formal political economics sees such capacity as the endogenous outcome of investment decisions made by governments subject to their political environment (Besley and Persson 2011). For example, underlying the development success story of South Korea was one of continual investment in bureaucratic capacity. This capacity was by no means exogenous. Under General Park Chung Hee in South Korea, "[t]he developmental state was not a given, but a human artifact" (Kim 2011, p. 86), one cultivated by continual investment and political choices. In fact, the postwar South Korean state was initially seen as weak; there was not a developmental state waiting to be helmed, and the state Park "inherited was a politically demoralized and technically backward institution" (Kim 2011, p. 86).

In this section, we begin with a case study of how Thailand attempted to reshape its industrial policy from the 1970s onwards with mixed success. Thailand's experience illustrates both of our main themes; namely, how the political realities discussed in the previous section constrain and shape the industrial policies that are chosen, and how state capacity affects the ability to execute the industrial policy successfully. With this case study for reference, we then focus on two dimensions of state capacity that dominate industrial policy considerations: (1) bureaucratic capacity, or the ability to implement policy, and (2) embeddedness and informational capacity, the ability of bureaucracies to interact with and exchange information with the private sector.

Case Study: Export-Led Industrial Policy in Thailand

Thailand is a useful lens for considering how political and capacity constraints shaped their ability to pursue a style of industrial policy popular in some East and Southeast Asian economies that focused on assisting certain industries—via instruments such as cheap loans, outright subsidies, and technological assistance—to develop so they could expand their exports.

In the 1950s, Thailand's military-dominated governments pursued an inchoate form of import substitution industrialization—a developmental strategy focused on replacing imported industrial goods with domestic production. Thai-style import substitution did not embody grand developmental strategies but served important practical (in terms of trade and revenue) and political purposes, helping maintain fragile post–World War II politics.

Export enthusiasm came to Thailand in the early 1970s under the government of Field Marshal Thanom Kittikachorn (Hewison 1987). The experience of earlier East Asian export-promotion success stories, like South Korea, Taiwan, and Japan, resonated with Thailand's technocrats, and export aspirations marked Thailand's Third Five-Year Plan (1972–1976) and the Export Promotion Act of 1972.

However, Thailand could not fully pursue an all-out export drive like those pursued by its East Asian counterparts. For a time, Thailand supported a contradictory mix of export promotion and import substitution, or what was called "export-oriented protectionism" (Poapongsakorn and Fuller 1997, p. 480). For example, where South Korean export policies allowed de facto import liberalization for exporters (Westphal 1990) so that exporters could purchase imported goods as inputs at world prices, Thai policy did not. Such import support was politically infeasible, given the industries reliant on protection were important constituencies. Although Thai exporters were given rebates from the import tariffs on their intermediate input purchases, the rebates were insufficient and mismanaged (Christensen et al. 1990). In fact, Thailand protected critical machinery and intermediate goods without adequate relief for exporters and even raised protection for capital goods throughout the decade (Wiboonchutikula 1987).

While the pressures of Thai politics made it difficult to roll back the tariffs of import-substituting industrialization, a weak development bureaucracy stymied the shift to export promotion. Despite spurts of reforms, Thailand had not invested in a developmental bureaucracy, and through the 1970s, oscillating military and civilian governments (mostly the former) politicized swaths of the economic bureaucracy. Developmental bodies, replete with duplication, were just vehicles for political patronage (Rock 1994; Doner and Ramsay 2000). The effect was a balkanized and fragmented developmental apparatus (Crouch 1984). The Thai Board of Investment, a key industrial strategy body, lacked the "capacity to monitor promoted firms, much less to impose any clear performance standards on them" (Doner and Ramsay 1997, p. 252). Where countries like South Korea developed systems for scrutinizing export incentives in the 1960s so that only firms who showed an ability to export were eligible for support, Thailand in the 1970s lacked the administrative capacity to condition support in this way.

Another important factor prevented an export push in the 1970s: the exchange rate for the Thai baht was overvalued. Thai political constraints made devaluation improbable, unlike in post–World War II Taiwan and South Korea, whose politics allowed—or even compelled—them to pursue politically difficult devaluations, which were preconditions for robust export promotion. In Thailand, powerful key political players, from business groups to military elites, favored a strong baht. For

instance, a strong baht favored the military's foreign procurement, and similarly, importers and firms borrowed US dollar-denominated capital (Doner and Ramsay 2000; Warr 1993, pp. 43–44). This status quo would remain until the 1980s.

Only in the 1980s did a coherent export-promotion policy emerge in Thailand, promulgated by a new regime that seized upon a window of opportunity. This shift was the by-product of multiple crises that emerged in the 1970s—civil unrest, coups, and deepening economic crisis. The chaotic interregnum led to a new semi-democratic political equilibrium helmed by Prime Minister Prem Tinsulanonda (1980–1988), who brokered power between newly empowered political parties and traditional military interests (Doner and Laothamatas 1994). Under this "Premocracy," technocrats and pro-reform parties emerged as salient political constituents. Together, these forces created the conditions to realize a true export promotion strategy. Muscat (1994, p. 195) summarized the situation: "[N]o previous Thai government had been under the kind of severe and sustained economic pressure that now brought the technocrats to the conclusion that a thoroughgoing shift to an export orientation could no longer be delayed, and . . . an export orientation of institutional factors would be central to a successful policy."

Export promotion—this time in earnest—became a top priority under Prem and "coincided with significant technical strengthening of the infrastructure of the Thai state"—choices supported by party politics and external international institutions (Rock 1995; Muscat 1994, p.753). A substantial institutional development program was initiated to improve the government's policy analysis and implementation capabilities. These investments in state capacity created the conditions necessary to rationalize economic and industrial policy. Combined with pressure from the World Bank's Structural Adjustment Programs and IMF assistance, the political climate allowed the Prem government to finally push through currency devaluations in 1981 and 1984, despite strong resistance from the military and incumbents. The move symbolized fledgling state autonomy. Broadly, under Prem, Thailand shifted from a clientelist state to a form of "liberal corporatism," where a relatively autonomous state bargained with key constituents (Laothamatas 1994).

For instance, private-public bodies proliferated through the 1980s and were seen as instrumental for promoting exports—and Thailand's development success more broadly (Doner and Ramsay 2000). Most famously, the Joint Public-Private Sector Consultative Committee was established in 1981 and was conspicuously modeled after Japanese institutions. Chaired by the prime minister, the Joint Public-Private Sector Consultative Committee convened monthly meetings between state agencies and business groups to coordinate policy and to elicit information on export incentives. Thailand also followed the path of Korean export agencies, launching a successful Department of Export Promotion. Such reforms facilitated a more robust export strategy; import protection offsets, ineffective in the 1970s, were now widely used by the 1980s, and export credit covered over 50 percent of exports by 1983 (Herderschee 1993). The state planning authority, the National Economic and Social Development Board, organized public-private partnerships to promote

investment in the hospitality sector, establishing what "may well have been the single most important export policy success of the 1980s" (Rock 1995, p. 752)—tourism.

While Thailand's political environment of the 1980s supported a more robust export-oriented policy, the Thai route was distinct. Although commentators drew parallels between Thai private-public efforts and other countries in East Asia, ascendant business groups and lobbies exercised far more power over the state in Thailand. Thailand's outward-oriented interventionism echoed aspects of South Korea's, however, Thailand could not fully pursue key pillars of Korean policy (for example, import liberalization for exporters), nor could it adopt the more complex industrial policies seen elsewhere, such as active state involvement in the research and development process that pushed firms further towards the technology frontier (Christensen and Siamwalla 1993). Nevertheless, policymakers acted on windows of opportunity to pursue a strategy—and invest in bureaucratic capacity—that worked within Thailand's political economy. By doing so, Thailand pursued an export-oriented industrial policy that was more successful than predicted (Doner and Ramsay 2000; Rock 1994).

The Thailand experience with industrial policy illustrates several of our main messages. First, the political environment and capacity constraints inhibited the ability of Thailand to adopt wholesale the East Asian–style export-oriented industrial policy in the 1970s. Second, once the political environment shifted in the 1980s, outward-oriented industrial policy became more workable, including relaxing political barriers to currency devaluation. Policymakers used windows of opportunity to pursue a form of export promotion that was workable within Thai politics. Third, the case underscores the importance of state capacity, which at first stymied the adoption of East Asian policies. The Thai experience of the 1980s showed the importance of investment in bureaucracy, including deliberative institutions that worked well within Thailand's political economy.

Bureaucratic Capacity and Autonomy

The Thai experience illustrates that implementing industrial policies requires *bureaucratic capacity*, that is, the ability of an administrative agency to execute and monitor the policies chosen by politicians. Administrations need resources, capital, staff, technology, and knowledge to do policy. Industrial policies can be particularly capacity-intensive to administer; they often require deep knowledge of the markets and firms with which they interact, regular data, technical expertise, and more. Where dimensions of bureaucracy capacity matter for economic development (Besley et al. 2022), the quality of bureaucracies becomes paramount in pursuing rational policies.

Bureaucratic autonomy, in particular, has been an essential feature of bureaucratic capacity in the world of industrial policy. By "autonomy," we mean the ability of bureaucratic agencies to have a meaningful degree of independent authority and discretion to implement policies (Bersch and Fukuyama 2023). Autonomy is promoted by limiting political interference in managerial procedures, staff hiring, and internal promotion decisions, reducing the constraints on bureaucratic operations, and more. Given the political temptations surrounding industrial policies

(discussed in the previous section), the autonomy bureaucracies have over policy has been vital for successful industrial policy.

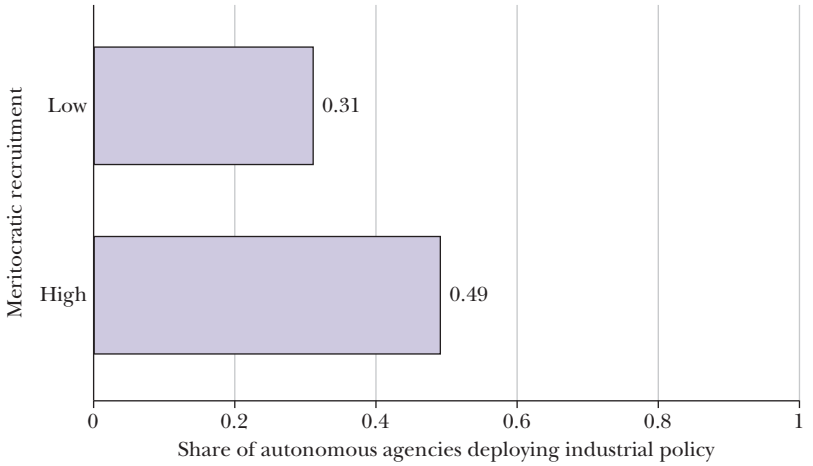
In practice, what does bureaucratic autonomy mean for industrial policy? To answer this question, we collect data on the public entities or formal bodies that implement industrial policies. We do so using the textual descriptions of industrial policies that appear in the Juhász et al. (2022) industrial policy dataset from the G20 group of jurisdictions. Specifically, we extract the names of public entities from policy descriptions with the help of prompts fed through OpenAI's ChatGPT application programming interface (for details, see online Appendix A.2). This yields a dataset of unique public entities that oversee industrial policy.

Our first observation is that industrial policy is deployed by many different types of public entities. Perhaps most familiar are government ministries (like Argentina's Ministry of Productive Development or South Africa's Ministry of Finance) that enact industrial policies, and state-owned financial institutions (like Brazil's National Bank for Economic and Social Development, or China Development Bank) that provide financing, often at below-market rates, for industrial policy projects. In addition, many G20 jurisdictions have public entities with much narrower mandates. Recent examples include Australia's Critical Minerals Facility, which finances projects aligned with the country's critical minerals strategy, and India's Solar Energy Corporation, tasked with facilitating the implementation of the country's "National Solar Mission." We also see instances of state-owned enterprises deploying industrial policy. For example, in 2010, Russian Railways, a state-owned railway company, implemented a policy providing subsidized transport rates for domestically produced cars to the Russian Far East. Similarly, the Saudi Arabian Saline Water Conversion Corporation (a state-owned utility) announced an import ban in 2009 on water desalination equipment to support the government's plan for increased domestic production in the sector.

Next, to capture one salient dimension of bureaucratic autonomy, we classify each public entity as "autonomous" if they are run by civil servants or other nonpoliticians (using the methodology developed by Field [2024], and discussed in online Appendix A.2). We consider an individual to be a politician if that person holds a political position or is affiliated with a political party. For example, government ministries and departments are typically headed by individuals holding the political position of "cabinet minister." These bodies are not classified as autonomous. On the other hand, Australia's Critical Minerals Facility (managed by Export Credit Australia) is led by the chief executive officer and chair of Australia's Critical Minerals Facility, and each have over 25 years' experience in the related fields of banking and financial services respectively and no easily identifiable political affiliation. We classify this entity as being autonomous from the government (for more examples, see online Appendix A.2).

In Figure 3, we plot our measure of autonomy against a standard measure of bureaucratic quality, the level of meritocratic recruitment, from the Varieties of Democracy (V-Dem) project. This country-level measure of meritocratic recruitment captures the extent to which state administrators are appointed based on credentials

Figure 3
Share of Autonomous Agencies Deploying Industrial Policies among G20 Countries, 2009–2022



Source: The public entities deploying industrial policy are collected from the text of industrial policies identified by Juhász et al. (2022), who use data from the Global Trade Alert. Data on meritocratic recruitment are from the Varieties of Democracy (V-Dem) project (Coppedge et al. 2024).
Note: We define a public entity as autonomous if it is run by civil servants or other nonpoliticians. We split our sample of G20 into two groups based on whether the country's meritocratic recruitment, as assessed by V-Dem, is above the G20 sample median. The share of autonomous entities deploying industrial policy within each group is defined as the mean share of autonomous agencies among the countries in the group.

rather than arbitrary criteria, such as personal or political connections (via Besley et al. 2022). Figure 3 splits countries into two groups based on whether their level of meritocratic recruitment is above or below the median level among the G20 sample, and the x-axis shows the mean share of autonomous industrial policy entities.

As Figure 3 shows, countries with high levels of meritocratic recruitment across the bureaucracy in general also tend to have more autonomous administrative bodies implementing industrial policy. Hence, for industrial policy, bureaucratic autonomy is higher in places where bureaucracies have higher levels of meritocratic recruitment. Indeed, historical evidence suggests that the pilot development agencies that deployed industrial policy in countries such as Japan, South Korea, or Taiwan evolved to have elite selection criteria, meritocratic promotion, and long, stable career paths (Johnson 1982). Highly trained civil servants staffed key economic institutions, and their incentives promoted longer-run policymaking.

Importantly, however, despite the strong case for delegating industrial policy to autonomous bureaucracies, Figure 3 shows that much of contemporary industrial policy tends to be guided by political bodies. Even among G20 countries characterized by higher levels of meritocratic recruitment, 49 percent of the agencies implementing industrial policy are headed by politicians. Thus, Figure 3 indicates

that much of industrial policy deployment is firmly in the political realm. Once again, we see that industrial policy is shaped by local political realities and constraints.

Up to this point, we have considered bureaucratic autonomy in terms of the nuts and bolts of *implementation*. However, the autonomy of bureaucracies to *formulate* policy may also be important for industrial policy design. Because industrial policies are complex, skill-intensive, and require careful design, there may be a case for delegating the details of policy formulation to higher-capacity bodies.[3] In post–World War II Japan, the pilot industrial policy agency—the Ministry of International Trade and Industry (MITI)—practiced what Chalmers Johnson (1982) famously called "administrative guidance," de facto power in shaping (and not simply implementing) the industrial policy of the 1950s and 1960s, which Johnson saw as consequential to policy success.

We can also see the power of delegating design in contemporary policymaking. Comparing the success of California's climate policies to Germany's more disappointing outcomes, Meckling and Nahm (2018) argue that bureaucratic autonomy in policy design was essential for crafting effective policy in California. Importantly, California's legislature set the policy goals, so politics was not entirely absent from the policy formulation. Similarly, Fernández-i-Marín, Knill, and Steinebach (2021) show that measures of environmental policy quality—and in particular, whether the policy mix is constructed to address the specific issue at hand—are associated with discretionary policy crafting power given to bureaucracies across OECD economies.

Of course, as discussed in the first section above, politicians will often wish for policymaking to remain firmly in the political realm. Modern political economic theory is filled with reasons for why sensible economic reforms may not come to fruition, particularly in the case of policies with distributive effects (Blinder 1997; Alesina and Tabellini 2007). Additionally, there are good reasons to believe democracies may want elected representatives involved in industrial policymaking rather than unelected civil servants.

We conclude this section by noting that industrial policy almost certainly requires additional and ongoing investment in bureaucratic capacity. For one, states are out of practice when it comes to the sweeping, capacity-intensive forms of industrial policy that are emerging across the post-COVID world. In all shapes and sizes, however, industrial policies have expanded rapidly across the globe (Juhász et al. 2022). Likewise, the existing bureaucratic capacity to perform industrial policy is low; underinvestment is seen in the OECD, in presidential systems, and in European democracies (Bednar and Lewis 2024; Fernández-i-Marín et al. 2024a, b). The history of industrial policy shows us that their success hinges critically on bureaucratic capability and autonomy. Yet, state capacity does not fall from the sky, nor is it static. That is, good industrial policies invest in bureaucracies. We will go so far as to make this claim: if industrial policies are to succeed, repeated investments in administrative capacity are a *must*.

[3] In keeping with the arguments in the previous section, some authors argue it is optimal to delegate design when policies have concentrated political stakes and are prone to time-inconsistency issues (Alesina and Tabellini 2007, 2008).

Embeddedness and Information

Implementing industrial policy not only requires a high-quality bureaucracy, but one that continually interacts, negotiates, and exchanges information with industry and stakeholders more broadly. Successful industrial policy is not passively deployed from commanding heights, nor is it inert. Rather, it is informed by and executed through continual interactions with market participants. Civil servants are not omnipotent, and uncovering the nature of market failures requires ongoing input from those with domain expertise. Firms may face a myriad of bottlenecks, including lack of finance, difficulties procuring land, skill shortages, and administrative barriers. New policies can bring heightened prominence to constraints: for example, across the United States, the rollout of the investment subsidies for clean energy in the Inflation Reduction Act of 2022 brought to the fore the concern that local permitting issues could hinder or block desired investment in new factories (Brouns 2023). Likewise, debates surrounding industrial policy often involve the informational limits of bureaucracies (Maloney and Nayyar 2018). This section examines the relationship between bureaucracies and private actors as a source of information exchange.

The idea of *embeddedness*—the extent to which bureaucracies have connections with the business sector—was developed by Peter Evans (1995) to describe a key feature of developmental bureaucracies. Evans explains how South Korea's dynamic random-access memory (DRAM) project, led by Korea's Electronics and Telecommunications Institute (ETRI), was not undertaken by the state in isolation. On the contrary, the *chaebol*, large Korean business groups, were incorporated into the decision-making process, including planning, implementation, and collaboration between government and private sector researchers. From South Korea's monthly export promotion meetings to Japan's use of deliberation councils, East Asian states purposefully cultivated embeddedness by institutionalizing interactions between firms and bureaucracy. At its height, East Asian industrial policy was marked by webs of collaboration between bureaucratic agencies and the private sector (Birdsall et al. 1993; Doner, Ritchie, and Slater 2005).

Embeddedness actually informs industrial policy practice across the income distribution, such as the US Advanced Research Projects Agency (ARPA) model or Peru's *Mesas Ejecutivas* (known as *mesas* or ME) (Juhász, Lane, and Rodrik 2023). The case of *mesas* is particularly instructive in how durable industrial policy bodies can be built in lower-capacity environments. Established in 2015, *mesas* are regular, weekly private-public working groups dedicated to solving sector-specific policy. Ghezzi (2017) explains how *mesas* help identify market and coordination failures and, importantly, can triage and expedite solutions across government bureaucracies. As coordinating bodies, institutions like the *mesas* have a low fiscal footprint and, in fact, were implemented as an alternative to costly external consultations (Ministry of Production 2016).

In seminal qualitative work, Breznitz (2007) provides a positive political economy description of how three small open economies—Israel, Ireland, and Taiwan—chose different forms of embedded bureaucracy to promote moves into

high technology industries. These embedded agencies were instrumental to each country entering fast-paced, competitive information technology markets, yet they did so with wide institutional variation. Where the Taiwanese state was directly involved in the industrial research and development process (for example, the Industrial Technology Research Institution), Irish agencies took a more advisory and advocacy role (for example, the National Software Directorate). These features shaped both the industrial policies that were chosen and where countries entered fragmented, hi-tech supply chains. Hence, there was no single recipe for success, but numerous ways in which small, open economies deployed embedded institutions to coordinate entry into dynamic, global industries.

Among the other benefits of embeddedness, it facilitates the flow of information between bureaucracy and industry. Doing so is essential for industrial policy given fundamental *informational asymmetries* between bureaucrats (principals) and the firms with which they interact (agents). Consider a green industrial policy, where a public agency subsidizes risky projects that, if successful, would generate both private and social benefits. How should the agency design conditional subsidies? Meunier and Ponssard (2024) show that when firms and public agencies have symmetric information about the probability of a project's success, rewarding success is optimal. However, under asymmetric information, where only the firm knows its probability of success, failure should be rewarded (!)—as it mitigates the windfall profit that arises when an agency subsidizes projects that would have received financing absent the subsidy. This insight speaks directly to the experience of the French Agency for Ecological Transition (ADEME), a public agency monitoring innovative activities for the energy transition funded by the Investments for the Future Programme. At the outset, ADEME used flat subsidies, but evidence of windfall profits quickly emerged in some projects. Therefore, the agency introduced "repayable advances," which are subsidies that need to be paid back in the case of success—that is, they are subsidies for failure.

Such informational asymmetries are not unique to industrial policy, but are inherent in many settings, particularly regulation and antitrust. These problems have inspired a storied literature on regulatory policy design and incentive mechanisms (for an overview, see Baron 1989; Armstrong and Sappington 2007). This literature highlights the importance of considering the institutional constraints bureaucracies face and the hard work necessary for designing policy under imperfect information. Depending on the challenge the government is trying to solve, embeddedness with the private sector may be an alternative to designing mechanisms that take the informational asymmetry as fixed, as Sabel (2004) and Rodrik (2014) argue. This outcome is particularly likely where the principal may not know what needs to be done to achieve public goals, and instead, the government and private sector work together in a discovery process. The *mesas* above offer one such example.

Embeddedness, however, can cut both ways. Dense links between the state and industry also introduce the potential for capture and predation. Among other things, embeddedness requires the bureaucratic independence and autonomy

described above. This balancing act is what Evans (1995) famously called "embedded autonomy," where both are required for industrial policy to succeed. Autonomy without embeddedness risks flying blind and constructing and deploying industrial policy in isolation from essential stakeholders. Embeddedness without autonomy risks incoherence and policies guided by private interests.

What then determines investment in state capacity, especially autonomous and embedded bureaucracies? Ultimately, these are political decisions. As Thailand's example shows, the political environment is key to understanding not only what industrial policy is chosen but also whether the accompanying investments in state capacity take place.

Conclusion

Variation in the practice of industrial policy is as much political as it is economic. Market failures and economic constraints often govern how economists view optimal policy, yet the political forces—especially the two dimensions of political constraints and state capacity—influence heavily how these interventions are realized. This conclusion is uncontroversial through the lens of modern political economy; in fact, it is the raison d'être of positive political economics (Persson and Tabellini 2002; Drazen 2000). Yet, in the specific realm of industrial policy, economists have paid far too little attention to the political conditions that have supported good industrial policy. If the empirical literature of industrial policy is far underdeveloped relative to practice, the positive political economy of industrial policy is even more anemic.

This paper has considered two prominent governance constraints to industrial policy, and our analysis highlights that productive industrial policy can and has been deployed within these constraints in various contexts. While working within these governance constraints is, in our view, necessary for industrial policy to succeed, it is not a sufficient condition. Thus, our take offers a pragmatic and carefully optimistic view of the possibility of overcoming government failure and the challenges of governance. We view these as important ingredients in producing more successful industrial policies.

■ *We thank the editorial board, Bentley Allan, Heather Boushey, Cristina Caffarra, George Dibb, Claudio Ferraz, Mark Lane, Weijia Lee, Jonas Meckling, Jonas Nahm, Dani Rodrik, Todd Tucker, Eric Verhoogen, and Katherine Wagner for helpful comments and conversations. We thank Lottie Field, Mikhael Gaster, Saumya Joshi, and Esha Vaze for their excellent research assistance and input. We are indebted to Gian Aswin Chansrichawla for guiding our focus to Thailand as a candidate case study, and for helpful conversations and references about the Thai case. Funding from the Social Sciences and Humanities Research Council of Canada and the Alfred P. Sloan Foundation is gratefully acknowledged.*

References

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2016. "Transition to Clean Technology." *Journal of Political Economy* 124 (1): 52–104.

Acemoglu, Daron, and James A. Robinson. 2013. "Economics versus Politics: Pitfalls of Policy Advice." *Journal of Economic Perspectives* 27 (2): 173–92.

Alesina, Alberto, and Guido Tabellini. 2007. "Bureaucrats or Politicians? Part I: A Single Policy Task." *American Economic Review* 97 (1): 169–79.

Alesina, Alberto, and Guido Tabellini. 2008. "Bureaucrats or Politicians? Part II: Multiple Policy Tasks." *Journal Of Public Economics* 92 (3–4): 426–47.

Armstrong, Mark, and David E. M. Sappington. 2007. "Recent Developments in the Theory of Regulation." In *Handbook of Industrial Organization*, edited by Mark Armstrong and Robert H. Porter, Vol. 3, 1557–1700. Amsterdam: Elsevier.

Ball, Jeffrey. 2018. "Why Carbon Pricing Isn't Working: Good Idea in Theory, Failing in Practice." *Foreign Affairs* 97 (4): 134–46.

Bardhan, Pranab K. 1971. "On Optimum Subsidy to a Learning Industry: An Aspect of the Theory of Infant-Industry Protection." *International Economic Review* 12 (1): 54–70.

Baron, David P. 1989. "Design of Regulatory Mechanisms and Institutions." In *Handbook of Industrial Organization*, edited by Robert Schmalensee and Robert D. Willig, Vol. 2, 1347–1447. Amsterdam: Elsevier.

Basseches, Joshua A., Rebecca Bromley-Trujillo, Maxwell T. Boykoff, Trevor Culhane, Galen Hall, Noel Healy, David J. Hess, et al. 2022. "Climate Policy Conflict in the U.S. States: A Critical Review and Way Forward." *Climatic Change* 170 (3): 32.

Bayer, Patrick, and Johannes Urpelainen. 2016. "It Is All about Political Incentives: Democracy and the Renewable Feed-in Tariff." *Journal Of Politics* 78 (2): 603–19.

Bednar, Nicholas R., and David E. Lewis. 2024. "Presidential Investment in the Administrative State." *American Political Science Review* 118 (1): 442–57.

Bersch, Katherine, and Francis Fukuyama. 2023. "Defining Bureaucratic Autonomy." *Annual Review of Political Science* 26: 213–32.

Besley, Timothy, Robin Burgess, Adnan Khan, and Guo Xu. 2022. "Bureaucracy and Development." *Annual Review of Economics* 14: 397–424.

Besley, Timothy, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton: Princeton University Press.

Biber, Eric. 2013. "Cultivating a Green Political Landscape: Lessons for Climate Change Policy from the Defeat of California's Proposition 23." *Vanderbilt Law Review* 66 (2): 399–462.

Birdsall, Nancy M., Jose Edgardo L. Campos, Chang-Shik Kim, W. Max Corden, Lawrence MacDonald, Howard Pack, John Page, Richard Sabor, and Joseph E. Stiglitz. 1993. *The East Asian Miracle: Economic Growth and Public Policy*. Washington, DC: Oxford University Press on behalf of the World Bank.

Blanchard, Olivier, Christian Gollier, and Jean Tirole. 2023. "The Portfolio of Economic Policies Needed to Fight Climate Change." *Annual Review of Economics* 15: 689–722.

Blinder, Alan S. 1997. "Is Government Too Political?" *Foreign Affairs* 76 (6): 115–26.

Boyce, James K. 1993. *The Political Economy of Growth and Improverishment in the Marcos Era*. Quezon City: Ateneo de Manila University Press.

Breznitz, Dan. 2007. *Innovation and the State: Political Choice and Strategies for Growth in Israel, Taiwan, and Ireland*. London: Yale University Press.

Brouns, Zoë. 2023. "Implementation Bottlenecks: Federal Talent Will Drive IIJA and IRA Success." Federation of American Scientists, July 27. https://fas.org/publication/implementation-bottlenecks-federal-talent-will-drive-iija-and-ira-success/ (accessed on April 27, 2024).

Brulle, Robert, and Christian Downie. 2022. "Following the Money: Trade Associations, Political Activity and Climate Change." *Climatic Change* 175 (3–4): 11.

Bueno de Mesquita, Ethan. 2016. *Political Economy for Public Policy*. Princeton: Princeton University Press.

Chibber, Vivek. 2002. "Bureaucratic Rationality and the Developmental State." *American Journal of Sociology* 107 (4): 951–89.

Christensen, Scott, David Dollar, Ammar Siamwalla, and Pakorn Vichyanond. 1990. *Thailand—The Lessons of East Asia: The Institutional and Political Underpinnings of Growth*. Washington, DC: World Bank.

Christensen, Scott R., and Ammar Siamwalla. 1993. *Beyond Patronage: Tasks for the Thai State: Policy Analysis*. Chonburi: Thailand Development Research Institute.

Climate Leadership Council. n.d. "Economists' Statement on Carbon Dividends Organized by the Climate Leadership Council." https://www.econstatement.org (accessed August 16, 2024).

Colmer, Jonathan, Ralf Martin, Mirabelle Muûls, and Ulrich J. Wagner. 2024. "Does Pricing Carbon Mitigate Climate Change? Firm-Level Evidence from the European Union Emissions Trading System." *Review of Economic Studies*.

Coppedge, Michael, John Gerring, Carl H. Knutsen, Staffan I. Lindberg, Joe Teorell, David Altman, Fabio Angiolillo, et al. 2024. "V-Dem Dataset V14." Varieties of Democracy (V-Dem) Project. https://v-dem.net/data/the-v-dem-dataset/country-yearv-dem-fullothers-v14/.

Cory, Jared, Michael Lerner, and Iain Osgood. 2021. "Supply Chain Linkages and the Extended Carbon Coalition." *American Journal of Political Science* 65 (1): 69–87.

Crouch, Harold. 1984. *Domestic Political Structures and Regional Economic Co-operation*. Pasir Panjang: Institute of Southeast Asian Studies.

Cullenward, Danny, and David G. Victor. 2020. *Making Climate Policy Work*. Cambridge, UK: Polity Press.

Dixit, Avinash. 1997. "Economists as Advisers to Politicians and to Society." *Economics and Politics* 9 (3): 225–30.

Dixit, Avinash. 2009. "Governance Institutions and Economic Activity." *American Economic Review* 99 (1): 5–24.

Doner, Richard F., and Ansil Ramsay. 1997. "Competitive Clientelism and Economic Governance: The Case of Thailand." In *Business and the State in Developing Countries*, edited by Sylvia Maxfield and Ben Ross Schneider, 237–76. Ithaca: Cornell University Press.

Doner, Richard F., and Anek Laothamatas. 1994. "Thailand: Economic and Political Gradualism." In *Voting for Reform: Democracy, Political Liberalization, and Economic Adjustment*, edited by Stephen Haggard and Steven B. Webb, 411–52. New York: Oxford University Press on behalf of the World Bank.

Doner, Richard F., and Ansil Ramsay. 2000. "Rent-Seeking and Economic Development in Thailand." In *Rents, Rent-Seeking and Economic Development: Theory and Evidence in Asia*, edited by Mushtaq H. Khan and Kwame Sundaram Jomo, 145–81. Cambridge, UK: Cambridge University Press.

Doner, Richard F., Bryan K. Ritchie, and Dan Slater. 2005. "Systemic Vulnerability and the Origins of Developmental States: Northeast and Southeast Asia in Comparative Perspective." *International Organization* 59 (2): 327–61.

Drazen, Allan. 2000. *Political Economy in Macroeconomics*. Princeton: Princeton University Press.

End, Nicolas. 2023. "Big Brother Is Also Being Watched: Measuring Fiscal Credibility." *Journal of Macroeconomics* 77: 103548.

Evans, Peter B. 1995. *Embedded Autonomy: States and Industrial Transformation*. Princeton: Princeton University Press.

Evenett, Simon J., and Johannes Fritz. 2022. *The Global Trade Alert Database Handbook*. St. Gallen, Switzerland: Global Trade Alert.

Fernández-i-Marín, Xavier, Markus Hinterleitner, Christoph Knill, and Yves Steinebach. 2024a. "Bureaucratic Overburdening in Advanced Democracies." *Public Administration Review* 84 (4): 696–709.

Fernández-i-Marín, Xavier, Markus Hinterleitner, Christoph Knill, and Yves Steinebach. 2024b. "Policy Growth, Implementation Capacities, and the Effect on Policy Performance." *Governance* 37 (3): 927–45.

Fernández-i-Marín, Xavier, Christoph Knill, and Yves Steinebach. 2021. "Studying Policy Design Quality in Comparative Perspective." *American Political Science Review* 115 (3): 931–47.

Field, Lottie. 2024. "The Political Economy of Industrial Development Organisations: Are They Run by Politicians or Bureaucrats?." SocArXiv. September 17. https://osf.io/preprints/socarxiv/a4ker.

Furceri, Davide, Michael Ganslmeier, and Jonathan Ostry. 2023. "Are Climate Change Policies Politically Costly?" *Energy Policy* 178: 113575.

Ghezzi, Piero. 2017. "Mesas Ejecutivas in Peru: Lessons for Productive Development Policies." *Global Policy* 8 (3): 369–80.

Global Trade Alert. n.d. *Data and Methodology: Download Data*. https://www.globaltradealert.org/data_extraction.

Haggard, Stephan. 1990. *Pathways from the Periphery: The Politics of Growth in the Newly Industrializing Countries*. Ithaca: Cornell University Press.

Hamilton, Alexander. 1791. "Report on Manufactures." In *The Early Republic and Antebellum America:*

An Encyclopedia of Social, Political, Cultural, and Economic History, edited by Christopher G. Bates. Milton: Taylor and Francis, 2010.

Harrison, Ann. Forthcoming. "What Makes Industrial Policy Work?." In Rethinking Economic Policy: Steering Structural Change, edited by Pierre-Olivier Gourinchas, Maurice Obstfeld, and Petia Topalova. Cambridge, MA: MIT Press.

Harrison, Ann, Leslie A. Martin, and Shanthi Nataraj. 2017. "Green Industrial Policy in Emerging Markets." *Annual Review of Resource Economics* 9: 253–74.

Harrison, Ann, and Andrés Rodríguez-Clare. 2010. "Trade, Foreign Investment, and Industrial Policy for Developing Countries." In *Handbook of Development Economics*, edited by Dani Rodrik and Mark Rosenzweig, Vol. 5, 4039–4214. Oxford: North-Holland.

Harrison, Kathryn. 2010. "The Comparative Politics of Carbon Taxation." *Annual Review of Law and Social Science* 6 (1): 507–29. https://doi.org/10.1146/annurev.lawsocsci.093008.131545

Herderschee, Han. 1993. "Incentives for Exports: The Case of Thailand." *ASEAN Economic Bulletin* 9 (3): 348–63.

Hess, David J. 2014. "Sustainability Transitions: A Political Coalition Perspective." *Research Policy* 43 (2): 278–83.

Hewison, Kevin. 1987. "National Interests and Economic Downturn: Thailand." In *South East Asia in the 1980s: The Politics of Economic Crisis*, edited by Richard Robison, Kevin Hewison, and Richard Higgott, 52–79. Sydney: Allen and Unwin.

Irwin, Douglas A. 1991. "Retrospectives: Challenges to Free Trade." *Journal of Economic Perspectives* 5 (2): 201–08.

Jenkins, Jesse D. 2014. "Political Economy Constraints on Carbon Pricing Policies: What Are the Implications for Economic Efficiency, Environmental Efficacy, and Climate Policy Design?" *Energy Policy* 69: 467–77.

Jenkins, Jesse D., and Valerie J. Karplus. 2017. "Carbon Pricing under Political Constraints: Insights for Accelerating Clean Energy Transitions." In *The Political Economy of Clean Energy Transitions*, edited by Douglas Arent, Channing Arndt, Mackay Miller, Finn Tarp, and Owen Zinaman, 39–59. Oxford: Oxford University Press.

Johnson, Chalmers. 1982. *MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925–1975*. Stanford: Stanford University Press.

Juhász, Réka, Nathan Lane, Emily Oehlsen, and Veronica Pérez. 2022. "The Who, What, When, and How of Industrial Policy: A Text-Based Approach." https://doi.org/10.31235/osf.io/uyxh9.

Juhász, Réka, Nathan Lane, and Dani Rodrik. 2023. "The New Economics of Industrial Policy." *Annual Review of Economics* 16: 213–42.

Juhász, Réka, and Nathan Lane. 2024. *Data and Code for: "The Political Economy of Industrial Policy."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E209090V1.

Karapin, Roger. 2016. *Political Opportunities for Climate Policy: California, New York, and the Federal Government*. Cambridge, UK: Cambridge University Press.

Killick, Tony. 2010. *Development Economics in Action: A Study of Economic Policies in Ghana*. 2nd ed. London: Routledge.

Kim, Hyung-A. 2011. "State Building: The Military Junta's Path to Modernity through Administrative Reforms." In *The Park Chung Hee Era: The Transformation of South Korea*, edited by Byung-Kook Kim and Ezra F. Vogel, 85–112. Cambridge, MA: Harvard University Press.

Kim, Sung Eun, Johannes Urpelainen, and Joonseok Yang. 2016. "Electric Utilities and American Climate Policy: Lobbying by Expected Winners and Losers." *Journal of Public Policy* 36 (2): 251–75.

Klenert, David, Linus Mattauch, Emmanuel Combet, Ottmar Edenhofer, Cameron Hepburn, Ryan Rafaty, and Nicholas Stern. 2018. "Making Carbon Pricing Work for Citizens." *Nature Climate Change* 8 (8): 669–77.

Krueger, Anne O. 1990. "Government Failures in Development." *Journal of Economic Perspectives* 4 (3): 9–23.

Laothamatas, Anek. 1994. "From Clientelism to Partnership: Business-Government Relations in Thailand." In *Business and Government in Industrializing Asia*, edited by Andrew MacIntyre, 194–215. Ithaca: Cornell University Press.

Larch, Martin, and Thomas Braendle. 2018. "Independent Fiscal Councils: Neglected Siblings of Independent Central Banks? An EU Perspective." *Journal of Common Market Studies* 56 (2): 267–83.

Maloney, William F., and Gaurav Nayyar. 2018. "Industrial Policy, Information, and Government

Capacity." *World Bank Research Observer* 33 (2): 189–217.
Marquardt, Jens, M. Cecilia Oliveira, and Markus Lederer. 2022. "Same, Same but Different? How Democratically Elected Right-Wing Populists Shape Climate Change Policymaking." *Environmental Politics* 31 (5): 777–800.
Marshall, Alfred. 1975. *The Early Economic Writings of Alfred Marshall, 1867–1890*. Edited by John K. Whitaker. London: Macmillan for the Royal Economic Society.
Matsuyama, Kiminori. 1990. "Perfect Equilibria in a Trade Liberalization Game." *American Economic Review* 80 (3): 480–92.
Meckling, Jonas, and Valerie J. Karplus. 2023. "Political Strategies for Climate and Environmental Solutions." *Nature Sustainability* 6 (7): 742–51.
Meckling, Jonas, Nina Kelsey, Eric Biber, and John Zysman. 2015. "Winning Coalitions for Climate Policy." *Science* 349 (6253): 1170–71.
Meckling, Jonas, and Jonas Nahm. 2018. "The Power of Process: State Capacity and Climate Policy." *Governance* 31 (4): 741–57.
Meckling, Jonas, Thomas Sterner, and Gernot Wagner. 2017. "Policy Sequencing toward Decarbonization." *Nature Energy* 2 (12): 918–22.
Meng, Kyle C., and Ashwin Rode. 2019. "The Social Cost of Lobbying over Climate Policy." *Nature Climate Change* 9 (6): 472–76.
Meunier, Guy, and Jean-Pierre Ponssard. 2024. "Green Industrial Policy, Information Asymmetry, and Repayable Advance." *Journal of Public Economic Theory* 26 (1): 12668.
Mildenberger, Matto. 2020. *Carbon Captured: How Business and Labor Control Climate Politics*. Cambridge, MA: MIT Press.
Mill, John Stuart. 1848. *Principles of Political Economy*. Vol. 2. London: John W. Parker.
Mill, John Stuart. 1868. "To Archibald Michie, of Victoria." In *The Letters of John Stuart Mill, 1910*, Vol. 2, edited by Hugh S. R. Elliot. London: Longmans Green and Co.
Ministry of Production. 2016. *Mesas Ejecutivas: A New Tool for Productive Diversification*. Lima: Ministry of Production of Peru.
Muscat, Robert J. 1994. *The Fifth Tiger: A Study Of Thai Development Policy*. Helsinki: United Nations University Press.
Pahle, Michael, Oliver Tietjen, Sebastian Osorio, Florian Egli, Bjarne Steffen, Tanja S. Schmidt, and Ottmar Edenhofer. 2022. "Safeguarding the Energy Transition against Political Backlash to Carbon Markets." *Nature Energy* 7 (3): 290–96.
Persson, Torsten, and Guido Tabellini. 1990. *Macroeconomic Policy, Credibility And Politics*. London: Harwood Academic.
Persson, Torsten, and Guido Tabellini. 1999. "Political Economics and Macroeconomic Policy." In *Handbook of Macroeconomics*, Vol. 1C, edited by John B. Taylor and Michael Woodford, 1397–1482. Amsterdam: Elsevier.
Persson, Torsten, and Guido Tabellini. 2002. "Political Economics and Public Finance." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein, Vol. 3, 1549–1659. Amsterdam: Elsevier.
Poapongsakorn, Nipon, and Belinda Fuller. 1997. "Thailand's Development Experience from the Economic System Perspective: Open Politics and Industrial Activism." In *East Asian Development Experience: Economic System Approach and Its Applicability*, edited by Susumu Sambommatsu and Toru Yanagihara, 466–518. Tokyo: Institute of Developing Economies.
Quirion, Philippe. 2021. "Tradable Instruments to Fight Climate Change: A Disappointing Outcome." *Wires Climate Change* 12 (3): 705.
Rathi, Akshat. 2024. *Climate Capitalism: Winning the Global Race to Zero Emissions and Solving the Crisis of Our Age*. London: John Murray Press.
Robinson, James A. 2010. "Industrial Policy and Development: A Political Economy Perspective." In *Annual World Bank Conference on Development Economics—Global*, edited by Justin Yifu Lin and Boris Pleskovic, 61–79. Washington, DC: World Bank.
Rock, Michael T. 1994. "Transitional Democracies and the Shift to Export-Led Industrialization: Lessons from Thailand." *Studies in Comparative International Development* 29: 18–37.
Rock, Michael T. 1995. "Thai Industrial Policy: How Irrelevant was it to Export Success?" *Journal Of International Development* 7 (5): 745–57.
Rodrik, Dani. 1995. "Getting Interventions Right: How South Korea and Taiwan Grew Rich." *Economic Policy* 10 (20): 53–107.

Rodrik, Dani. 2008. "Normalizing Industrial Policy." World Bank Commission on Growth and Development Working Paper 3.

Rodrik, Dani. 2014. "When Ideas Trump Interests: Preferences, Worldviews, and Policy Innovations." *Journal of Economic Perspectives* 28 (1): 189–208.

Sabel, Charles F. 2004. "Beyond Principal-Agent Governance: Experimentalist Organizations, Learning and Accountability." In *De staat van de democratie: Democratie voorbij de staat*, edited by Ewald R. Engelen and Monika Sie Dhian Ho, 173–95. Amsterdam: Amsterdam University Press.

Sato, Misato, Ryan Rafaty, Raphael Calel, and Michael Grubb. 2022. "Allocation, Allocation, Allocation! The Political Economy of the Development of the European Union Emissions Trading System." *WIREs Climate Change* 13 (5): 796.

Staiger, Robert. 1995. "International Rules and Institutions for Trade Policy." In *Handbook of International Economics*, Vol. 3, edited by Gene M. Grossman and Kenneth Rogoff, 1495–1551. Amsterdam: Elsevier.

Stern, Nicholas. 2022. "A Time for Action on Climate Change and a Time for Change in Economics." *Economic Journal* 132 (644): 1259–89.

Taussig, Frank William. 1914. *The Tariff History of the United States*. New York: GP Putnam's Sons.

Tornell, Aaron. 1991. "Time Inconsistency of Protectionist Programs." *Quarterly Journal of Economics* 106 (3): 963–74.

Vihma, Antto, Gunilla Reischl, and Astrid Nonbo Andersen. 2021. "A Climate Backlash: Comparing Populist Parties' Climate Policies in Denmark, Finland, and Sweden." *Journal of Environment and Development* 30 (3): 219–39.

Vormedal, Irja, and Jonas Meckling. 2023. "How Foes Become Allies: The Shifting Role of Business in Climate Politics." *Policy Sciences* 57 (1): 101–24.

Wade, Robert H. 1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*. 2nd ed. Princeton: Princeton University Press.

Warr, Peter G. 1993. "The Thai Economy." In *The Thai Economy in Transition*, edited by Peter G. Warr, 1–80. Cambridge: Cambridge University Press.

Westphal, Larry E. 1990. "Industrial Policy in an Export-Propelled Economy: Lessons from South Korea's Experience." *Journal of Economic Perspectives* 4 (3): 41–59.

Wiboonchutikula, Paitoon. 1987. "Second Phase Import Substitution in Thailand." Bangkok: Thailand Development Research Institute.

World Bank. 2024. *State and Trends of Carbon Pricing 2024*. Washington, DC: World Bank.

World Bank. 2024. *World Bank Analytical Classifications*. Washington, DC: World Bank. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-group.

Zingales, Luigi. 2020. "The Political Limits of Economics." *AEA Papers and Proceedings* 110: 378–82.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 55–80

Industrial Policy: Lessons from Shipbuilding

Panle Jia Barwick, Myrto Kalouptsidi, and Nahim Bin Zahur

Industrial policy refers to a government agenda to shape industry structure by promoting certain industries or sectors. Although casual observation suggests that industrial policy can boost sectoral growth, researchers and policymakers have not yet mastered predicting or evaluating the efficacy of different types of government interventions, nor how to measure the overall short-run and long-run welfare effects. In this article, we focus on one particular example of industrial policy, which we believe serves as a revealing case study: government support for shipbuilding in general, and China's industrial policy to support shipbuilding in particular.

Shipbuilding has been historically a classic target of industrial policy, pursued by several countries that devised national programs for heavy industrialization, such as Japan in the 1950s and South Korea in the 1980s. Interestingly, shipbuilding has now entered industrial policy agendas in both the European Union (Folkman 2024) and the United States (Forooher 2024), with calls for reshoring shipbuilding production. We begin this essay with an overview of global production patterns in shipbuilding, and how these patterns have shifted in the last century or so. The rich and tumultuous history of shipbuilding presents puzzles and leaves us with open questions: Why have governments subsidized shipbuilding throughout history? This is not obvious at first glance: the global market for sales of newly built ships is about $120 billion annually, which by global standards is not large. Was industrial policy

■ *Panle Jia Barwick is Professor of Economics, University of Wisconsin, Madison, Wisconsin. Myrto Kalouptsidi is Professor of Economics, Harvard University, Cambridge, Massachusetts. Nahim Bin Zahur is Assistant Professor of Economics, Queen's University, Kingston, Ontario, Canada. Their email addresses are pbarwick@wisc.edu, myrto@g.harvard.edu, and nz17@queensu.ca.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.55.

successful? Although several national shipbuilding programs, such as in Japan or South Korea, seem to be behind tremendous sectoral growth, others, like support for shipbuilding in the United States, have seemingly failed.

We offer a brief overview of the two primary research approaches that have been used to study industrial policy: (1) a largely reduced-form empirical analysis has explored the impact of specific policies on available outcomes (primarily output growth and employment) exploiting natural or quasi-experimental variation; and (2) work based on strategic trade theory. However, we focus primarily on a third approach: structural modeling of the industry, and in particular an application of this approach to China's subsidies in support of shipbuilding. China is of particular interest more generally, given its strong advocacy for industrial policy and its numerous trade conflicts with many countries across various industries. In recent years, China's government has been explicitly targeting sectors with the goal of turning its firms into world leaders. We outline the combination of structural modeling and data work needed to tackle the challenge of measuring explicit and implicit subsidies; our answer to how the global industry would have evolved if China had not subsidized shipbuilding, and in particular, how China's subsidies for shipbuilding affected shipbuyers, shipping costs, and world trade; the low economic payoffs that China received from its subsidies in the beginning and how it altered the policy design over time; and the broader economic and geopolitical reasons China may have chosen to subsidize shipbuilding. In conclusion, we offer some provisional lessons for researchers and policymakers considering industrial policies, whether in shipbuilding or in other industries.

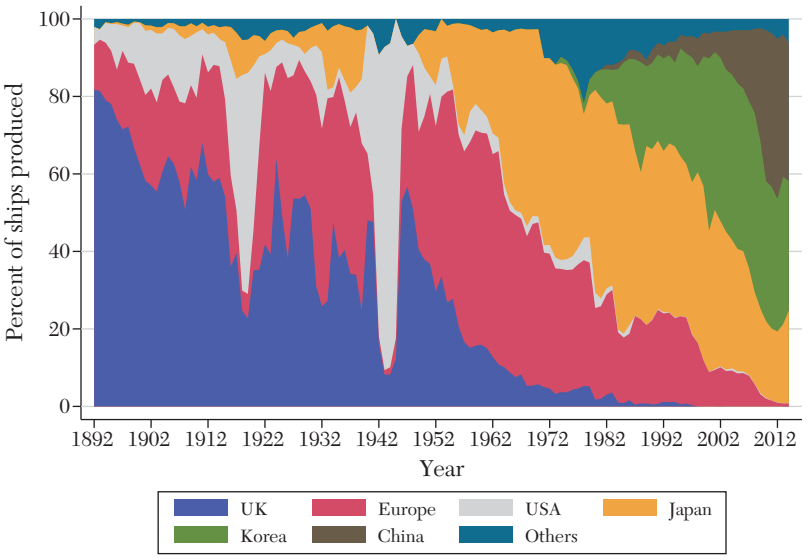
A Brief History of Global Shipbuilding

The history of shipbuilding is as tumultuous as the seas themselves. Shipbuilding has always held an allure for governments, in its real and perceived interactions with industrialization, maritime trade, and military strength.

The Changing Leader

Figure 1 shows the succession of countries as the world's dominant shipbuilding nation. The United Kingdom holds the lion's share of the industry for the better part of the nineteenth and twentieth centuries, fending off competition from other Western European economies (mainly Germany and Scandinavia) at times. After World War II, it is swiftly overtaken by Japan, which prevails as a world leader until the 1980s, when South Korea dominates the global market. In the mid-2000s, China has staked its claim. This succession is made even more dramatic by the notorious "shipbuilding cycles," a succession of booms and busts (Kalouptsidi 2014). Building large ships takes years, but demand for ships is governed by volatile macroeconomic fluctuations. When demand spikes but shipbuilding capacity is fixed and sluggish, shipping rates skyrocket, and shipbuilders pile up orders for new ships. Ship prices soar. But in the bust, shipbuilding capacity idles, and prices hit rock bottom.

Figure 1
Share of Commercial Ships Produced by Each Country, 1892–2014



Source: Data for 1892–1997 were obtained from historical issues of the World Fleet Statistics published by Lloyd's Register, while the data from 1998 onwards are based on Clarksons data. We group together all European shipbuilding countries except for the United Kingdom under "Europe."
Note: This figure plots the world market share in terms of the number of ships delivered from 1892 to 2014 for the major ship producing countries.

Why have certain countries dominated global shipbuilding? The basic economics of shipbuilding certainly plays a role in fostering the high concentration of production in certain locations. Ship production requires a base on land, easy access to water (sea or rivers), materials such as steel and engines, and (skilled) labor. Countries that dominate shipbuilding often have a strong maritime tradition and have often been important players in global shipping and trade. In addition, the shipbuilding leaders are often undergoing a phase of heavy industrialization. But a quick inspection of the history of shipbuilding also reveals the decisive role that industrial policy, of different forms and with different motivations, has played in shaping the global market for ships.[1] The United Kingdom's leading position in shipbuilding, which lasted for decades up until the 1950s, initially stemmed from access to cheap iron and steel in the late 1800s (Hanlon 2020).[2] From the late nineteenth century onwards, though, UK dominance in shipbuilding was in large part due to

[1] The major types of ships currently produced include containerships, (oil) tankers, bulk carriers, as well as more niche products like cruise ships, liquefied natural gas carriers, and "Ro-Ro's," which are ships that allow vehicles to be rolled on and off the ship.
[2] The United States dominated wood shipbuilding in the first half of 1800s until wood ships were replaced by metal ships in the 1850s.

its strong maritime trade: UK trade flows dominated global trade, and the British Empire needed ships to execute this trade volume with its colonies and other trading partners. The United Kingdom also benefitted from strong integration between its shipowners and shipbuilders: British owners almost never bought foreign ships and instead maintained close relationships with domestic shipbuilders (Pollard 1957; Stopford 2009). Access to a large and loyal home market allowed British shipbuilders to enjoy economies of scale, both internal and external: it facilitated the formation of a skilled pool of labor, enabled shipbuilders to specialize in producing specific ship types, and reduced their exposure to shipbuilding boom-and-bust cycles, allowing them to utilize their capacity more effectively (Pollard 1957; Hanlon 2020).

While government policy may have played a role in explaining the United Kingdom's leading position in global trade and shipping (which in turn boosted its shipbuilding sector), it is unclear if the United Kingdom directly subsidized the shipbuilding sector itself during this earlier period. But this would turn out to be a historical anomaly. Every major shipbuilding player that subsequently emerged did so at least in part with the aid of industrial policy.

The first major challenge to the United Kingdom's hegemony came from other Western European countries with strong maritime traditions, such as Scandinavian countries. Initially, many of these competitors were unable to compete globally with the significantly more productive UK shipyards and instead were propped up through the use of subsidies and various protective policies (Pollard 1957). Over the first half of the twentieth century, other European shipbuilders built up a substantial market share in ship production, while British shipbuilding declined. Another part of this shift was that British shipbuilders failed to transition away from the "craft" style of production popular in the early twentieth century to the more "industrial" style of production that became more common in the mid-twentieth century as ship sizes increased, ships became more standardized, and ship production became more capital-intensive; that said, the protectionist policies used by Britain's competitors (including generous subsidies to shipbuilders) also contributed to Britain's decline (Lorenz 1991).

During World War I and II, the United States both assembled and dismantled massive shipbuilding capacity in a very short time span (Thornton and Thompson 2001; Thompson 2001). Beyond these two periods, the United States was never globally competitive as a ship producer, despite the fact that the United States had long been a proponent of industrial policy in shipbuilding. For instance, the Jones Act prohibits foreign vessels from transporting goods between two US ports, and the "construction differential subsidies" that were provided to US shipbuilders until the 1980s amounted to between 30 and 50 percent of the cost (Stopford 2009). Once high wartime demand for ships had evaporated, US shipyards, with their much higher costs, were unable to compete commercially with European and Japanese yards (Stopford 2009). That said, the US example shows that a massive shipbuilding program can be set up at an astonishing speed, but there is no guarantee that such a program can be sustained in the long run.

After World War II, Japan developed national programs for its shipbuilding industry as part of its efforts to rebuild its industrial base alongside several other heavy industries, such as steel and coal (which were viewed as complementary), chemical fertilizers, petrochemicals, and automobiles (Okuno-Fujiwara 1991; Flath 2022). An island nation with a very strong maritime tradition, Japan swiftly became the world's dominant ship producer. Through a series of interventions, including subsidized financing, export credits, and protectionist measures, a number of large Japanese conglomerates—Mitsubishi, Kawasaki, Sasebo, and others—became global leaders in shipbuilding (and other sectors). Indeed, during the 1950s, 30 percent of the total loans made by the Japan Development Bank were for marine transportation (Stopford 2009). Japanese shipbuilders were also well-positioned to take advantage of structural changes underway in global shipping (such as increased demand for transportation of oil) that increasingly favored the use of larger ships: Japan's shipyards were larger and more integrated with steelmaking facilities than their European competitors, and the Japanese industry was at the forefront of innovation in ship construction that significantly reduced construction time (Stråth 1987). By 1970, Japan's global market share in shipbuilding had increased to 48 percent (from only 4.7 percent in 1949), while Europe's market share had fallen from 75 percent to 48 percent, in spite of the heavy European subsidies to the shipbuilding industry as a response to increased Japanese competition (Stråth 1987).

By the 1980s, however, Japan's shipbuilding was losing ground to South Korea. South Korean shipbuilding, similar to Japanese shipbuilding, grew as part of the government's large-scale push for heavy industrialization in the late 1970s (Choi and Levchenko 2021; Lane forthcoming). Recognizing shipbuilding as a strategic industry, Korea's government provided support primarily in the form of favorable financing (such as low-interest loans and government debt guarantees), as well as direct investment in shipbuilding facilities. Also like Japan, major South Korean conglomerates such as Hyundai, Samsung, and Daweoo grew quickly. Indeed, Korea's shipbuilding program was from the outset focused on producing very large ships in a small number of large shipyards owned by these conglomerates (Stopford 2009). Within 20 years, by 1995, South Korea's market share grew to 28 percent (from less than 1 percent in the early 1970s), reducing Japan's market share from 50 percent in 1975 to 41 percent and Europe's share from 32 percent to 23 percent.

However, unlike its predecessors in global shipbuilding, South Korea did not have a maritime tradition nor a large national fleet. Its maritime trade was much smaller than that of Europe or Japan. Even as South Korea emerged as one of the two leading shipbuilding countries, the share of the global shipping fleet registered in South Korea never exceeded 2 percent, whereas Japan's share had reached 10 percent by 1984 (Stopford 2009).[3] In other words, this was the first time that shipbuilding had been targeted as a primarily exporting sector. In 1995, 78 percent

[3]This estimate based on whether ships are registered in certain countries is likely a gross underestimate of the ships owned by companies from a given country, because of flags of convenience; for instance, according to Stopford (2009) in 2005, 90 percent of Japanese-owned ships sailed under foreign flags.

of Korea's ships were exported, compared to 42 percent for Japan (Lloyd's 1892–1999). This shift can be partially attributed to the growing adoption of "flags of convenience": from the late 1960s, shipowners began choosing flags of countries that provided tax and licensing benefits, such as Panama or Liberia, instead of remaining listed in their national fleet registry. This trend rendered shipowning a more "global" industry, breaking the link between shipowning and shipbuilding. Today, demand for ships remains globally fragmented and comes from many different countries and hundreds of different shipowning firms or fleet operators.[4]

Why did South Korea choose to subsidize shipbuilding? The development of heavy industries—including not just shipbuilding but also steel/metals, machinery, electronics, and petrochemicals—was seen by Korea's government as a prerequisite for long-term economic growth (Lane forthcoming). National security was another key motive: changing US foreign policy in the early 1970s, and in particular the withdrawal of one-third of all US troops from South Korea in 1971, led the Korean government to prioritize sectors perceived as being important for defense (Bruno and Tenold 2011). Finally, shipbuilding may have been targeted because the shipbuilding production process required skilled labor and sophisticated capital and machinery, or maybe because of South Korea's desire to follow the steps that Japan had taken.

By 2000, European shipbuilders were focusing on niche high-tech products, such as cruise ships. Their overall market share is 14 percent. Japan and South Korea compete head-to-head, with an overall market share of 38 percent each, but focused on different segments. Japanese yards dominate the production of bulk carriers (with a share of 70 percent), while South Korean yards lead the production of higher-end, specialized oil tankers (61 percent) and containerships (50 percent).

In the 2000s, China enters the shipbuilding scene. In 2002, former Premier Zhu inspected the China State Shipbuilding Corporation (CSSC), one of the two largest shipbuilding conglomerates in China, and pointed out that "China hopes to become the world's largest shipbuilding country (in terms of output) [. . .] by 2015." Within a few years, China overtook Japan and South Korea to become the world's leading ship producer in terms of output. By 2009, China's market share had reached 53 percent, from less than 10 percent in 2000; the combined market share of Japan and South Korea decreased from 75 percent in 2000 to 42 percent in 2009.

Industrial Policy and Shipbuilding: Some Key Questions

This narrative suggests that industrial policy has played a pivotal role in shaping the evolution of the modern shipbuilding industry. Since the start of the twentieth century, each new successor to the throne of the world's biggest shipbuilding region—continental Europe, Japan, South Korea, and China—has done so on the

[4]Industry structure for bulk carriers (tankers and dry bulk carriers) is highly fragmented with hundreds of firms operating globally (Kalouptsidi 2014). In container shipping, although operators are fairly concentrated, they often lease their vessels from a large number of small shipowners.

back of a deliberate program of supporting shipbuilding. Yet this dive into history also leaves at least three open questions.

First, why do governments subsidize shipbuilding? Our narrative suggests a wide variety of reasons: the connection between trade, shipping, and shipbuilding; the development of heavy manufacturing as a strategy for promoting economic growth; employment; national security and military considerations; and the desire for national prestige (or "pride and machismo," as Stråth (1987) puts it). Yet, in none of the historical cases is it self-evident exactly what mix of objectives led to industrial policy in shipbuilding.

Second, was industrial policy successful? It is challenging to evaluate if industrial policy worked. There are certainly examples of "apparent success" in Japan, South Korea, and China, where a country with a negligible initial share of the global industry embarks on a program of industrial policy and rapidly becomes a global leader. But the history of shipbuilding is also filled with examples of unsuccessful industrial policy, such as the long-standing US policy of protecting its shipbuilding sector through cabotage laws, European governments' prolonged and costly attempts to subsidize their shipbuilders in the face of Japanese and Korean competition (Stråth 1987), or an earlier attempt by South Korea to promote shipbuilding in the 1960s (Amsden 1989). Other countries have failed to launch a shipbuilding industry as well, as in the case of Brazil's failed attempt to launch its own shipbuilding sector in the late 1970s (Bruno and Tenold 2011). Even the apparent success stories required massive support, leading to the question (rarely answered in the literature) of whether the benefits from subsidizing shipbuilding are worth its large cost.

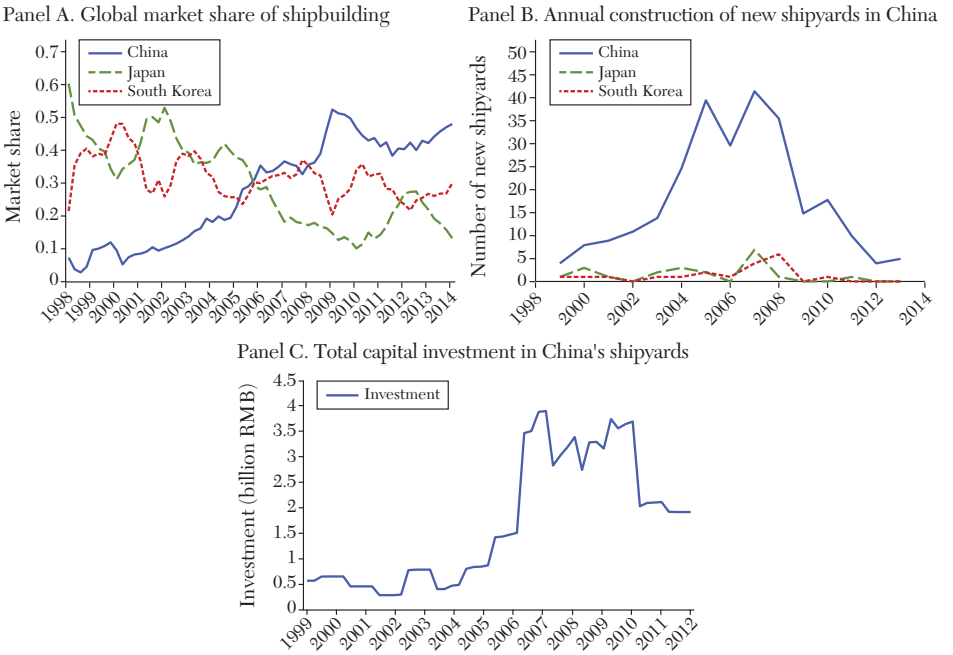
Third, how large is government support for shipbuilding? Even this seemingly straightforward question does not have an easy answer due to the "obscure jungle of subsidies" (Stråth 1987) that governments have used to subsidize shipbuilding. After all, governments have little incentive to be transparent about policy support, especially if there is a risk of triggering retaliatory policy by rivals. Many subsidies to shipbuilding favored by governments—such as preferential access to land or favorable financing—are implicit in nature, with their true cost difficult to uncover.

In our research, we address each of these questions through a detailed study of industrial policy in the context of the Chinese shipbuilding industry. We begin by describing China's industrial policy in shipbuilding.

Chinese Industrial Policy in Shipbuilding

As early as 2003, China's National Marine Economic Development Plan proposed constructing three shipbuilding bases centered in the Bohai Sea area (Liaoning, Shandong, and Hebei), the East Sea area (Shanghai, Jiangsu, and Zhejiang), and the South Sea area (Guangdong). However, China's 11th National 5-year Economic Plan 2006–2010 was the first to appoint shipbuilding as a "strategic industry" in need of "special oversight and support"; the central government "unveiled an official shipbuilding blueprint to guide the medium and long-term development of the industry." As part of the national plan, the central government

Figure 2
The Rapid Expansion of China's Shipbuilding Industry



Source: Barwick, Kalouptsidi, and Zahur (forthcoming), using data from Clarksons Research and China's National Bureau of Statistics.
Note: Market shares by country are computed from quarterly ship orders. Number of new shipyards is computed annually and by country. Industry aggregate quarterly investment by Chinese shipyards in billions of 2000 yuan.

set specific output and capacity goals: annual production was to reach 15 million deadweight tons by 2010 and 22 million deadweight tons by 2015. Remarkably, both goals were met several years in advance.

Panel A of Figure 2 shows the rise in China's global market share of shipbuilding by plotting China's total shipbuilding output as a share of global output. During this period, a boom-and-bust cycle took place in global shipbuilding. In the early 2000s, China's international imports (mostly commodities) and exports (mostly manufacturing) boomed, commodity prices soared, and as a result shipping rates spiked to a historical high. Shipowners placed heaps of new ship orders, and shipyard backlogs grew exponentially; by the end of 2008, the global ship backlog was more than five times larger than in 2001. But the shipbuilding boom was stopped short by the Great Recession of 2008–2009. The crisis led to an idling of the existing fleet, while at the same time another 70 percent of that fleet was still scheduled for delivery by 2012. Ship prices plummeted and threatened the survival of many shipyards.

China's national and local governments provided numerous subsidies for shipbuilding, which we classify into three groups. First, below-market-rate land prices along the coastal regions, in combination with simplified licensing procedures, acted as "entry subsidies" that incentivized the creation of new shipyards. As shown in panel B of Figure 2, between 2006 and 2008, the annual construction of new shipyards in China exceeded 30 new shipyards per year; in comparison, during the same time period, Japan and South Korea averaged only about one new shipyard per year each. In the booming mid-2000s, many of the orders for new ships were placed in these Chinese "greenfields," which were taking orders as they were getting built themselves.

Second, regional governments set up dedicated banks to provide shipyards with "investment subsidies" in the form of favorable financing, including low-interest long-term loans (a common industrial policy tool, as illustrated also by the programs in Japan and South Korea) and preferential tax policies. China's rise in total capital invested in shipyards is illustrated in panel C of Figure 2.

Third, China's government also employed "production subsidies" of various forms, such as subsidized material inputs, export credits, and buyer financing. The government-buttressed domestic steel industry provided cheap steel, which is an important input for shipbuilding. Export credits and buyer financing by government-directed banks made the new and unfamiliar Chinese shipyards more attractive to global buyers.

The combination of these policies was followed by a sharp expansion in China's shipbuilding production, market share, and capital accumulation. China's market share grew from 14 percent in 2003 to 53 percent by 2009, while Japan shrunk from 32 percent to 10 percent and South Korea from 42 percent to 32 percent. This impressive output growth was partially achieved via a massive entry wave of new firms: there were 173 new Chinese shipbuilding firms, a 230 percent increase in five years. Indeed, one intriguing characteristic of China's industrial policy in shipbuilding—which also applies to its industrial policy in other sectors such as solar panels, auto manufacturing, and steel—is that the industrial policy led to a large number of small firms. This pattern is in sharp contrast to the policies adopted by Japan and South Korea, which relied on promoting a handful of large conglomerates that became global industry leaders. Most of China's shipbuilding growth at this time was concentrated in the least high-tech ships (50 percent global share in bulk carriers) rather than in oil tankers or containerships (28 percent and 23 percent, respectively), where it also concentrated on smaller sizes.

Then came the Great Recession of 2008–2009, which drove the global shipping industry to a historic bust. The large number of new Chinese shipyards exacerbated low capacity utilization and contributed to plummeting ship prices around the world. The effectiveness of China's industrial policy was questioned. In response to the crisis and in an effort to promote industry consolidation, the government unveiled the "2009 Plan on Adjusting and Revitalizing the Shipbuilding Industry," which resulted in an immediate moratorium on entry and subsequently shifted support towards only selected firms in an issued "White List." Concentration in China's shipbuilding industry started increasing.

Since 2013, China has been the world's leading shipbuilder, accounting for 54 percent of all tonnage delivered in 2022, compared to 28 percent for South Korea and 18 percent for Japan. Moreover, China has begun to slowly move up the product ladder: between 2018 and 2022, Chinese shipbuilders delivered 45 percent of all new containerships (versus 36 percent for Korea), up from 23 percent between 2006 and 2010. Chinese shipyards still tend to build smaller containerships, although they have begun to build bigger ships in simpler ship types, such as bulk carriers.

How to Estimate Effects of Sectoral Industrial Policy

The case of shipbuilding illustrates how industrial policy can at least in some cases lead to rapid and substantial sectoral growth. But how can researchers evaluate the welfare impact of such growth, both domestically and internationally, as well as the efficacy of different types of government interventions? Such questions pose considerable challenges.

Any methodology for estimating the effects of industrial policy must grapple with a basic data challenge: government subsidies to industries are notoriously difficult to detect and measure. Indeed, "systematic data are non-existent; reliable sources of information are scarce and mostly incomplete [. . .] because governments do not systematically provide the information" (WTO 2006). Researchers and policymakers end up relying mostly on data reported by governments, like budgetary subsidies that are generally exempt from regulation, including research and development, environmental, and agricultural subsidies. However, reported measures of output subsidies tend to be crude and untrustworthy. Even worse, certain industrial support initiatives may be almost unmeasurable: consider a government-built airport in a small city, which builds a longer-than-needed runway (at considerable cost) that is used by an aircraft manufacturer for large plane trials.[5]

Beyond the data limitations, we need a methodology to assess the effect of industrial policy. The research in this area falls into three broad categories. An earlier approach traditionally relied on descriptive analyses that regress firm or sectoral outcomes on available measures of industrial policy (Noland and Pack 2003; Pack and Saggi 2006). But more recently, there has been a proliferation of studies that leverage natural or quasi-experimental variation in historical contexts to derive plausibly causal estimates of industrial policies on outcomes such as industry output, revenue, employment, exports, and sometimes productivity. As one example, Juhász (2018) uses a natural experiment, the Napoleonic blockade

[5] To address this challenge, researchers have in recent years compiled information on industrial policies by different countries across sectors using Global Trade Alert (Juhász et al. 2022; Evenett et al. 2024). These sources fill an important data gap because it is often hard to even know which sectors are targeted or which policies are in place, but these databases do not in general have complete information on the size of government subsidies.

of Britain from 1806–1813, to estimate the causal effect of temporary trade protection on long-term economic development. Though the blockade was not an explicit industrial policy, it created exogenous and differential variation in trade costs, effectively protecting northern French cotton spinners more than southern ones against British imports. In another notable example, Lane (forthcoming) studied South Korea's Heavy and Chemical Industry Drive in 1973–1979, in which industrial policy reshaped Korea's long-term dynamic comparative advantage.[6]

Other studies suggest more nuanced findings. For example, in another study of South Korea's Heavy and Chemical Industry Drive, Kim, Lee, and Shin (2021) report that misallocation worsened significantly during the period.[7] Using exogenous variation in the eligibility criteria of place-based investment subsidies in the United Kingdom, Criscuolo et al. (2019) find that these subsidies increased manufacturing employment for small firms but not large firms, with positive effects on investment but not productivity.[8] Similarly, exploiting changes in the eligibility criteria for India's small-firm subsidy programs, Rotemberg (2019) finds almost complete output crowd-out among domestically consumed products but much less crowd-out for those that were exported. Aggregate productivity rose by 1–2 percent due to the higher productivity of newly eligible firms. This finding echoes an obvious but often forgotten message: firms are heterogeneous, so policies that promote the participation of more efficient firms tend to be more effective (Barwick, Kalouptsidi, and Zahur forthcoming).

Such studies have greatly enriched our understanding of the effects of industrial policy and mark a substantial improvement over the earlier, largely correlational studies (Juhász, Lane, and Rodrik 2024). On the other hand, these studies are bound by their research designs. None of them can readily conduct a cost-benefit analysis, leaving open the question of whether these industrial policies passed the "Bastable test"—that is, whether the discounted future gains in consumer and producer surplus exceed the costs of protection (Harrison and Rodríguez-Clare 2010). These studies are limited in their ability to make counterfactual predictions, such as how industries would have evolved in the absence of industrial policy, and sometimes uninformative in terms of mechanisms by which industrial policy affects the economy. Finally, these papers are silent on the comparison across alternative policy instruments and how to design effective and welfare-enhancing policies.

A second approach builds on the significant body of work on "strategic trade," whereby government interventions alter the strategic interactions of firms that compete globally. The literature focused on when such policies benefit the domestic economy. Early prominent work in this area, including Helpman and Krugman

[6] See also Aghion et al. (2015) and Harris, Keay, and Lewis (2015). Lane (2020) and Juhász, Lane, and Rodrik (2024) provide reviews of recent studies. There is also a growing literature that examines the effect of industrial policy on quality upgrading (Bai et al. 2024). See Verhoogen (2023) for a review.

[7] An earlier literature (for example, Krueger and Tuncer 1982; Birdsall et al. 1993) and plenty of anecdotal evidence indicate that industrial policies often lead to misallocation and excess capacity, a topic we return to below.

[8] See Neumark and Simpson (2015) for a review of place-based policies.

(1989) and Brander (1995), was based on models that were simple enough to remain analytically tractable. However, it turned out that different specifications of the model could entirely alter the conclusions; for instance, the theoretical predictions about the welfare implications of industrial policies depend on whether firms compete in prices or quantities. If firms are choosing quantities produced (Cournot competition), subsidies can improve national welfare. But subsidies would reduce the welfare of national players if firms competed in prices (Bertrand competition), because that would intensify competition. This ambiguity in terms of when it is optimal to use subsidies or taxes to help the national champion in a globally oligopolistic industry obstructed the evolution of this literature.

That said, in follow-up literature, empirical studies applied these models to study the effect of industrial policies in specific industries, often using calibration or simulation to assess welfare effects (for an excellent review, see Harrison and Rodríguez-Clare 2010).[9] This literature acts as a precursor to the third approach we discuss below. The lessons of this approach are more tempered. While industrial policies led to growth in the targeted industry, several of the policies evaluated (Baldwin and Krugman 1988a; Luzio and Greenstein 1995; Irwin 2000) did not pass the Bastable test. Given the multiple objectives of industrial policy—including distributional considerations, national security, and international competitiveness—the Bastable test may be too narrowly focused. Nevertheless, welfare calculations serve as a valuable benchmark. Finally, these studies often relied on aggregate data, stylized models, and calibration techniques. Few examined investment and capacity decisions directly, which are arguably the most critical margins that industrial policy seeks to influence.

A third approach to the study of industrial policy applies structural methodology in the field of industrial organization (Ackerberg et al. 2007; Ho, Hortaçsu, and Lizzeri 2021). This approach essentially seeks to combine the two previous approaches, harnessing the benefits of each. Unlike reduced-form analysis, model-based empirical work looks at the data through the lens of an empirical industry equilibrium model. Unlike the strategic trade models, it relies on devising the simplest model that is realistic enough to capture the main features of the environment under study. Most importantly, the combination of data and theory allows the researcher to test or validate the modeling assumptions imposed (Nevo and Whinston 2010). Finally, structural methodology allows the researcher to infer a measure of industrial subsidies that are otherwise unobserved (more on that below), as well as to evaluate welfare and other (policy) counterfactuals of interest.

[9] Baldwin and Krugman (1988a) and Baldwin and Krugman (1988b) analyze the US-Japan rivalry in the 16K RAM sector during 1978–1983 and Airbus' rise of world market share in the 1970s, as a result of trade protection and subsidies, respectively. Head (1994) studies the effect of tariff protection and learning-by-doing on the emergence of the steel rail industry in the United States between the Civil War and World War I. Irwin (2000) analyzes the effect of tariff protection on the growth of the US tinplate industry in the 1890s, and Luzio and Greenstein (1995) study the effects of Brazil's protection of the microcomputer industry in the 1980s.

In this approach, the researcher builds a "custom" model for the industry under study—say, shipbuilding. This modeling requires a deep understanding of the environment; for example, the market structure of producers and buyers, how firms compete, how prices are formed, the production cost function, and other important firm decisions, such as whether the industry sells a homogeneous or differentiated good, and finally how government subsidies affect firms. Possible answers to these questions come from extensive discussions with industry participants, industry press or reviews, as well as the prior academic literature. The model must account for the key features of the industry under study and allow for the key mechanisms the researcher is interested in analyzing. The model must then "meet" the data. The researcher collects data, usually in the form of firm actions such as quantities produced, prices, investments, and product characteristics. Then, one estimates key relationships like the demand curve and the firm's cost curve. The goal is to "assign numbers" to key parameters of interest, such as the demand elasticity and the marginal cost of production or investment. Finally, this quantitative model is used to compute counterfactuals of interest; for example, to predict the evolution of an industry with subsidies (as observed) or without, or to change the policy mix and compare the new equilibrium in the model to the observed outcomes.

A small but growing literature applies this third approach to evaluate industrial policies in various contexts; for example, Spain's decade-long effort to promote its domestic automobile industry (Miravete and Moral 2024); the optimal design of China's electric vehicle subsidies (Barwick, Kwon, and Li 2024); the positive global spillovers of the electric vehicle subsidies implemented in China, Europe, and the United States, as well as the negative implications of local content requirements, in the presence of steep learning-by-doing in the upstream electric vehicle battery sector (Barwick et al. 2024); and China's research and development subsidies in its InnoCom Program (Chen et al. 2021).

The structural approach is not without its concerns. Does the model capture the key characteristics of the industry? Does it allow for potential confounding factors, and is there rich enough data variation to consider such factors? Put differently, the findings are conditional on modeling assumptions, and the data must allow one to test these assumptions. Finally, this approach focuses on partial equilibrium or sector- specific analysis, which allows the researcher to exploit rich data and institutional details to answer important "what-ifs." However, general equilibrium analysis may be important in exploring spillovers of industrial policy (Liu 2019; Choi and Levchenko 2021).

Illustration: China's Twenty-First Century Shipbuilding Program

We illustrate the structural approach to studying the impact of industrial policy on industry evolution and global welfare in the context of China's twenty-first century shipbuilding program in Kalouptsidi (2018) and Barwick, Kalouptsidi, and

Zahur (forthcoming). To our knowledge, this work is the first attempt at evaluating quantitatively industrial policy in shipbuilding globally and among the first papers employing the structural industrial organization methodology to understand the welfare implications and effective design of industrial policy more generally.

Framework

We build a model of a global market for ships. On the demand side, a large number of shipowners across the world are deciding whether to buy new vessels. Their willingness-to-pay for new ships depends on present and expected future market conditions, notably world trade and the current fleet level. On the supply side, our model considers shipyards located in China, Japan, and South Korea (which account for 90 percent of world production). Each shipyard decides how many ships to produce by comparing the market price of a ship, dictated by the willingness-to-pay of shipowners on the demand side, and its production costs. Shipyards are price-takers (an assumption we relax below) and will keep producing as long as the price exceeds the marginal cost of an additional vessel. Thus, we can use the assumption of profit-maximizing behavior, along with observed ship prices and firm-level production, to uncover the shipyard's cost function. We assume the cost function is convex—that is, marginal costs of production rise with quantity—a feature that may capture capacity constraints.

To bring this model to data, we employ a rich dataset consisting of firm-level quarterly ship production between 1998 and 2014, firm-level investment, entry and exit, and new ship market prices by ship type (containerships, tankers, and dry bulk carriers, which together account for 90 percent of global sales). In other words, for each shipyard in the world we know how many ships are ordered and delivered every quarter, how much new capital is invested, whether the shipyard exits or is a new entrant, as well as the prevailing global ship price. Such data on the shipbuilding industry are available from several data providers, at a relatively small cost to the researcher, such as Clarksons Research (which we use), Lloyds, and S&P Global.

As a starting point, we need to measure the size of China's subsidies for shipbuilding. But how can we do this when the subsidies are both direct and indirect and poorly measured? We know that China's subsidies for shipbuilding started in 2006. We use data from all countries before 2006 to estimate a shipbuilding cost function. This cost function should then continue to predict shipbuilding outside of China after 2006; however, it can only match China's shipbuilding after 2006 by including a measure of what subsidies must have been. With the model in hand, we can compare China's observed shipbuilding production with subsidies to the outcome that would have arisen had China not subsidized shipbuilding.

In other words, our methodology relies on combining available data on firm choices and an economic model. It aims at uncovering a "gap" between the observed production and the optimal production the economic model implies (that is, the production that equates price to marginal production cost). We estimate the marginal cost from variation in prices and examine its behavior around

2006.[10] We are particularly interested in whether this cost function exhibits an abrupt and otherwise inexplicable change around 2006 that makes Chinese shipyards' production costs all of a sudden lower. The idea is to essentially ask whether Chinese firms are "over"-producing, compared to our prediction from the earlier production function.

How Big Were China's Production Subsidies?

Our estimates suggest that China provided $23 billion in production subsidies between 2006 and 2013. This finding is driven by the cost function obtained from this analysis, which exhibits a significant drop for Chinese producers equal to about 13–20 percent of the cost per ship. Simply put, Chinese shipbuilding firms were "over"-producing after 2006 compared to our prediction of output without subsidies.

Might some confounding factor, like a technological advance or a particular type of ship being built, explain this pattern? A general change in shipbuilding technology does not seem to be the cause because the cost reduction is only present for Chinese shipyards—there are no "breaks" in the estimated cost functions of Japanese or South Korean shipyards. In addition, the results are robust to many different specifications, as well as different ways of accounting for temporal changes. For example, the results hold when only shipyards that existed prior to 2001 are considered—which in turn suggests that cost declines are not driven by new shipyards, which may have a different technology or may be learning-by-doing. Moreover, the results hold if we focus on the smallest size category of bulk vessels (called Handysize), where China was already an important producer before 2006. China's production process for shipbuilding does not seem to be characterized by significant technological advances, and product differentiation is very limited.

Finally, one may wonder about certain modeling assumptions, such as the price-taking assumption for shipyards. This industry is globally fragmented, with the Herfindahl-Hirschman Index ranging from 230 to 720 between 2006 and 2013, and China alone having more than 250 shipyards during the peak of the boom. Nonetheless, we carry out a version of our analysis where we relax the price-taking assumption, as industrial policy is often motivated by strategic trade considerations and these considerations only come into play if firms can exert market power. Here we face a choice between assuming Bertrand competition (firms compete in prices) and Cournot competition (firms compete in quantities). In the shipbuilding industry, capacity constraints are first-order: a typical shipyard can only work on a handful of ships at a time, therefore it makes more sense to assume ships compete by choosing quantity. When we estimate a model of Cournot competition and

[10] The cost function relates output to operating expenditures. As in many industries, however, costs of production are not readily observed. Standard methodology in industrial organization (Bresnahan 1982; Berry, Levinsohn, and Pakes 1995) estimates costs from rich variation in demand. In our context, price is equal to marginal cost of production. If we impose a functional form on the marginal cost (say quadratic in quantity), we can use the observed price, and the level of output to estimate the cost function coefficients.

calculate the implied markups, we find these are tiny: the average markup for bulk carriers is only 6.39 percent of the price of a new bulk carrier, and markups are even smaller for tankers (4.26 percent) and container ships (2.12 percent). Similarly, the assumption that ships are homogeneous (conditional on category and size) is motivated by the empirical pattern that prices of new ships are almost perfectly predictable from ship type, ship size, and time fixed effects.

Taking Dynamics into Account

The basic framework as we have sketched here misses several important features, which can affect the long-run behavior of firms. China's industrial policy in shipbuilding included subsidies not just for production but for entry and investment as well; Figure 2 showed a dramatic rise in investment and the number of new shipbuilding firms in China. Entry and investment have long-run implications for industry structure, especially with the severe volatility that characterizes this industry, because entry, exit, and investment are sluggish to adjust. For instance, during busts, firms do not necessarily exit or divest: these decisions are largely irreversible and firms may delay them in the hope that demand will recover. Even the decision of how many ships to produce is subject to dynamic considerations: building a ship takes two to five years and thus shipyards accumulate backlogs, which can affect their future production costs, either negatively (capacity constraints) or positively (expertise acquisition or larger input orders).

The model used in our empirical analysis of industrial policy is flexible enough to capture these dynamic features of the market for ships. Both the demand for shipping and the supply of ships are at the mercy of large macroeconomic swings, and firms operate in the shadows of severe uncertainty regarding both demand for international trade as well as input cost shocks (for example, steel prices). Demand for new ships is driven by demand for international sea transport, which is uncertain and volatile. On the supply side, we employ a dynamic model of industry evolution, where firms can enter, exit, invest to increase their capital, and compete by producing ships. Shipyards decide whether to enter by comparing their lifetime expected profitability to entry costs, which include the costs to set up a new firm (such as the cost of land acquisition, shipyard construction, and any initial capital investments) and the implicit cost of obtaining regulatory permits. They exit if expected profitability from remaining in the industry falls below a given threshold, capturing the shipyard's "scrap" value (that is, the proceeds from liquidating the business, as well as any option values of the firm). Optimal production decisions involve comparing current margins to expected costs, given input price fluctuations and backlog accumulation. The industry is globally fragmented enough that we assume firms do not engage in strategic dynamic interactions; that said, they form expectations about the evolution of the industry, and in equilibrium these are correct on average (Hopenhayn 1992).

Entry and investment subsidies are identified following the same strategy as for production subsidies. As before, we use firms' observed entry and investment behavior to back out entry and investment costs. This is done by finding the cost

parameters that bring observed behavior as close as possible to the optimal behavior implied by the dynamic model. Once entry and investment costs are estimated, subsidies are estimated by comparing Chinese to non-Chinese firms, before and after 2006. The three types of subsidies can be separately identified because they affect different decisions that firms make, on which we have rich firm-level data.

Our estimates suggest that China provided $91 billion in subsidies along all three margins—production, entry, and investment—between 2006 and 2013, averaging over $11 billion per year, which totaled nearly 50 percent of Chinese shipbuilding industry revenue over that period. This is considerably larger than the $23 billion that (according to our estimates) was provided in production subsidies alone. Thus, all firm decisions—entry, exit, and capacity investment in addition to production—matter in evaluating the impact and effectiveness of industrial policy. Indeed, entry subsidies were 69 percent of total subsidies, while production subsidies were 25 percent, and investment subsidies accounted for the remaining 6 percent. This empirical pattern reflects that shipbuilding firms "over-entered (recall the astonishing entry rates during the boom years of 2006–2008) and "over-invested" (recall the striking increase in investment *during the bust*) as shown earlier in Figure 2.

Output and Welfare

Perhaps the biggest advantage of building a quantitative model for an industry is that it can evaluate hypothetical scenarios. For example, what would have happened to China's shipbuilding industry absent Chinese subsidies? Presumably, China's market share would still have increased, especially during the boom years, but by how much?

Our structural model suggests that China's industrial policy in support of shipbuilding boosted China's domestic investment in shipbuilding by 140 percent, and more than doubled the entry rate: 143 shipbuilding firms entered with subsidies versus 64 without subsidies from 2006 to 2013. It also depressed exit. Overall, industrial policy raised China's world market share in shipbuilding by more than 40 percent.

Calculating whether this increase in sectoral output should be counted as an increase in welfare is a more delicate question. Here are several slices at an answer. First, 70 percent of China's output expansion occurred via taking business from rival countries. From a global perspective, Chinese subsidies reduced South Korea's world market share from 48 percent to 39 percent and Japan's market share from 23 percent to 20 percent during 2006–2013, with profits earned by shipyards in these two countries falling by ¥144 billion (in US dollars, roughly $21 billion). There is evidence (backed by our cost estimates) that Chinese shipyards are less efficient than their Japanese and South Korean counterparts; thus, the transfer of shipbuilding to China that occurred constitutes a misallocation of global resources.

Second, China's industrial policy for shipbuilding led to considerable declines in ship prices. Lower ship prices benefited world ship-buyers somewhat, though only a modest amount accrues to Chinese ship-buyers, as they accounted for a small fraction of the world fleet.

Table 1
Welfare Effects of Industrial Policy

	Δ Net Profit/Subsidy	*Δ Revenue/Subsidy*
Panel A. Comparison of different policy instruments		
Investment Subsidy	74%	153%
Production Subsidy	50%	153%
Entry Subsidy	32%	66%
All Subsidies	18%	72%
Panel B. Industrial policy and the business cycle		
Procyclical subsidies	38%	189%
Counter-cyclical subsidies	70%	168%
Panel C. Targeted industrial policy		
Subsidize all firms	37%	85%
Subsidize "White List" firms	71%	105%

Source: Barwick, Kalouptsidi, and Zahur (forthcoming).
Note: Panel A compares the actual policy mix where firms received all three kinds of subsidies ("All Subsidies"), with counterfactual policy mixes where the government provides only one type of subsidy. In the counterfactuals reported in panel B and panel C, firms receive a combination of production and investment subsidies. Panel B compares the effect of providing subsidies during the boom (2006–2008) versus the bust (2009 onwards). Panel C compares a policy where all firms are eligible to receive subsidies versus a "White List" policy where only selected firms (chosen on the basis of how profitable they are) are eligible for subsidies. "ΔNet Profit/Subsidy" is our measure of the gross return on the policy, and equals the change in the discounted sum of net profits (relative to the scenario with no subsidies), divided by the discounted sum of subsidies. Likewise, "ΔRevenue/Subsidy" refers to the change in the discounted sum of revenue (relative to the scenario with no subsidies), divided by the discounted sum of subsidies.

Finally and most importantly, although China's shipbuilding subsidies were highly effective at achieving output growth and market share expansion, we find that they were largely unsuccessful in terms of welfare measures. The program generated modest gains in domestic producers' profit and domestic consumer surplus. In the long run, the gross return rate of the adopted policy mix, as measured by the increase in lifetime profits of domestic firms divided by total subsidies, is only 18 percent, meaning that for every $1 the government spends, it gets back 18 cents in profitability. In other words, the net return when incorporating the cost to the government was a negative 82 percent, with entry subsidies explaining a lion's share of the negative return.

Alternative Design of Industrial Policy

Our structural model also allows us to consider alternative industrial policy designs. Policy design is bound to be crucial, especially when some producers are more efficient than others, and when demand in the industry is highly volatile.

We first begin by comparing the efficacy of production, investment and entry subsidies (panel A of Table 1). Although none of the policies (in isolation nor

together) yield positive returns in terms of lifetime profits,[11] production and investment subsidies can be justified if the goal is industry revenue maximization (the ratio of increased industry revenue to subsidies is 153 percent, meaning that $1 in subsidies generates a $1.53 increase in lifetime revenue). This finding might explain the popularity of these subsidies in China, because the promotions of local individuals are often linked to quantity and revenue targets. Indeed, several shipbuilding programs were local, motivated in part by the important implications of increased shipbuilding output on local industrial growth. This led to regional duplication, with several provinces having their own local shipbuilding industry.

Entry subsidies are wasteful—even by the revenue metric—and lead to increased industry fragmentation and idleness. Entry subsidies attract small and inefficient firms. In contrast, production and investment subsidies increase the backlog and capital stock, which lead to economies of scale and drive down both current and future production costs. As such, they favor large and efficient firms. Indeed, the take-up rate for production and investment subsidies is much higher among efficient firms: 82 percent of production subsidies and 68 percent of investment subsidies is allocated to firms that are more efficient than the median firm, whereas only 49 percent of entry subsidies goes to more efficient firms.

Another important consideration is the volatile nature of the shipbuilding industry, which is subject to boom and bust cycles. Our model suggests that a counter-cyclical policy would outperform the procyclical policy that was adopted by a large margin: strikingly, subsidizing firms in production and investment during the boom leads to a gross rate of return of only 38 percent (a net return of −62 percent), whereas subsidizing firms during the downturn leads to a much higher gross return of 70 percent (a net return of −30 percent), as shown in panel B of Table 1. In boom periods, the industry is operating close to full capacity, so further expansion is costly and entails the utilization of high-cost resources. During a bust, the industry operates well below capacity and subsidies mobilize underutilized facilities, resulting in smaller distortions. In addition, subsidies during a boom attract inefficient firms, which pushes down the rate of return. Despite the benefits of a countercyclical policy, the actual policy mix was overwhelmingly procyclical: 90 percent of total subsidies was handed out between 2006 and 2008 versus 10 percent between 2009 and 2013.

Finally, we examine the consolidation policy adopted in the aftermath of the financial crisis, whereby the Chinese government implemented a moratorium on entry into shipbuilding and issued a "White List" of firms prioritized for government support (panel C of Table 1). Indeed, China adopted this strategy in several industries to curb excess capacity and create national champions that can compete globally, following the examples of Japan and South Korea. In our calculations, if an "optimal White List" is formed—that is, the most productive firms are chosen for subsidies—the gross rate of return climbs to 71 percent versus 37 percent when all firms are subsidized. Why? Subsidizing all firms encourages suboptimal entry; in

[11] Benefits to domestic consumers are negligible and not included for simplicity.

contrast, the White List subsidizes existing firms and does not distort entry. Moreover, firms on the White List have lower production costs; shifting support to more efficient firms reduces misallocation. However, China's actual White List was suboptimal, as it favored state-owned enterprises. This illustrates a further difficulty with designing industrial policy: regulatory capture.

Our results highlight why industrial policies have worked better for some countries. In East Asian countries where industrial policy was often considered successful, the policy support was often conditioned on firm performance. In contrast, in Latin America where industrial policies often aimed at import-substitution, no mechanisms existed to weed out nonperforming beneficiaries (Rodrik 2009). In China's modern-day industrial policy in the shipbuilding industry, the policy's return was low in earlier years when output expansion was primarily fueled by the entry of inefficient firms, but increased over time as the government relied on "performance-based" criteria via its White List. Such targeted industrial policy design can be substantially more successful than open-ended policies that benefit all firms.

China's Industrial Policy in Shipbuilding: Why?

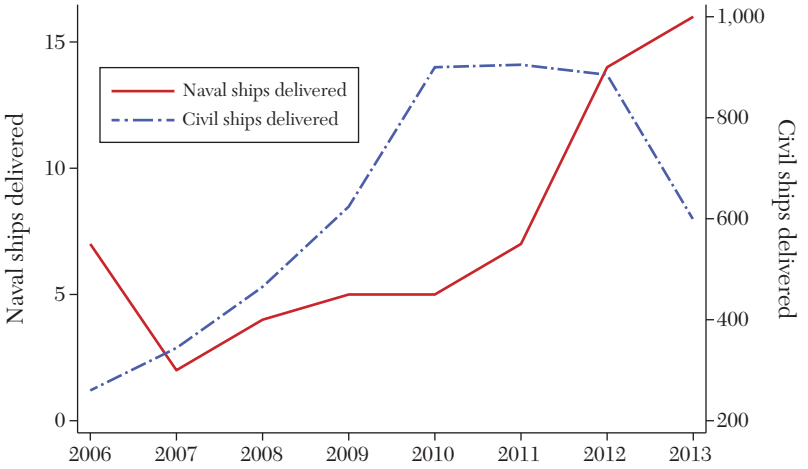
If industrial policy related to shipbuilding had a low payoff, then what objectives, economic or noneconomic, was China's government trying to attain?

Many of the standard arguments for industrial policy do not seem to apply especially well to shipbuilding in our sample period. The shipbuilding industry is fragmented globally, market power is limited, and markups are slim; thus, there are no "rents on the table" that, when shifted from foreign to domestic firms, outweigh the cost of subsidies. The possibility of industry-wide learning-by-doing ("Marshallian externalities") is another common rationale for industrial policy, but we find little evidence of learning-by-doing, perhaps because the production technology for the ship types that China expanded the most, such as bulk ships, was already mature. Another industrial policy argument focuses on spillovers to other domestic sectors (like steel production or the labor market), but we find limited evidence for such spillovers in this context. In addition, more than 80 percent of ships produced in China are exported, which limits the fraction of subsidy benefits that is captured domestically. A scenario whereby Chinese output growth in shipbuilding eventually forces competitors to exit does not seem first-order either: by 2023, no substantial foreign exit has been observed.

Of course, one possibility already noted is that the structure of China's incentives for local political leaders rewards readily observed results, so the observed growth in shipbuilding output and global market share is sufficient to offer a political justification for the subsidies. However, we also find support for two different rationales.

First, as China became the world's biggest exporter and a close second largest importer during our sample period, transport cost reductions from increased shipbuilding and reduced shipping costs can lead to substantial increases in its trade volume. China's imports consist mainly of raw materials and are carried

Figure 3
Military Ship Production



Source: Barwick, Kalouptsidi, and Zahur (forthcoming), using data from Clarksons and IHS Jane's.
Note: This figure plots the number of commercial ships and naval ships delivered by Chinese shipyards over time.

by bulk carriers and tankers, while its exports are mostly manufactured goods and are transported in containerships. To evaluate this argument, we carry out a back-of-the-envelope calculation of the subsidies' impact on China's trade flows. Shipbuilding subsidies reduced bulk carrier freight rates by 6 percent and containership freight rates by 2 percent between 2006 and 2013. Using trade elasticities from the literature (Brancaccio, Kalouptsidi, and Papageorgiou 2020; Jeon 2022), the industrial policy raised China's annual trade volume by 5 percent ($144 billion) between 2006 and 2013. This increase in trade was certainly large relative to the size of the subsidies (which averaged $11.3 billion annually between 2006 and 2013). Of course, "more trade" does not translate directly into economic well-being, but the relative magnitudes are suggestive.

Second, China's military ship production might have benefited from industrial policy with regard to shipbuilding. Military ship production is concentrated at state-owned yards, especially at 13 subsidiaries of China State Shipbuilding Corporation (CSSC) and China Shipbuilding Industry Corporation (CSIC), the two largest conglomerate shipyards that are state-owned.[12] These subsidiaries are typically dual-use, producing both commercial and military ships in the same complex. Figure 3 plots the annual deliveries of naval and commercial ships from 2006 to

[12] Our primary sources are the yearly report known as IHS Jane's Fighting Ships, produced by the intelligence company IHS Jane's (Saunders 2015), as well as *Chinese Naval Shipbuilding: an Ambitious and Uncertain Course*, a 2017 book about the Chinese naval shipbuilding industry (Erickson 2016). We are grateful to Elliott Mokski for discovering and collecting these datasets.

2013. Both types of deliveries experienced a several-fold increase during this period, although military production appears to have accelerated after the financial crisis and continued to increase throughout the sample period, providing suggestive evidence that China's supportive policy might have benefited its military production as well.

Ongoing Challenges of Research on Industrial Policy

In many ways, the example of China's industrial policy with regard to shipbuilding echoes patterns observed in other countries and industries, and thus illustrates the academic and policy-making challenges of industrial policy.

First, studying industrial policy in shipbuilding showcases the issue of scant and mismeasured data on government interventions; thus, we had to derive estimates of subsidies from a structural model and firm-level data. More generally, government subsidies to industries are notoriously difficult to detect and measure. Indeed, partly because international trade agreements prohibit direct and in-kind subsidies, "systematic data are non-existent" (WTO 2006), and thus the presence and magnitude of industrial subsidies are often unknown. This lack of both information and compliance has obstructed the role of global policymakers and prompted them to reevaluate their guidelines.

Second, it showcases the problem of designing industrial policy; indeed, given the pressures that politicians face to support certain industries, the relevant question they face with regard to industrial policy may not be whether to do it, but how (Rodrik 2009). China's industrial policy for shipbuilding started with large subsidies, such as cheap land and subsidized credit, without many restrictions on the firms that could access them. This policy approach proved costly and inefficient, as it led to a massive entry wave of new firms that were not high performers. The negative consequences of the poor design were hidden during the boom years but painfully revealed when excess capacity came to plague the industry during the Great Recession of 2008–2009. The government then subsidized only firms on a White List; this pattern seems to continue today in the "Made in China 2025" program.

Third, economists need an improved methodology for assessing the welfare impact of industrial policy, domestically and globally. Measuring welfare effects necessitates a modeling framework, with nontrivial choices to be made as to what is included (for example, partial versus general equilibrium). Most important, in our view, is that evaluations of industrial policy need to reach beyond purely economic objectives; for instance, how do researchers incorporate the geopolitical considerations that are so common in industrial policy agendas today? Understanding noneconomic objectives requires economists to think outside their standard toolbox and thus poses both a great challenge and an opportunity for research.

■ *We are thankful to Elhanan Helpman for helping parse through the literature on industrial policy. We are also grateful to Leran Qi for excellent research assistance.*

References

Ackerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes. 2007. "Econometric Tools for Analyzing Market Outcomes." In *Handbook of Econometrics*, Vol. 6A, edited by James J. Heckman and Edward E. Leamer, 4171–4276. Amsterdam: North-Holland.

Aghion, Philippe, Jing Cai, Mathias Dewatripont, Du Luosha, Ann Harrison, and Patrick Legros. 2015. "Industrial Policy and Competition." *American Economic Journal: Macroeconomics* 7 (4): 1–32.

Amsden, Alice H. 1989. *Asia's Next Giant: South Korea and Late Industrialization*. New York and Oxford: Oxford University Press

Bai, Jie, Panle Jia Barwick, Shengmao Cao, and Shanjun Li. 2024. "Quid Pro Quo, Knowledge Spillover, and Industrial Quality Upgrading: Evidence from the Chinese Auto Industry." CEPR Discussion Paper 19440.

Baldwin, Richard, and Paul Krugman. 1988a. "Market Access and International Competition: A Simulation Study of 16K Random Access Memories." In *Empirical Methods for International Trade*, edited by Robert C. Feenstra. Cambridge, MA.: MIT Press.

Baldwin, Richard, and Paul Krugman. 1988b. "Industrial Policy and International Competition in Wide-Bodied Jet Aircraft." In *Trade Policy Issues and Empirical Analysis*, edited by Robert E. Baldwin, 45–78. Chicago: University of Chicago Press.

Barwick, Panle Jia, Myrto Kalouptsidi, and Nahim Bin Zahur. Forthcoming. "Industrial Policy Implementation: Empirical Evidence from China's Shipbuilding Industry." *Review of Economic Studies*.

Barwick, Panle Jia, Hyuk-Soo Kwon, and Shanjun Li. 2024. "Attribute-Based Subsidies and Market Power: An Application to Electric Vehicles." NBER Working Paper 32264.

Barwick, Panle Jia, Hyuk-soo Kwon, Shanjun Li, and Nahim Bin Zahur. 2024. "Drive Down the Cost: Learning by Doing and Government Policy in the Global Electric Vehicle Battery Industry." Unpublished.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.

Birdsall, Nancy M., Jose Edgardo L. Campos, Chang-Shik Kim, W. Max Corden, Lawrence MacDonald, Howard Pack, John Page, Richard Sabor, and Joseph E. Stiglitz. 1993. *The East Asian Miracle: Economic Growth and Public Policy*. Oxford: Oxford University Press.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2020. "Geography, Transportation, and Endogenous Trade Costs." *Econometrica* 88 (2): 657–91.

Brander, James A. 1995. "Strategic Trade Policy." In *Handbook of International Economics*, Vol. 3, edited by Gene M. Grossman and Kenneth Rogoff, 1395–1455. Amsterdam: North-Holland.

Bresnahan, Timothy F. 1982. "The Oligopoly Solution Concept Is Identified." *Economics Letters* 10 (1–2): 87–92.

Bruno, Lars, and Stig Tenold. 2011. "The Basis for South Korea's Ascent in the Shipbuilding Industry, 1970–1990." *The Mariner's Mirror* 97 (3): 201–17.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. 2021. "Notching R&D Investment with Corporate Income Tax Cuts in China." *American Economic Review* 111(7): 2065–2100.

Choi, Jaedo, and Andrei A. Levchenko. 2021. "The Long-Term Effects of Industrial Policy." NBER Working Paper 29263.

Criscuolo, Chiara, Ralf Martin, Henry G. Overman, and John Van Reenen. 2019. "Some Causal Effects of an Industrial Policy." *American Economic Review* 109 (1): 48–85.

Erickson, Andrew S. 2016. *Chinese Naval Shipbuilding: An Ambitious and Uncertain Course*. Annapolis, MD:

Naval Institute Press.
Evenett, Simon, Adam Jakubik, Fernando Martín, and Michele Ruta. 2024. "The Return of Industrial Policy in Data." *World Economy* 47 (7): 2762–88.
Flath, David. 2022. *The Japanese Economy*. Oxford: Oxford University Press.
Folkman, Varg. 2024. "Shipbuilders Ask for EU Help on Chinese Subsidies." *Politico*, January 23. https://www.politico.eu/article/ship-builders-eu-help-china-foreign-subsidies/.
Foroohar, Rana. 2024. "Shipbuilding: The New Battleground in the US-China Trade War." *Financial Times*, March 12. https://www.ft.com/content/4e2d5bb7-e4d5-4b98-b1a8-895c0d493b07.
Hanlon, W. Walker. 2020. "The Persistent Effect of Temporary Input Cost Advantages in Shipbuilding, 1850 to 1911." *Journal of the European Economic Association* 18 (6): 3173–3209.
Harris, Richard, Ian Keay, and Frank Lewis. 2015. "Protecting Infant Industries: Canadian Manufacturing and the National Policy, 1870–1913." *Explorations in Economic History* 56: 15–31.
Harrison, Ann, and Andrés Rodríguez-Clare. 2010. "Trade, Foreign Investment, and Industrial Policy for Developing Countries." In *Handbook of Development Economics*, Vol. 5, edited by Dani Rodrik and Mark Rosenzweig, 4039–4214. Amsterdam: North-Holland.
Head, Keith. 1994. "Infant Industry Protection in the Steel Rail Industry." *Journal of International Economics* 37 (3–4): 141–65.
Helpman, Elhanan, and Paul Krugman. 1989. *Trade Policy and Market Structure*. Cambridge, MA: MIT Press.
Ho, Kate, Ali Hortaçsu, and Alessandro Lizzeri, eds. 2021. *Handbook of Industrial Organization*. Vol. 5. Amsterdam: North-Holland.
Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–50.
Irwin, Douglas A. 2000. "Did Late-Nineteenth-Century US Tariffs Promote Infant Industries? Evidence from the Tinplate Industry." *Journal of Economic History* 60 (2): 335–60.
Jeon, Jihye. 2022. "Learning and Investment under Demand Uncertainty in Container Shipping." *RAND Journal of Economics* 53 (1): 226–59.
Juhász, Réka. 2018. "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade." *American Economic Review* 108 (11): 3339–76.
Juhász, Réka, Nathan Lane, Emily Oehlsen, and Verónica C. Pérez. 2022. "The Who, What, When, and How of Industrial Policy: A Text-Based Approach." Unpublished.
Juhász, Réka, Nathan Lane, and Dani Rodrik. 2024. "The New Economics of Industrial Policy." *Annual Review of Economics* 16: 213–42.
Kalouptsidi, Myrto. 2014. "Time to Build and Fluctuations in Bulk Shipping." *American Economic Review* 104 (2): 564–608.
Kalouptsidi, Myrto. 2018. "Detection and Impact of Industrial Subsidies: The Case of Chinese Shipbuilding." *Review of Economic Studies* 85 (2): 1111–58.
Kim, Minho, Munseob Lee, and Yongseok Shin. 2021. "The Plant-Level View of an Industrial Policy: The Korean Heavy Industry Drive of 1973." NBER Working Paper 29252.
Krueger, Anne O., and Baran Tuncer. 1982. "An Empirical Test of the Infant Industry Argument." *American Economic Review* 72 (5): 1142–52.
Lane, Nathan. 2020. "The New Empirics of Industrial Policy." *Journal of Industry, Competition and Trade* 20 (2): 209–34.
Lane, Nathan. Forthcoming. "Manufacturing Revolutions: Industrial Policy and Networks in South Korea." *Quarterly Journal of Economics*.
Liu, Ernest. 2019. "Industrial Policies in Production Networks." *Quarterly Journal of Economics* 134 (4): 1883–1948.
Lloyd's. 1892–1999. *Lloyd's Register World Fleet Statistics, 1999*. London: Heritage and Education Center. https://hec.lrfoundation.org.uk/archive-library/world-fleet-statistics (accessed in August 2024).
Lorenz, Edward H. 1991. "An Evolutionary Explanation for Competitive Decline: The British Shipbuilding Industry, 1890–1970." *Journal of Economic History* 51 (4): 911–35.
Luzio, Eduardo, and Shane Greenstein. 1995. "Measuring the Performance of a Protected Infant Industry: The Case of Brazilian Microcomputers." *Review of Economics and Statistics* 77 (4): 622–33.
Miravete, Eugenio J., and María J. Moral. 2024. "Trade Protection, Industrial Policy, and the Shaping of Local Preferences." Unpublished.
Neumark, David, and Helen Simpson. 2015. "Place-Based Policies." In *Handbook of Regional and Urban Economics*, Vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange,

1197–1287. Amsterdam: North-Holland.

Nevo, Aviv, and Michael D. Whinston. 2010. "Taking the Dogma Out of Econometrics: Structural Modeling and Credible Inference." *Journal of Economic Perspectives* 24 (2): 69–82.

Noland, Marcus, and Howard Pack. 2003. *Industrial Policy in an Era of Globalization: Lessons from Asia*. Washington, DC: Institute for International Economics.

Okuno-Fujiwara, Masahiro. 1991. "Industrial Policy in Japan: A Political Economy View." In *Trade with Japan: Has the Door Opened Wider*, edited by Paul Krugman, 271–304. Chicago: University of Chicago Press.

Pack, Howard, and Kamal Saggi. 2006. "Is There a Case for Industrial Policy? A Critical Survey." *World Bank Research Observer* 21(2): 267–97.

Pollard, Sidney. 1957. "British and World Shipbuilding, 1890–1914: A Study in Comparative Costs." *Journal of Economic History* 17 (3): 426–44.

Rodrik, Dani. 2009. "Industrial Policy: Don't Ask Why, Ask How." *Middle East Development Journal* 1 (1): 1–29.

Rotemberg, Martin. 2019. "Equilibrium Effects of Firm Subsidies." *American Economic Review* 109 (10): 3475–3513.

Saunders, S. C. 2015. *Jane's Fighting Ships 2014-2015: Yearbook*. London: IHS Inc.

Stopford, Martin. 2009. *Maritime Economics*. New York: Routledge.

Stråth, Bo. 1987. *The Politics of De-industrialisation: The Contraction of the West European Shipbuilding Industry*. London: Croom Helm.

Thompson, Peter. 2001. "How Much Did the Liberty Shipbuilders Learn? New Evidence for an Old Case Study." *Journal of Political Economy* 109 (1): 103–37.

Thornton, Rebecca Achee, and Peter Thompson. 2001. "Learning from Experience and Learning from Others: An Exploration of Learning and Spillovers in Wartime Shipbuilding." *American Economic Review* 91 (5): 1350–68.

Verhoogen, Eric. 2023. "Firm-Level Upgrading in Developing Countries." *Journal of Economic Literature* 61 (4): 1410–64.

WTO. 2006. *World Trade Report 2006: Exploring the Links between the Subsidies, Trade and the WTO*. Geneva: WTO.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 81–110

Semiconductors and Modern Industrial Policy

Chad P. Bown and Dan Wang

Semiconductors, also known as chips or integrated circuits, are the tiny pieces of machine-crafted silicon that play an essential role in all digital technologies. These include everything from microwaves and toasters to smartphones and 5G communications networks, as well as automobiles, advanced weapons systems, and emerging tools for artificial intelligence. Semiconductors are, in short, intertwined with technological leadership, economic prosperity, jobs, and even national security.

For governments, the semiconductor industry has been an irresistible target for industrial policy. The sector is science-based and fast-moving. It has vast capital requirements: a new semiconductor fabrication plant, or "fab," or "foundry," now costs on the order of $20 billion. Learning-by-doing is important in the manufacturing process, having the potential to spill over to other parts of the economy and thus creating a possible efficiency role for government intervention. Thus, with an industrial policy based on subsidies for research and development or for capital/construction costs—or perhaps even for short-term protection from import competition—policymakers hope to gain a lasting first-mover advantage for their local chip sector.

Some policymakers may hope for an even greater ultimate prize. The origins of semiconductor manufacturing in Silicon Valley have become the canonical example of "agglomeration externalities." For economists, this term describes

■ *Chad P. Bown is Reginald Jones Senior Fellow, Peterson Institute for International Economics, Washington, DC. Dan Wang is a Fellow, Paul Tsai China Center, Yale Law School, New Haven, Connecticut. Their email addresses are cbown@piie.com and danwyd@gmail.com.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.81.

the phenomenon of economies of scale at the level of local industry; that is, industry-wide average costs falling as more output was produced. In an agglomeration economy, knowledge grows and spreads as workers share ideas within and across firms. Multiple companies enjoy access to the same local pool of specialized workers and input suppliers, as well as access to customers for cutting-edge products. In Silicon Valley, over time this mixture would expand to include upstream toolmakers, chip manufacturers themselves, downstream users like computer and telecommunications companies, and now also digital platforms and software companies at the forefront of artificial intelligence. The success of Silicon Valley is one that many other countries would like to replicate. They too want a self-sustaining ecosystem for generating, producing, and then regenerating cutting-edge technologies.

But modern industrial policy is also grappling with one other central challenge—that these agglomeration externalities may lead to the excessive geographic concentration of semiconductor manufacturing. Something has arguably gone too far. Today's heightened risk of localized shocks stemming from climate change (say, in the form of extreme storms or droughts), public health emergencies (like shutdowns tied to the COVID-19 pandemic), or even geopolitical tensions (blockades, invasions, or war) have spurred policymakers into seeking more diversified sources of semiconductor production.

In this essay, we begin with a review of the early US dominance of the semiconductor industry, and then the move to globalization of the sector in the 1980s and 1990s. We consider three main traits that define the modern industry: the rise of the fabless foundry model, in which chip design and manufacturing are done by different firms; the fragmentation of the semiconductor supply chain; and the global shifts in demand for and supply of semiconductors. We then describe two recent issues that are driving concerns about the chip sector: the rise of the semiconductor industry in China and the riskiness of concentrated production in certain other parts of East Asia. These issues have implications beyond the standard industrial policy topics of market and technological leadership. They raise issues related to risks of supply disruptions and weaponizing trade dependencies, as well as the future use of semiconductors in areas like weapons, surveillance, and artificial intelligence.

We then review how governments are implementing industrial policy. In the United States, a primary tool is the CHIPS Act of 2022—formally called the Creating Helpful Incentives to Produce Semiconductors Act. Meanwhile, China and other major economies in East Asia and Europe are deploying industrial policies of their own. Though subsidies and import tariffs have retained importance in the activist government tool kit, additional policies like export controls, foreign investment screening, and even merger reviews are increasingly used as well. In the conclusion, we point out that the semiconductor industry, despite its public prominence, is under-researched. We suggest some of the questions and topics that might usefully be investigated as the current global wave of semiconductor industrial policy proceeds.

Early US Dominance of the Semiconductor Industry

Semiconductor chips evolved from the transistor, which was invented in the late 1940s in New Jersey at Bell Labs—the research arm of American Telephone and Telegraph (AT&T)—by a team of scientists who would later win the 1956 Nobel Prize in physics. Packing large numbers of transistors onto a small chip ultimately resulted in the integrated circuit etched on a silicon wafer, and packing more and more transistors onto those integrated circuits meant faster and more powerful electronic applications. By 1965, Gordon Moore, who later founded Intel, would famously predict that the number of transistors on a chip would double roughly every two years, a pattern that became known as Moore's Law and which has held true for half a century (Roser, Ritchie, and Mathieu 2023).

In this early period, the primary form of US industrial policy toward semiconductors was through direct purchases for the military and space programs—the destination for about half of US production of integrated circuits in the early 1960s (Tilton 1971, Tables 4–8). However, these forms of industrial policy became relatively less important as private sector demand for semiconductors surged. Pocket calculators, for example, were an early driver of chip demand. Semiconductors became a standard input into telecommunications equipment, consumer electronics, computers, and more.

Through the 1970s, US firms dominated the semiconductor industry. Texas Instruments, National Semiconductors, Motorola, and Intel were among the top five firms globally in 1980, by revenue. Yet, even the numbers shown in Table 1 underrepresented the size of US semiconductor manufacturing, which featured another set of vertically integrated companies that made chips only for their in-house needs. These "captive" semiconductor manufacturers that produced for internal demand included AT&T and IBM—at the time, the latter was one of the largest semiconductor manufacturers in the world (USITC 1993, p. 7). Because the "captive" business model did not involve arms-length sales, these companies were often omitted from industry lists defined in terms of revenues. Companies like Intel or Motorola that sold semiconductors on the open market were "merchant" firms. Some, like Texas Instruments, did both.

The Japanese Challenge, Activist Industrial Policy, and the Rise of a Global Semiconductor Market

While European companies like Philips had long been major players in the sector, the emergence of Japanese firms in the 1970s and 1980s posed the first major threat to US chipmakers' dominance of the US semiconductor market.[1] Two differences between the Japanese and US models stood out.

[1] This section draws heavily from Irwin (1996), the seminal political-economy study of the US-Japan trade dispute over semiconductors during the 1980s.

Table 1
Top 10 Global Semiconductor Firms, by Sales Revenue, 1980–2020

Ranking	*1980*	*1990*	*2000*	*2010*	*2020*
1	Texas Instruments	NEC (Japan)	Intel	Intel	Intel
2	National Semiconductor	Toshiba (Japan)	Samsung (South Korea)	Samsung (South Korea)	Samsung (South Korea)
3	Motorola	Intel	NEC (Japan)	TSMC (Taiwan, foundry)	TSMC (Taiwan, foundry)
4	Philips (Europe)	Hitachi (Japan)	Texas Instruments	Texas Instruments	SK Hynix (South Korea)
5	Intel	Motorola	Toshiba (Japan)	Toshiba (Japan)	Micron
6	NEC (Japan)	Texas Instruments	STMicro (Europe)	Renesas (Japan)	Qualcomm (fabless)
7	Fairchild Semiconductor	Fujitsu (Japan)	Motorola	SK Hynix (South Korea)	Broadcom (fabless)
8	Hitachi (Japan)	Mitsubishi (Japan)	Micron	STMicro(Europe)	Nvidia (fabless)
9	Toshiba (Japan)	National Semiconductor	Hyundai (South Korea)	Micron	Texas Instruments
10	Mostek	Philips (Europe)	Hitachi (Japan)	Qualcomm (fabless)	Apple* (fabless)

Source: Compiled by the authors from Brown and Linden (2009, Table 1.1); IC Insights Research Bulletin (2012, 2021).
Note: Shaded companies are domiciled in the United States. In 2001, SK Hynix completed its separation from Hyundai. In 2009, NEC and Renesas Technology merged, forming Renesas Electronics. In 2018, Broadcom redomiciled from Singapore to the United States.
*Custom devices for internal use.

One difference involved the role of government, with Japan taking a more activist form of industrial policy toward the semiconductor sector. Japanese government support included industry tax breaks—since replicated elsewhere—as well facilitating a consortium of domestic firms that would pool resources in an effort to prevent redundant spending on research and development, through the Very Large Scale Integrated Circuits (VLSI) project of 1976–1979 (Okuno-Fujiwara 1991).

A second difference stemmed from industrial structure. Many Japanese semiconductor firms were vertically integrated—similar to, say, IBM or AT&T in the United States—and thus benefited from internal demand for their chips. However, the vertically-integrated US suppliers mostly kept their production in-house, worried about antitrust authorities questioning the terms of their sales to competitors. Unlike their US counterparts, Japanese firms also sold their semiconductors on the US market. Meanwhile, the other half of the US chipmaking industry that only manufactured for arms-length sales found it difficult to penetrate the Japanese market, where demand was driven by those vertically integrated Japanese companies. Some "captive" US semiconductor firms like IBM did have foreign

direct investment operations in Japan and accessed the Japanese market in this way (Irwin 1996).

Furthermore, Japanese firms were part of *keiretsu*, or business conglomerates. These included affiliations with a large bank that helped facilitate investments into capital expenditure—which Japanese firms did much more than US firms during this period (OECD 1992, 146–47). Access to credit would allow Japanese companies to expand production even during market downturns, which was important for an industry characterized by boom-bust cycles, and not something that US companies could match (Irwin 1996; Hoshi, Kashyap, and Sharfstein 1990).

In general, exports from Japan became ever more visible in the US economy during this time, starting with less technologically sophisticated sectors like clothing and footwear in the 1960s, and then in the 1970s and 1980s proceeding to steel, consumer electronics, automobiles, and ultimately chips. Japan's increasing industrial competitiveness stemmed from many sources, including its very high rates of domestic saving and investment, as well as elevated US interest rates and a strengthening US dollar that made imports from Japan relatively cheaper. When the United States began to run a large and growing trade deficit in the 1980s, Japan was the country with the largest bilateral surplus. Imports from Japan were a tremendous source of trade conflict at the time, leading to alarmist predictions of decline for the entire US economy (Prestowitz 1988; Thurow 1992). This situation culminated in the United States pursuing "aggressively unilateral" trade policy, including toward Japan (Bhagwati and Patrick 1990; Bergsten and Noland 1993).

For semiconductors, the result was a highly interventionist US policy, in which the nature of industrial policy shifted to attempts to manage and regulate foreign trade. Under the threat of US import tariffs, Japan "voluntarily" agreed in the US-Japan Semiconductor Trade Agreement of 1986 to limit exports to both the US and third country markets. The Japanese government also "voluntarily" agreed to expand Japan's *imports* of chips—specifically, US firms were to supply 20 percent of the Japanese market by 1992. When goals were not met, the US government retaliated with import tariffs, including on Japanese computers and televisions that used semiconductors as inputs. Such aggressive use of trade policy against an ally was unusual and partly made possible because of Japan's reliance on the United States for military protection.

The US government also decided to emulate some elements of the Japanese approach toward the semiconductor industry. For example, the US Department of Defense provided $100 million annually for five years beginning in 1988 to SEMATECH (SEmiconductor Manufacturing TECHnology), a public-private partnership. Though its performance would face mixed reviews, SEMATECH involved 14 US-based semiconductor firms forming a consortium designed to share the burden of research and development costs that some felt were holding back the US industry (Irwin and Klenow 1996).

One short-term effect of the US-Japan Semiconductor Trade Agreement was to stabilize semiconductor prices, which had been an objective of US semiconductor firms. But using bilateral trade policy as an industrial policy tool to push for higher

price levels also created opportunities for chip companies in other economies—who had often benefited from their governments' industrial policies—to enter the market profitably.

For example, Taiwan Semiconductor Manufacturing Corporation (TSMC) emerged in 1987 as the world's first contract manufacturer of chips designed by other companies. The Taiwanese government provided $100 million to help TSMC construct a foundry that would focus on manufacturing chips designed by Philips and other companies (Landler 2020; Chang and Hsu 1998). Other firms also invested in Taiwan at the time, including Texas Instruments. All told, Taiwan's share of US semiconductor imports doubled between 1989 and 1999, from 4.5 percent to 9 percent.[2]

In South Korea, Samsung, Goldstar and Hyundai—the latter two would one day combine to become part of today's SK Hynix—also emerged as global competitors in making semiconductors.[3] During this era, the South Korean government provided support to conglomerate *chaebols*, a form of industrial structure similar to Japan's *keiretsu*; for example, Samsung and Hyundai received subsidized credit (Kim 1998). South Korea's share of the US import market for semiconductors grew from 6 percent in 1987 to 16 percent by 1999. By 2000, Samsung and Hyundai had joined the ten largest semiconductor companies in the world (again, see Table 1). They started by focusing on "memory" chips, which store data for retrieval rather than "logic" chips that process data. It is noteworthy that Japan and South Korea each developed their chip industries via mastery of memory chips. Industry success has long been characterized by process technology improvements, where learning-by-doing meant increasing "yields," or getting more usable chips from each batch of production. In a study of dynamic random access memory (DRAM) over 1974–1992, Irwin and Klenow (1994) found per unit production costs fell 20 percent every time cumulative output doubled (for additional discussion of learning-by-doing for semiconductor production, see also Baldwin and Krugman 1988; Dick 1991).

In the United States, Europe, and Japan, the new competition from South Korea and Taiwan resulted in renewed industry demands for import protection. Though Micron had licensed some of its technology to Samsung in the early 1980s, in 1992 it changed tack and filed an antidumping petition against the Korean firm, seeking import tariffs and alleging injury caused by underpriced chips. Around the same time, Motorola's UK subsidiary and the German firm Siemens asked for similar protection in Europe from the Korean memory chipmakers. In 1997, Micron demanded (and received) US antidumping duties on imports of Taiwanese semiconductors. Following the Asian financial crisis in the late 1990s, Micron and its affiliates also requested and received anti-subsidy (countervailing) duties on Korean memory imports across three different markets—the United States, Japan, and European Union.

[2] This section draws from the trade and policy data presented in detail in Bown (2020).

[3] In 1995, Goldstar changed its name to LG Electronics, which then merged with Hyundai in 1999 to form Hynix. In 2012, Hynix partnered with SK and changed its name to SK Hynix.

The memory segment of the semiconductor industry was also consolidating heavily, with Micron, Samsung, Hynix, and Infineon (a spinoff of Siemens) dominating the market by the early 2000s. This led to conflicting US policy signals. While one part of the federal government was worried about subsidized imports being priced too low, US antitrust authorities became simultaneously concerned that memory chip manufacturers were colluding to raise prices, hurting computer companies such as Dell, Hewlett-Packard, Apple, and IBM. Between 2004 and 2005, Samsung, Hynix and Infineon all pled guilty to fixing memory chip prices and paid criminal fines, while firm executives served prison terms (US Department of Justice 2005).

Any industrial policy history of the semiconductor sector from this era runs a risk of sounding like nothing but trade barriers and disputes. But in the mid-1990s and early 2000s, two trade-facilitating policy developments would also ultimately serve as a counterweight to help globalize the industry's supply chains. First, these major economies implemented international agreements featuring a general reduction of import tariffs for semiconductors and critical inputs, such as semiconductor manufacturing equipment. This included the 1997 Information Technology Agreement, a deal to cut tariffs to zero on a wide range of high-tech products, as well as China and Taiwan joining the World Trade Organization (WTO), which locked in their low tariffs, in 2001 and 2002, respectively. The second policy innovation involved the TRIPs (Trade Related Aspects of Intellectual Property Rights) Agreement, implemented as part of the establishment of the WTO in 1995. Improved protection of patents and other trade secrets would help facilitate the fabless-foundry model—one firm licensing its technology to another firm for manufacturing purposes, without fear that the first firm would lose its intellectual property to another of the manufacturer's clients.

The Modern Semiconductor Industry: Three Characteristics

The semiconductor industry has changed in terms of how and where chips would be produced. Here, we emphasize three characteristics of the modern industry that subsequently shaped the context for today's policymakers.[4]

The Rise of the Fabless Foundry Model

The structure of the semiconductor industry has evolved considerably over the past three decades. Begin with memory chips that store data, which continued to make up roughly 23 percent of global industry sales in 2022 (SIA 2023a). The most noticeable feature of the memory chip segment is its continued consolidation—Samsung, SK Hynix, and Micron, for example, currently make up nearly all of the lucrative DRAM market. Furthermore, memory chips are the most commoditized

[4]For book-length treatments of the evolution of the semiconductor industry, see Brown and Linden (2009) and Miller (2022).

of semiconductor technologies: products from Samsung or Micron are largely interchangeable, and some memory chips are even sold on a spot market. Still, they are likely the most complex of any products that are commoditized.

Aside from memory, the rest of the semiconductor market includes logic, analog, and a variety of other kinds of chips that perform different functions. These types of semiconductors are also sometimes characterized by their vintage. At one extreme might be the latest graphics processing unit (GPU) chip that is needed to run today's most powerful large language model for artificial intelligence. At the other extreme are "legacy" or "mature" semiconductors that have been around for a while. Though firms may not require the latest technology to manufacture these older types of chips, the products remain complex. A semiconductor that goes into an automobile that powers a window, for example, may be little different from one of ten years ago. Yet, it has other critical characteristics—such as reliability and durability—allowing it to survive extreme temperature changes over long periods of time without replacement.

For the nonmemory segment of the market, integrated device manufactures such as Intel and Texas Instruments remain important players. Nevertheless, perhaps the greatest change in industry structure has been the rise of the fabless-foundry model. Some companies (especially in Silicon Valley) have decided to focus solely on the design of logic chips, while contracting with specialized foundries (mostly in Asia) to manufacture them. The most prominent foundry company is Taiwan's TSMC, the industry pioneer, which emerged as a top ten firm by revenue by the 2000s (as shown in Table 1). Other examples of such "pure-play" foundries include UMC (Taiwan), GlobalFoundries (United States), and SMIC (China). A large share of their expenditures involves the physical plants and capital equipment needed to run state-of-the-art facilities. Again, by the 2020s, the cost of building and equipping a leading-edge fab was above $20 billion.

The complements to foundries are "fabless" chip companies (as shown in Figure 1). They focus only on design. For example, Broadcom developed chips for modems, routers, and telecommunications networks, while Qualcomm designed semiconductors for smartphones and other devices (Nellis and Mehta 2023). Apple has become a major semiconductor player in its own right by replacing chips from Intel, Qualcomm, and others in its computers and mobile phones. Nvidia has become prominent for its GPU chips that are in demand with the growth of artificial intelligence applications (Waters 2023). These design companies devote the bulk of their costs to research and development, while letting the foundries worry about capital-intensive manufacturing. By 2020, four of the top ten semiconductor firms by revenue were fabless. But without the emergence of contract manufacturers such as TSMC, fabless firms likely would not exist.

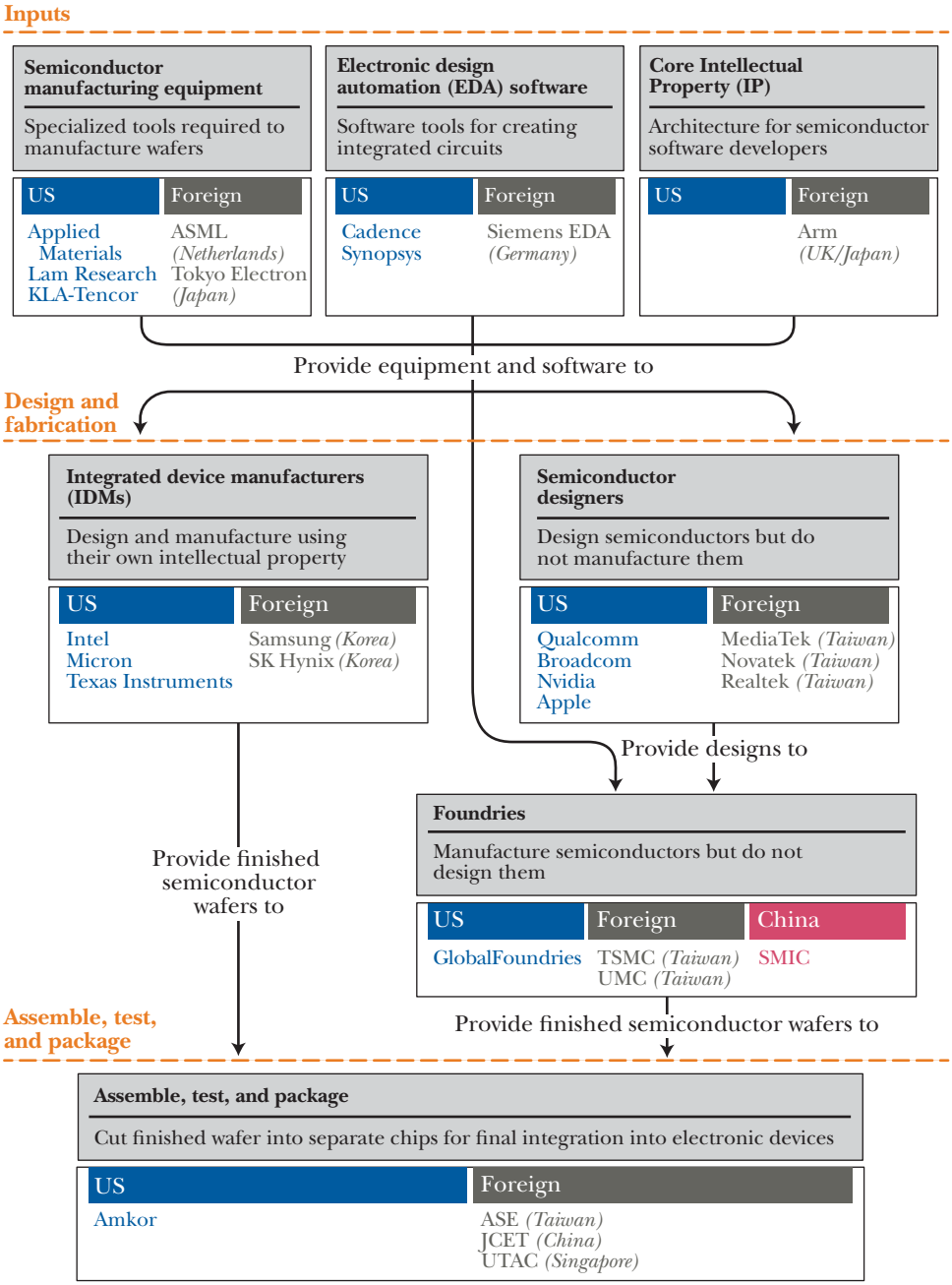
The Fragmentation of Semiconductor Supply Chains

Beyond design and manufacturing of semiconductors, Figure 1 illustrates other critical elements of the modern semiconductor supply chain. For example, the last step involves taking a finished "wafer" and putting it through a process known as

Figure 1

Modern semiconductor manufacturing is a globally integrated, multi-stage process

The stages and examples of companies involved in the semiconductor design and manufacturing supply chain



Source: Constructed by the authors.
Note: Examples of companies are illustrative.

"assembly, test, and package" by cutting the wafer into separate chips for final integration into electronic devices. This phase is relatively worker-intensive and thus is not only often outsourced to a different company, but is also often offshored to a labor-abundant country where wage costs are lower. Indeed, the assembly, test, and package segment was one of the first parts of the semiconductor supply chain to be moved overseas in the 1960s, when Fairchild Semiconductor set up such a process in Hong Kong.

The upstream direction of the supply chain includes key input providers. One input is the software from electronic design automation firms, currently dominated by two US-based companies, Cadence and Synopsys, as well as the German firm Siemens EDA.[5] Many semiconductor companies are also reliant on the intellectual property input—or "Core IP"—of Arm, a firm headquartered in the United Kingdom and owned by a Japanese financial institution (Softbank).

For physical inputs, five companies—three in the United States (Applied Materials, Lam Research, KLA-Tencor), one in the Netherlands (ASML), and one in Japan (Tokyo Electron) dominate the provision of capital equipment used in these $20 billion fabs. ASML plays an outsized role as the only firm to make the extreme ultraviolet lithography equipment required to produce the most advanced semiconductors, including those used in artificial intelligence and weapons systems (Bounds 2023; Bradshaw and Gross 2023).

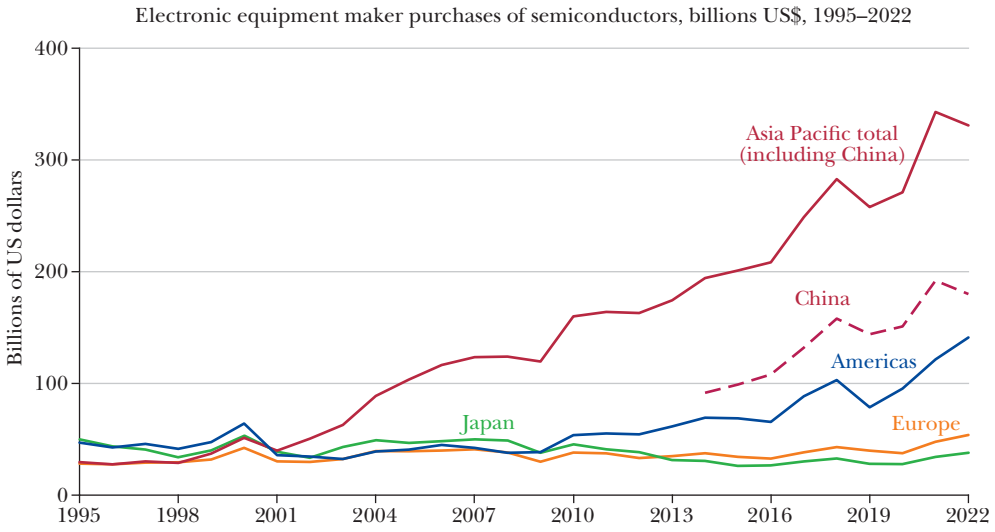
Geography of Semiconductor Manufacturing

While the overall semiconductor supply chain meanders around the world, the physical manufacturing at foundries and the tasks of assembly, test, and package have gravitated both toward each other and toward the location of downstream demand for many of those chips. SIA (2023a) estimates that 70 percent of end users of chips are companies making consumer electronics, computers, and telecommunications equipment. Assembly of such products became increasingly concentrated in China over the 2000s, as firms like Apple hired contract manufacturers to put together iPhones and other devices using low-cost labor.

Figure 2 illustrates the geographic distribution of demand emanating from such electronic equipment end users. By 2020, 62 percent of demand from these end users was located in Asia, and 34 percent of global demand came from China alone. Furthermore, most of the major US-headquartered chipmakers—both integrated device manufacturers and fabless designers—counted China as a major destination for their sales; for example, China accounts for roughly 30 percent of semiconductor sales of Intel, Broadcom, and Nvidia, and half or more of sales of Texas Instruments and Qualcomm (based on authors' calculations from annual 10K company reports).

[5] In 2017, Siemens acquired what was then called Mentor Graphics—a firm focused on electronic design automation tools for chip-making headquartered in the United States—and renamed it Siemens EDA in 2021.

Figure 2
China has become a major source of demand for chips



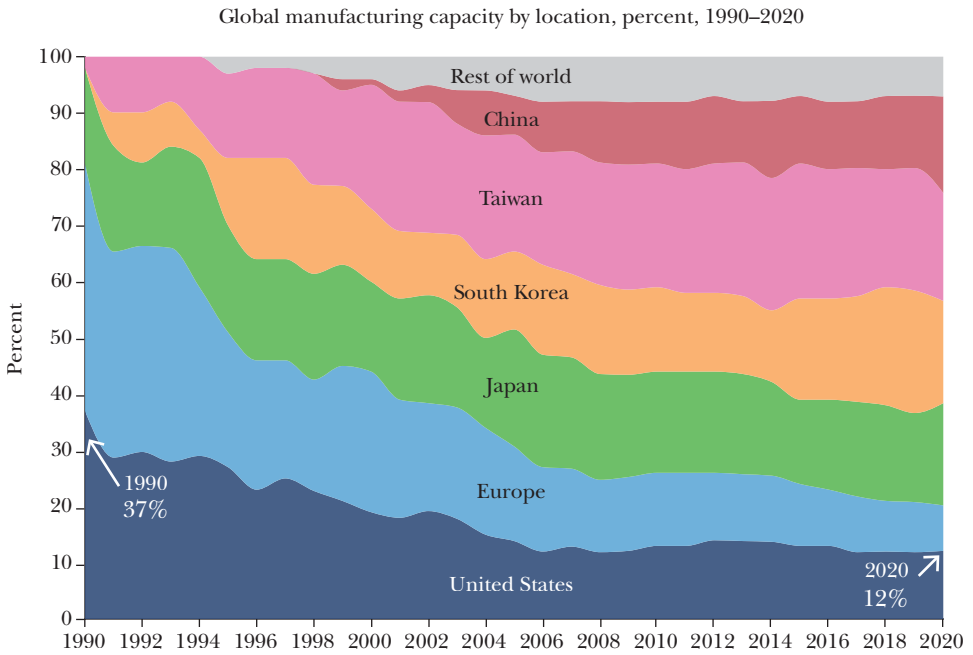
Source: SIA (2023a).
Note: China-specific data are only available since 2014.

On the supply side, the location of semiconductor manufacturing also became very concentrated in East Asia, as shown in Figure 3. As a byproduct, the US share of global semiconductor manufacturing capacity fell dramatically—from 37 percent in 1990 to 12 percent in 2020. The United States was not alone: Europe experienced a similarly sizeable decline in its share of global manufacturing; Japan's share also fell. However, US firms still play a very important role in the global semiconductor industry. For chip production in Taiwan by TSMC, for example, the foundry is often manufacturing chips designed by American firms like Qualcomm and Nvidia. Also, US-headquartered companies that continued to manufacture chips have expanded outside of the United States through foreign direct investment—like Intel (Ireland, Israel) and Micron (Japan, Singapore, Taiwan)—either by building plants or by acquiring foreign facilities.

The increasing concentration of manufacturing in Asia was not due to only market forces: foreign industrial policy continued to play a role. The Semiconductor Industry Association, for example, argued in 2020 that it was 30 percent more costly to operate a fab in the United States relative to Taiwan or South Korea and up to 50 percent more costly than in China. The SIA also estimated that 40–70 percent of that cost differential was due to relatively higher foreign government subsidies (Varas et al. 2020, pp. 14–20).[6]

[6]The Semiconductor Industry Association is certainly not a dispassionate observer. However, whereas the SIA of the 1980s was clearly an industry association dominated by American semiconductor

Figure 3
Over time, the supply of semiconductor manufacturing capacity shifted toward Asia



Source: Varas et al. (2020, Exhibit 2).
Note: Values for 2020 are estimates.

Overall, US-headquartered firms continued to play essential roles in the global semiconductor industry. Intel remains one of the leading chip manufacturers, and four US fabless firms also made the top ten of all semiconductor companies in 2020 by revenue. Other US companies dominated electronic design automation tools, and a third set provided much of the most essential semiconductor manufacturing equipment. Yet, the reduced share of chip manufacturing plants has proved worrisome for US policymakers.

manufacturing firms, by the 2020s, its membership had evolved considerably. In 2023, SIA counted among its members integrated device manufacturers (Intel, Micron), fabless firms (Broadcom, Qualcomm, Nvidia), pure-play foundries (GlobalFoundries), equipment manufacturers (Applied Materials, Lam Research, KLA-Tencor), and electronic design automation tool providers (Cadence, Synopsys). SIA also had international members including TSMC, Samsung, SK Hynix, Infineon, NXP, Arm, Tower, and ASML (SIA Members, https://www.semiconductors.org/about/members/, accessed December 31, 2023).

Current Issues in the Global Semiconductor Industry: China and the Concentration of Production

By the mid-to-late 2010s, the emerging global semiconductor industry was raising two interrelated sets of risks for many politicians: What about China? What about the increasing geographic concentration of production for the most advanced chips?

China

The Chinese government has been blatant about its desire to achieve self-sufficiency and technological leadership across a range of industries. The semiconductor industry is perhaps the most critical sector in which China is neither self-sufficient nor much of a technological leader (Wang 2023a). In 2014, the State Council developed its "Guidelines to Promote the Development of the National Integrated Circuit Industry," which established major funding for domestic chip companies. Beijing also revealed the "Made in China 2025" industrial policy in 2015, which set aggressive numerical targets for the future market shares of Chinese chip firms in China as well as globally, heightening concerns of policymakers elsewhere.

China's industrial policies for semiconductors were also now taking place in a changing geopolitical climate under President Xi Jinping. Flashpoints included China's gradual subjugation of Hong Kong, its military provocations through the shipping lanes of the South and East China Seas, its increasingly aggressive "wolf warrior" diplomacy, and its intention of annexing Taiwan (Harrell, Rosenberg, and Saravalle 2018). Beijing's "Military-Civil Fusion" policy also explicitly encouraged companies in China to share their technologies to upgrade the military readiness of the People's Liberation Army (Ford 2019). Other concerns included Chinese state-sponsored espionage for military, intelligence, or corporate gain, as well as its growing efforts at large-scale surveillance that threatened human rights, including with respect to repression of Uyghurs in Xinjiang.

China offers a standard menu of subsidies for the semiconductor industry common to more advanced economies, including tax credits for research and development, land concessions, and direct subsidies. For example, in 2020, China's State Council (2020) announced that it would eliminate corporate income taxes of advanced semiconductor fabs for ten years. Public companies listed on stock markets report the subsidies they have received from various levels of the Chinese government; by one tally, these reached $1.75 billion for 190 firms in 2022 (Cao 2023).

In addition, China assists its national champions by more opaque means. The China Integrated Circuit Industry Investment Fund, also known as the Big Fund, has raised tens of billions of dollars to support the local industry since its launch in 2014 (Liu and Leng 2023). OECD (2019), for example, found this mechanism has provided an especially high amount of below-market debt and equity financing to Chinese companies, with SMIC and Tsinghua Unigroup receiving government support over five years that exceeded 30 percent of their annual revenue. In addition, the Chinese government maintains talent recruitment programs targeting

engineers—from Taiwan, South Korea, and elsewhere—to work for its domestic companies. Finally, the US government has accused Beijing of running broader cyber-intrusion campaigns, seeking to steal secrets from technology companies.

China has also sought access to foreign technology by acquiring western companies, though rarely with success. In 2015, Tsinghua Unigroup attempted to buy Micron (Baker and Roumeliotis 2015). Fujian Grand Chip Investment Fund wanted, in 2016, to purchase Aixtron, a German company whose technology was used to upgrade the Patriot missile defense systems (Sheahan 2016). These and other potential deals were either discouraged or prevented by the US government. The Committee on Foreign Investment in the United States (CFIUS) has legal authority to stop company mergers or acquisitions that threaten national security. Other examples of CFIUS blocking semiconductor industry acquisitions include the attempted takeover by the Chinese investment firm Canyon Bridge Capital Partners of Lattice Semiconductor Corporation, as well as the efforts by Broadcom—headquartered in Singapore at the time—which sought to acquire Qualcomm in 2018 (for more on CFIUS, see Congressional Research Service 2018).[7]

China did not take these actions lying down. In 2015, it imposed fines of nearly $1 billion against Qualcomm in an antitrust action (Dou 2015). In 2018, the Chinese government refused to greenlight Qualcomm's potential acquisition of NXP, a Dutch firm (Martina and Nellis 2018). More than half of Qualcomm's sales at the time were to companies in the Chinese market, as mentioned earlier, and at risk if the Chinese government objected to the acquisition.

Given China's economic size, its industrial policies can also disrupt global production and the allocation of resources. China's earlier efforts to dominate global production capacity—for example, in shipbuilding, steel, aluminum, and solar panels—have often created serious political-economic problems for other major economies whose firms and workers were put under unrelenting economic stress by a nonmarket actor.[8]

The US government has been uneasy about China's semiconductor ambitions for some time. At the end of the Obama administration, the White House published a report on semiconductors to warn that "a concerted push by China to reshape the market to favor their needs threatens the competitiveness of US industry" (Mundie and Otellini 2017). In principle, World Trade Organization rules can help trading partners to address some of China's actions, such as its high levels of subsidies, its treatment of foreign intellectual property, and conditioning access to the Chinese market on transferring technology to local firms. However, under the Trump administration, the United States turned away from using the WTO to tackle concerns with China (Bown and Keynes 2020; Bown 2021a) and instead deployed import

[7] See CRS (2018, pp. 5–7) for a discussion of the US Congressional backlash following a Japanese firm's attempted acquisition of Fairchild Semiconductor in 1987 and how the "Exon-Florio" amendment to the Defense Production Act changed the Committee on Foreign Investment in the United States process by which foreign investments are reviewed.

[8] For details on the shipbuilding example, which mostly impacted firms in Japan and South Korea, see Barwick, Kalouptsidi, and Zahur (forthcoming), as well as their paper in this symposium.

tariffs (USTR 2018). Without US backing, no other country was willing to invoke the WTO to formally question China's industrial policies.

However, for all of China's efforts to become a leader in semiconductor technologies, its track record to date is decidedly mixed. In no segment of the semiconductor supply chain can Chinese firms claim leadership, although there are a few where they are not many years behind. Khan, Mann, and Peterson (2021) are an early attempt that relies on revenue-based measures to assess China's role in different segments of the global semiconductor supply chain. Chinese firms are competitive in assembly, test, and package, for example, though this is a low value-added part of the supply chain. They also play some role in the design of logic chips and the production of memory chips. Yet, in the manufacturing of logic chips, China's SMIC remains several years behind TSMC. Chinese firms are also weak in the production of semiconductor equipment and the software companies creating electronic design automation tools. The main area where they are a global player is in the volume of production for the less-complex legacy chips, as further described below.

China's inability to catch up to the global technological frontier thus far is likely due to several factors. It is a latecomer. China's chip industry began in earnest only in the late 1990s, which is decades behind the leading firms from the United States, Europe, Japan, South Korea, and even Taiwan. The semiconductor industry has tended to favor incumbents, as the pace of innovation in the industry is rapid and unforgiving. Chinese semiconductor firms have struggled, in part due to their smaller commercial scale and lack of experienced personnel. Finally, the United States and other governments have, since 2015, more aggressively wielded export controls in ways that may hobble China's chip progress (as described further below).

The Geographic Concentration of High-End Semiconductor Production

The other main emerging worry for western policymakers was the extreme geographic concentration for the manufacturers of the most advanced semiconductors. Two companies—TSMC in Taiwan (92 percent) and Samsung in Korea (8 percent)—dominated world production of the smallest and fastest chips, defined as semiconductor nodes below ten nanometers (Varas et al. 2021). This concentration raises various risks: geographically focused shocks due to extreme weather events, earthquakes, or public health emergencies, as well as geopolitical shocks due to risks of military confrontation with China or North Korea.

The global semiconductor shortage of 2021 stoked these fears. Especially frustrating was the unavailability of legacy chips—semiconductors that were not the most difficult to manufacture—but still essential to produce a toaster, refrigerator, microwave, washing machine, or car (Horwitz 2021; Jung-a and Olcott 2021). Indeed, in the United States and Germany, chip shortages shut down parts of the politically influential automobile industry for a time, furloughing workers (Grossman 2021; Miller and Arnold 2021).

Much of the chip shortage was clearly caused by disruptions related to the pandemic, not weaponization of supply chains. Indeed, some of the US chip shortage was even self-inflicted, due to the new 25 percent US import tariffs and Chinese

hoarding of chips induced by the Trump administration's export controls described below (Bown 2021b). Ironically, the geographic concentration of US chipmaking facilities domestically—around Austin, Texas—also contributed to the shortage problems when a freak winter storm hit the region in February 2021, knocking down the electrical grid and throwing offline facilities belonging to Samsung, Infineon, and NXP (Fitch 2021) rather than the geographic concentration of "foreign" production. Nonetheless, the shortage experience spooked policymakers, who argued that firms' private evaluations of geographic location and supply chains did not fully recognize broader social tradeoffs.

This motivation for industrial policy is notably different from the classic arguments about how such a targeted government intervention might be able—through exclusive focus on the *benefits* of agglomeration—to improve firm-level productivity growth and possibly national economic growth.[9]

Current Industrial Policies for the Semiconductor Sector

When it comes to industrial policy for semiconductor manufacturing, the United States, Japan, and Europe have largely been supportive of each other's policies to date. Some modest and informal institutional arrangements have even emerged to facilitate communication across governments seeking to "coordinate" these policies—including through the US-EU Trade and Technology Council as well as US-Japan and EU-Japan bilateral initiatives, which have now also been extended to South Korea (Hayashi 2022; Nagao 2023; Sullivan 2023). That these economies have not (much) challenged each other's subsidies is likely because they have common concerns: China and the geographic concentration of production of advanced nodes in Taiwan by TSMC. Here, we discuss current industrial policies for the semiconductor industry for the United States, other major producers, and China.

The CHIPS Act of 2022 and Other US Policies

US industrial policy for semiconductors is evolving in a number of ways—for example, adjusting its rules for inbound foreign direct investment under the Committee on Foreign Investment in the United States, creating notification requirements impacting outbound foreign investment (Biden 2023), and changing its use of export controls. But a major additional policy change is clearly the Creating Helpful Incentives to Produce Semiconductors (CHIPS) Act of 2022 (for details of the law, see CRS 2023).

The headline provisions of the bill involved $52 billion of subsidies and tax incentives, of which $39 billion were to be spent over five years for building,

[9]This focus on countering the geographic concentration of production was also not unique to semiconductors but has also arisen in industrial policy for supply chains ranging from personal protective equipment (Bown 2022a) to critical minerals needed for electric vehicles (Bown 2024a).

expanding, and equipping fabrication facilities in the United States. Federal expenditures were limited to up to $3 billion per project (although higher amounts could be dispersed with presidential certification to Congress). This amount would only offset a portion of the construction and outfitting costs for a new fab—again, for the most advanced chips, a fab could cost $20 billion or more. For context, in 2021, TSMC alone announced it would spend $100 billion over three years to expand its global production capacity (*Reuters* 2021).

Implementing the CHIPS Act would first require creating administrative capacity within the US government. The Department of Commerce had to establish a new office and hire staff to solicit and evaluate private sector proposals so as to disburse its funds (Shepardson 2022, 2023). Once operational, the office created a five-step application process to allocate funding. A potential applicant for these funds would begin by submitting a "statement of interest" that would be followed by an iterative process with Department of Commerce officials, with companies offering more detailed information before any decisions were finalized (CHIPS for America 2023a). In February 2023, Commerce announced its first "Notice of Funding Opportunity" for semiconductor fabs (NIST 2023a). It would receive over 200 statements of interest over the first six weeks of the program and nearly 600 by the end of 2023 (CHIPS for America 2023b; US Department of Commerce 2024).

There are significant strings attached to CHIPS funding, however, including novel elements that sparked controversy. Funding recipients were expected to offer high-quality childcare to their employees, and also to share any "windfall profits" with US taxpayers (Swanson 2023). The Commerce Department also established a rule that companies could not use CHIPS Act funding to "directly or indirectly benefit foreign countries of concern," including China (NIST 2023b, c). This rule limited what companies could do in China and was especially important to potential funding recipients like TSMC, Samsung, and SK Hynix—firms whose production facilities US policymakers were attempting to attract—each of which already operated multi-billion dollar chip-making plants in China.

Four other novel elements of the CHIPS Act are worth mentioning. First, as part of the $39 billion for manufacturing incentives, $2 billion was set aside to increase US production capacity of legacy chips. These mature semiconductors are especially important for automobiles (and certain military applications); indeed, legacy-chip shortages were some of the most problematic in 2021. However, these chips are not particularly profitable—as such, China's high-volume, low-profit margin, state-supported fabs make unsubsidized production noncompetitive and could end up dominating this segment of the market, resulting in a new concern involving the geographic concentration of production (Hawkins and Leonard 2023).

Second, the CHIPS Act allocated up to $500 million to subsidize the environment for assembly, test, and package facilities in countries *outside* of the United States. This provision recognized the economic difficulty of relocating certain parts of the supply chain to the United States, especially when labor costs play an important role in assembly, test, and package. In 2023, the US Department of State (2023a, b, c, d) announced that it was exploring such partnerships with Panama,

Costa Rica, and Vietnam—Intel already had such facilities in the latter two nations (Guarascio 2023; *Reuters* 2023a).

Third, the CHIPS Act included a 25 percent investment tax credit for capital expenses for manufacturing semiconductors and semiconductor manufacturing equipment.

Fourth, $13 billion of the $52 billion was included for research and development and for workforce development. This program may be able to draw lessons from both the SEMATECH experience as well as more recent efforts to promote research and development elsewhere, including the successful IMEC (Interuniversity Microelectronics Centre) research and development hub in Europe (Beattie 2022).

Implementing a new industrial policy program as set out by the CHIPS Act would take time. Thus, the first funding announcement was not made until December 2023. In a nod to the importance of national security motivating the legislation, the Department of Commerce (2023) awarded the first $35 million to the defense contractor BAE Systems to expand the type of chips used in F-35 fighter jets.

Nevertheless, long before any announcement of CHIPS Act grants, many of the companies expecting to receive funding under the program had already begun construction of new or expanded facilities, publicly expressing their expectation of federal funding (for example, see the press releases from TSMC 2020; Samsung 2021; Intel 2022; Micron 2022), seemingly with the support of policymakers.[10] Indeed, President Joe Biden's official visit to South Korea in May 2022 included a stop at a Samsung plant where he highlighted the company's already announced $17 billion new investment in Texas as well as the need for Congress to quickly pass and appropriate funding for the CHIPS Act to facilitate the completion of that project (Biden 2022).

More generally, the CHIPS Act was only one—and far from the first—of numerous US policies seeking to modify the incentives that affect the decisions of these global companies regarding where to locate production. Preceding the subsidies were the Trump administration's 25 percent import tariffs, as semiconductors were one of the first products caught up in the US-China trade war that began in summer 2018. Over the subsequent three years, the volume of US semiconductor imports from China fell by roughly half (Bown 2022b, Fig. 9). As US chip buyers were unable to completely shift purchases to other foreign suppliers, this, of course, contributed to the shortage of chips available in the United States in 2021.

The new US export controls of October 7, 2022, were another policy designed to affect the location of semiconductor production (BIS 2022; Schuman 2023; Bown 2022c). In the name of national security, the United States planned to limit exports

[10] Furthermore, the broad parameters of how much the industry might be seeking under the CHIPS Act date back at least as far as March 2020, when the Semiconductor Industry Association released a commissioned study finding that the US government would need to replace an estimated $49 billion of lost revenue from purchases by Chinese device manufacturers due to decoupling (Varas and Varadarajan 2020, p. 16).

of the most sophisticated chips and advanced semiconductor manufacturing equipment. (Facilities in China producing older nodes would not be affected.) One year later, the United States further tightened these rules, partially in reaction to US companies like Nvidia and Intel designing chips for the Chinese market that met the letter, but apparently not the spirit, of the original US export controls (BIS 2023; Hayashi 2022). Another contribution to the tightening of the rules may have been the announcement that China's SMIC had managed to manufacture certain advanced semiconductors despite US export controls on SMIC dating back to the Trump administration (O'Keeffe and Fitch 2023).

These announcements were the latest in a deepening set of US export controls involving semiconductors and China. Under the Trump administration, in 2019, the United States began to limit chip exports to China to address national security concerns related to a different sector—critical infrastructure and telecommunications. The US government worried about the Chinese company Huawei's provision of 5G telecommunications equipment, including base stations and cell towers (Bown 2020); indeed, the export controls followed a US Department of Justice (2019) indictment of Huawei for conspiracy, attempted theft of trade secrets, wire fraud, and obstruction of justice. Controls sought to limit advanced node semiconductors being made outside of China, which were an essential input into such Chinese-made 5G equipment.

The initial versions of the United States's export control policies met a fatal flaw. Even if many advanced chips were designed by US companies, they were physically manufactured in Taiwan or South Korea, and thus outside the jurisdiction of the first round of export controls. As an update, US officials then announced that foreign fabs could not use US-made technologies to produce chips for Huawei. Given the United States's dominance of semiconductor production equipment and electronic design automation software, that would prove devastating to foreign fabs. Legally, the US government deployed the "foreign direct product rule," which gave foreign fabs a choice—if they wanted to continue to access US-made inputs (like equipment from Applied Materials, Lam Research, and KLA-Tencor), then they would have to give up selling chips to Huawei and other worrisome Chinese companies. The discovery of this equipment choke point was also key to the US government's later application of export controls on October 7, 2022, affecting China's semiconductor manufacturing sector itself.

Semiconductor Policies by Other Major Economies

Some elements of the US industrial policy toward semiconductors need international cooperation; for example, US export controls would be ineffective if done unilaterally, because companies in other countries would provide the goods instead. Thus, for controls on semiconductor manufacturing equipment, governments of the Netherland and Japan eventually adopted policies similar to the US controls of October 7, 2022—restricting exports of ASML and Tokyo Electron (shown in Figure 1)—in 2023 (Kelly and Uranaka 2023; Government of the Netherlands 2023).

Other countries have also acted alongside US efforts to diversify the location of production globally. Japan, for example, subsidized over $3 billion for TSMC to build a plant on the island of Kyushu (Inagaki 2023). Japan is also providing $1.3 billion to Micron to build a new factory (Nohara 2023), and it has backed Rapidus, a newly formed semiconductor manufacturer, to produce advanced-node chips in Japan, including in partnership with IBM (Kelly and Lee 2022).

Similarly, the European Union passed the European Chips Act in 2023 (Norton Rose Fulbright 2023). TSMC has received an additional €5 billion from the German government, as part of an arrangement with NXP, Infineon, and Bosch, to build a fab in Dresden (Wu and Cantrill 2023; Pasquini 2023). Germany is also providing Intel nearly €10 billion of subsidies for two plants (Heine, Mukherjee, and Rinke 2023), and the Polish government has subsidized Intel's construction of a new assembly, test, and package facility in Poland (Badohal and Mukherjee 2023; Intel 2023). Despite the outbreak of war in nearby Gaza, Intel also announced it was spending $25 billion on expansion of its facilities in Israel after receiving a commitment of over $3 billion of subsidies from the Israeli government (Lu 2023). Finally, the French government provided GlobalFoundries with €2.9 billion to build a new plant with STMicroelectronics in southeastern France (Vidalon and Kar-Gupta 2023).

Taiwan and South Korea have not remained idle as other countries seek to lure their manufacturers in the name of supply chain diversification. As the chip facilities operated in China by their multinationals (TSMC, Samsung, and SK Hynix) have dimming long-term prospects—given the US, Japanese, and Dutch export controls on equipment—these leading-edge firms faced decisions of where to locate production next. To incentivize reshoring, in January 2023, Taiwan passed a law allowing its local semiconductor companies to convert 25 percent of their research and development spending into tax credits (Wang 2023). The South Korean National Assembly similarly agreed to legislation in March 2023 known as the "K-Chips Act," designed to boost the domestic semiconductor industry by expanding investment tax credits available to manufacturers like Samsung and SK Hynix (Kim 2023).

With industrial policies in play across all of these major industrialized economies, the ultimate global footprint of the industry remains highly uncertain. As one example, what happens if the hundreds of billions of industry and government dollars invested in new semiconductor facilities leads to excess global capacity? When supply of semiconductors exceeded demand in the 1980s and 1990s, the industry was known for infighting and turning to trade remedies such as antidumping and countervailing duties that sometimes ended up further limiting competition. In addition, no one is coordinating their semiconductor industrial policies with China.

Prospects for China's Semiconductor Industry

On one hand, China's technological catch-up in semiconductors is undeniably more difficult in the face of new export controls imposed by the United States, Japan, and the Netherlands. There is a plausible scenario in which Chinese firms fail to grow much from their present scale. Yet, it is also possible that the technological landscape evolves to the strengths of Chinese firms. If Chinese companies are

blocked from purchasing the equipment to make high-end chips, but it turns out that there is relatively less demand for high-end chips, which go into smartphones and data servers, and more demand for low-end chips, which go into electric vehicles and consumer electronics, then the Chinese industry with its advantages in low-end chips may be able to outcompete incumbents on volume (Wang 2023b).

Furthermore, just as the US-Japan trade pact to "stabilize" the chip market in the 1980s led to opportunities for Taiwan and South Korea to enter the market, the US-led export controls begun in October 2022 may also have unintended consequences. Many Chinese companies that had previously bought from US firms now have an incentive to buy domestic chips instead (Wang 2021). Similarly, the US efforts to keep China several generations behind US technological capabilities has created undeniably higher obstacles for China's leading chipmakers; on the other hand, these firms are now being forced to work more intensively to break this bottleneck.

The future of China's policy support for its semiconductor industry is also not clear. While some press reports emphasize that Beijing is prepared to spend more than ever on semiconductor subsidies, others suggest that Beijing is pausing chip investments, given their enormous cost and the country's economic problems elsewhere (compare reporting in Zhu 2022; *Bloomberg* News 2023). Nevertheless, many recent Chinese policies have been a retaliatory response to new foreign actions.

As one example, the Cybersecurity Administration of China announced in May 2023 that Micron had failed a security review, and barred Chinese companies involved in key infrastructure projects from buying from the US memory chipmaker (*Reuters* 2023b). The implied preferential access provided to Samsung and SK Hynix also works to drive a wedge between the interest of South Korean and US policymakers. In another example, following a June 2023 announcement by the Netherlands of export controls on chipmaking equipment, China retaliated with new export restrictions on gallium and germanium—materials critical to semiconductor manufacturers everywhere (Liu and Bradshaw 2023). According to the US Geological Survey (2023a, b), China was the source of 98 percent of global gallium production in 2022 and the source of 54 percent of US germanium metal imports over 2018–2021.

Finally, China blocked Intel's takeover of Tower Semiconductor, an Israel-headquartered company, by refusing to act on the proposed acquisition by August 2023 (Clark and Bradsher 2023). As Intel pivots to becoming more of a contract manufacturer—to compete with the likes of TSMC—it has attempted to acquire other foundries (Yu and Cheng 2022), and so China's denial of the Tower acquisition puts an obstacle in the way of this strategy.

Research Opportunities in Semiconductors and Industrial Policy

Although the modern semiconductor industry has extraordinary prominence in economics, politics, and foreign policy, it has hardly been studied by economists

(for an exception, see Thurk 2022). Admittedly, empirical research into the semiconductor industry faces data constraints. There are some data available from national statistical agencies, some from companies themselves in annual reports, and some from industry sources like the Semiconductor Industry Association, SEMI, and the Global Semiconductor Alliance, as well as consulting and market intelligence firms. Furthermore, there is also a lack of information on policy actions. Industrial policy deployed by the likes of the United States, Europe, Japan, South Korea, and Taiwan is relatively transparent, though understanding its effect must now account for not only direct subsidies, but also the near simultaneous imposition of export controls, import tariffs, foreign investment screening, and sometimes antitrust actions (Evenett et al. 2024). The data for Chinese industrial policy with regard to semiconductors suffer from all of those challenges and more, due to even more opaque features of China's underlying economic and political system, which may require novel research approaches to "back out" the size and impact of its policies.[11]

Both the modern story of the chips sector as well as current industrial policies are substantially different than their predecessors. As the latest wave of industrial policies with regard to semiconductors starts to take effect, it will raise some familiar questions but also some new ones.[12]

For example, the modern semiconductor industry has been reorganized. Chip firms mostly did everything in-house in the 1980s. The contemporary set-up features a long and highly-fragmented supply chain, with companies specializing in tasks, buying from some and selling to others, focusing on what they do best. Today's supply chains are global; thus, where firms locate geographically has also changed. Is it possible in this environment for policymakers to establish and sustain "national" champions? Are the agglomeration economies that attract policymakers—the chance of contributing to a new Silicon Valley—likely to work the same way in this new industrial structure?

Governments have been involved in semiconductor technologies since the beginning of the industry. However, their role has grown far more complex since the 1980s, when trade policy was the main tool used to tackle semiconductor issues. The potential role for governments to use industrial policy to address market imperfections such as agglomeration externalities, learning by doing, and technological leadership remains relevant. But today's officials have other motives that are not driven by economic efficiency: promoting geographic diversification, blunting China's ability to make technological gains, and limiting the spread of the most advanced chips (especially those potentially involved in national security issues). Given China's enormous internal market and capable scientific community, it

[11] See the approach taken in Barwick, Kalouptsidi, and Zahur (forthcoming) as applied to industrial policy for Chinese shipbuilding, as well as the accounting approach by OECD (2019) applied to semiconductors.

[12] Harrison and Rodríguez-Clare (2010) provide a classic introduction to the economics of industrial policy, albeit from the historical approach of market failures in developing countries. See also Grossman (1990). For recent surveys of the economics of industrial policy, see Juhász, Lane, and Rodrik (2024) and Bown (2024b).

remains to be seen how effective efforts to limit China's semiconductor technology will be. In response to US export controls, Chinese firms are investing heavily in the production of legacy chips, which might require a further US policy response (Wang 2023b). Also, diversification of supply chains for insurance purposes is likely to be costly, including if it results in plants operating (and supply chains clustering) at smaller scale or in more places that result in fewer agglomeration externalities.

The macroeconomic climate and market structure in which industrial policy happens is also likely to affect the outcomes. During the increasingly intense competition from imported semiconductors in the 1980s, segments of the US chip sector repeatedly suffered through jarring downturns. Yet, as industrial policy really got rolling in 2021 and 2022, global semiconductor sales were the highest ever recorded—during a period of growing demand associated in part with the economic peculiarities driven by the COVID-19 pandemic (SIA 2023b). Then came another positive demand shock driven by chips needed for artificial intelligence. There is now an enormous ongoing effort from companies and governments around the world to expand semiconductor fabrication capacity, but the details of future demand for semiconductors—in total, across sectors, and across types of chips—is highly uncertain. If the current investment expansion in semiconductor manufacturing leads to overcapacity and overproduction, at least in certain product segments of the market, will future governments view taxpayer support of financial losses in those areas as a price worth paying?

Economists should seek to evaluate the extent to which industrial policy is achieving its intended outcomes, in semiconductors and other industries, the costs of doing so, all the while remaining alert to the near-certainty of its unintended consequences.

■ *Thanks to Martin Chorzempa, Douglas Irwin, Ángel Ubide, Alan Wolff, and the journal's editors for valuable comments. Thanks to the Semiconductor Industry Association and Boston Consulting Group for sharing data. Jing Yan provided outstanding research assistance. Nia Kitchin, Alex Martin, and Sam Elbouez at the Peterson Institute assisted with graphics.*

References

Badohal, Karol, and Supantha Mukherjee. 2023. "Focus: How Poland Snagged Intel's Multi-billion Dollar Investment." *Reuters*, June 22. https://www.reuters.com/technology/how-poland-snagged-intels-multi-billion-dollar-investment-2023-06-22/.

Baker, Liana B., and Greg Roumeliotis. 2015. "Exclusive: Micron Does Not Believe Deal with Tsinghua Is Possible—Sources." *Reuters*, July 21. https://www.reuters.com/article/idUSKCN0PU1X1/.

Baldwin, Richard E., and Paul Krugman. 1988. "Market Access and International Competition: A

Simulation Study of 16K Random Access Memories." In *Empirical Methods for International Trade*, edited by Robert C. Feenstra. Cambridge, MA: MIT Press.

Barwick, Panle Jia, Myrto Kalouptsidi, and Nahim Zahur. Forthcoming. "Industrial Policy Implementation: Empirical Evidence from China's Shipbuilding Industry." *Review of Economic Studies*.

Beattie, Alan. 2022. "Stopping the Silicon Chips Falling Where They May: CEO of Global Semiconductor Research Hub IMEC Warns against Wasting Subsidies." *Financial Times*, April 4. https://www.ft.com/content/4feebf80-d9a9-4917-820c-5ec1ea83c436.

Bergsten, C. Fred, and Marcus Noland. 1993. *Reconcilable Differences? United States–Japan Economic Conflict*. Washington, DC: Institute for International Economics.

Bhagwati, Jagdish, and Hugh T. Patrick, eds. 1990. *Aggressive Unilateralism: America's 30I Trade Policy and the World Trading System*. New York and London: Harvester Wheatsheaf.

Biden, Joe. 2022. "Remarks by President Biden after Touring Samsung Electronics Pyeongtaek Campus." Speech, Samsung Electronics Pyeongtaek Campus, Pyeongtaek, May 20, 2022. https://www.whitehouse.gov/briefing-room/speeches-remarks/2022/05/20/remarks-by-president-biden-after-touring-samsung-electronics-pyeongtaek-campus/.

Biden, Joseph R., Jr. 2023. "Executive Order on Addressing United States Investments in Certain National Security Technologies and Products in Countries of Concern." White House, August 9. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/08/09/executive-order-on-addressing-united-states-investments-in-certain-national-security-technologies-and-products-in-countries-of-concern/.

BIS. 2022. "Implementation of Additional Export Controls: Certain Advanced Computing and Semiconductor Manufacturing Items; Supercomputer and Semiconductor End Use; Entity List Modification." Federal Register, October 13. https://www.federalregister.gov/documents/2022/10/13/2022-21658/implementation-of-additional-export-controls-certain-advanced-computing-and-semiconductor.

BIS. 2023. "Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use; Updates and Corrections." Federal Register, October 25. https://www.federalregister.gov/documents/2023/10/25/2023-23055/implementation-of-additional-export-controls-certain-advanced-computing-items-supercomputer-and.

Bloomberg News. 2023. "Battered by Covid, China Hits Pause on Giant Chip Spending Aimed at Rivaling US." *Bloomberg*, January 3. https://www.bloomberg.com/news/articles/2023-01-04/battered-by-covid-china-hits-pause-on-giant-chip-spending?sref=ATN0rNv3.

Bounds, Andy. 2023. "Made in Eindhoven: The Small Dutch City That Became a Tech Powerhouse." *Financial Times*, February 13. https://www.ft.com/content/c702c9f0-7f45-47d2-a0b3-6a6525e4b583.

Bown, Chad P. 2020. "How the United States Marched the Semiconductor Industry into Its Trade War with China." *East Asian Economic Review* 24 (4): 349—88.

Bown, Chad P. 2021a. "The US–China Trade War and Phase One Agreement." *Journal of Policy Modeling* 43 (4): 805–43.

Bown, Chad P. 2021b. "The Missing Chips: How to Protect the Semiconductor Supply Chain." *Foreign Affairs*, July 6. https://www.foreignaffairs.com/united-states/missing-chips.

Bown, Chad P. 2022a. "How COVID-19 Medical Supply Shortages Led to Extraordinary Trade and Industrial Policy." *Asian Economic Policy Review* 17 (1): 114–35.

Bown, Chad P. 2022b. "Four Years into the Trade War, Are the US and China Decoupling?" *PIIE Realtime Economics*, October 20. https://www.piie.com/blogs/realtime-economics/four-years-trade-war-are-us-and-china-decoupling.

Bown, Chad P. 2022c. "National Security, Semiconductors, and the US Move to Cut off China." *Trade Talks*, podcast, Episode 170, November 2. https://tradetalkspodcast.com/podcast/170-national-security-semiconductors-and-the-us-move-to-cut-off-china/.

Bown, Chad P. 2024a. "How the United States Solved South Korea's Problems with Electric Vehicle Subsidies under the Inflation Reduction Act." *Journal of World Trade* 58 (1): 1–34.

Bown, Chad P. 2024b. "Modern Industrial Policy and the World Trade Organization." *Annual Review of Economics*. https://doi.org/10.1146/annurev-economics-100223-041958.

Bown, Chad P., and Soumaya Keynes. 2020. "Why Trump Shot the Sheriffs: The End of WTO Dispute Settlement 1.0." *Journal of Policy Modeling* 42 (4): 799–819.

Bradshaw, Tim, and Anna Gross. 2023. "The Big Question of How Small Chips Can Get." *Financial Times*, May 23. https://www.ft.com/content/fbf52ede-e1a9-4797-9e52-3f80a7d855d9.

Brown, Clair, and Greg Linden. 2009. *Chips and Change: How Crisis Reshapes the Semiconductor Industry*.

Cambridge, MA: MIT Press.

Cao, Ann. 2023. "China Gave 190 Chip Firms US$1.75 Billion in Subsidies in 2022 as It Seeks Semiconductor Self-Sufficiency." *South China Morning Post*, May 7. https://www.scmp.com/tech/tech-war/article/3219697/china-gave-190-chip-firms-us175-billion-subsidies-2022-it-seeks-semiconductor-self-sufficiency.

Chang, Pao-Long, and Chiung-Wen Hsu. 1998. "The Development Strategies for Taiwan's Semiconductor Industry." *IEEE Transactions on Engineering Management* 45 (4): 349–56.

CHIPS for America. 2023a. "Frequently Asked Questions: Commercial Fabrication Facilities." Department of Commerce, National Institute of Standards and Technology. https://www.nist.gov/chips/frequently-asked-questions-commercial-fabrication-facilities (accessed December 27, 2023).

CHIPS for America. 2023b. "CHIPS for America's Strong Start: Department Received 200 Statements of Interest for Semiconductor Projects Across 35 States." Press release, April 14. https://www.nist.gov/news-events/news/2023/04/chips-americas-strong-start.

Clark, Don, and Keith Bradsher. 2023. "China Scuttles a $5.4 Billion Microchip Deal Led by U.S. Giant Intel." *New York Times*, August 16. https://www.nytimes.com/2023/08/16/business/intel-tower-semiconductor-china.html.

CRS. 2018. The Committee on Foreign Investment in the United States (CFIUS). CRS Report RL33388. Washington, DC: Congressional Research Service.

CRS. 2023. *Frequently Asked Questions: CHIPS Act of 2022 Provisions and Implementation*. CRS Report R47523. Washington, DC: Congressional Research Service.

Dick, Andrew R. 1991. "Learning by Doing and Dumping in the Semiconductor Industry." *Journal of Law and Economics* 34 (1): 133–59.

Dou, Eva. 2015. "Qualcomm Spared Some of Worst Penalties by Chinese Authorities." *Wall Street Journal*, February 10. https://www.wsj.com/articles/qualcomms-fine-signals-chinas-toughened-stance-1423532319.

Evenett, Simon, Adam Jakubik, Fernando Martín, and Michele Ruta. 2024. "The Return of Industrial Policy in Data." *World Economy* 47 (7): 2762–88.

Fitch, Asa. 2021. "Texas Winter Storm Strikes Chip Makers, Compounding Supply Woes." *Wall Street Journal*, February 17. https://www.wsj.com/articles/texas-winter-storm-strikes-chip-makers-compounding-supply-woes-11613588617.

Ford, Christopher A. 2019. "Technology and Power in China's Geopolitical Ambitions." Testimony to the U.S.–China Economic and Security Review Commission, Washington, DC, June 20, 2019. https://2017-2021.state.gov/technology-and-power-in-chinas-geopolitical-ambitions/.

Government of the Netherlands. 2023. "Government Publishes Additional Export Measures for Advanced Semiconductor Manufacturing Equipment." Press release, June 30. https://www.einpresswire.com/article/642343929/government-publishes-additional-export-measures-for-advanced-semiconductor-manufacturing-equipment.

Grossman, Gene M. 1990. "Promoting New Industrial Activities: A Survey of Recent Arguments and Evidence." OECD Economic Study 14.

Grossman, Matt. 2021. "GM Plans to Idle Factories Longer Amid Chip Shortage." *Wall Street Journal*, September 16. https://www.wsj.com/articles/gm-plans-to-idle-factories-longer-amid-chip-shortage-11631819780.

Guarascio, Francesco. 2023. "Vietnam Eyes First Semiconductor Plant, US Officials Warn of High Costs." *Reuters*, October 30. https://www.reuters.com/technology/vietnam-eyes-first-semiconductor-plant-us-officials-warn-high-costs-2023-10-31/.

Harrell, Peter, Elizabeth Rosenberg, and Edoardo Saravalle. 2018. *China's Use of Coercive Economic Measures*. Washington, DC: Center for a New American Security.

Harrison, Ann, and Andrés Rodríguez-Clare. 2010. "Trade, Foreign Investment, and Industrial Policy for Developing Countries." In *Handbook of Development Economics*, Vol. 5, edited by Dani Rodrik and Mark Rosenzweig, 4039–4214. Amsterdam: Elsevier.

Hawkins, Mackenzie, and Jenny Leonard. 2023. "US to Gather Intelligence on Chinese Chipmakers as Biden Mulls Tariffs." *Bloomberg*, December 21. https://www.bloomberg.com/news/articles/2023-12-21/us-to-gather-intelligence-on-chinese-chipmakers-as-biden-mulls-tariffs?sref=ATN0rNv3.

Hayashi, Yuka. 2022. "US, EU Agree to Coordinate Semiconductor Subsidy Programs." *Wall Street Journal*, December 5. https://www.wsj.com/articles/u-s-eu-agree-to-coordinate-semiconductor-subsidy-programs-11670284917.

Hayashi, Yuka, and Asa Fitch. 2023. "U.S. Tightens Curbs on AI Chip Exports to China, Widening Rift

with Nvidia and Intel." *Wall Street Journal*, October 17. https://www.wsj.com/tech/u-s-tightens-curbs-on-ai-chip-exports-to-china-widening-rift-with-u-s-businesses-3b9983df.

Heine, Friederike, Supantha Mukherjee, and Andreas Rinke. 2023. "Intel Spends $33 Billion in Germany in Landmark Expansion." *Reuters*, June 19. https://www.reuters.com/technology/berlin-sign-agreement-with-intel-after-chip-plant-talks-2023-06-19/.

Horwitz, Josh. 2021. "Fridges, Microwaves Fall Prey to Global Chip Shortage." *Reuters*, March 29. https://www.reuters.com/world/china/fridges-microwaves-fall-prey-global-chip-shortage-2021-03-29/.

Hoshi, Takeo, Anil Kashyap, and David Scharfstein. 1990. "The Role of Banks in Reducing the Costs of Financial Distress in Japan." *Journal of Financial Economics* 27 (1): 67–88.

IC Insights Research Bulletin. 2012. "Extreme Results in Top 25 2011 Semiconductor Sales Ranking." April 4. https://www.icinsights.com/data/articles/documents/424.pdf.

IC Insights Research Bulletin. 2021. "17 Semiconductor Companies Forecast to Have >$10.0 Billion in Sales This Year." December 20. https://www.icinsights.com/news/bulletins/17-semiconductor-companies-forecast-to-have-100-billion-in-sales-this-year/.

Inagaki, Kana. 2023. "How TSMC's Chip Plant Is Shaking up Japan." *Financial Times*, September 25. https://www.ft.com/content/09f0ae79-8935-4070-ab52-dc828b770dce.

Intel. 2022. "Intel Announces Next US Site with Landmark Investment in Ohio." Press release, January 21. https://www.intel.com/content/www/us/en/newsroom/news/intel-announces-next-us-site-landmark-investment-ohio.html#gs.d0pzbt.

Intel. 2023. "Intel Plans Assembly and Test Facility in Poland." Press release, June 16. https://www.intc.com/news-events/press-releases/detail/1627/intel-plans-assembly-and-test-facility-in-poland.

Irwin, Douglas A. 1996. "Trade Policies and the Semiconductor Industry." In *The Political Economy of American Trade Policy*, edited by Anne O. Krueger, 11–72. Chicago: University of Chicago Press.

Irwin, Douglas A., and Peter J. Klenow. 1994. "Learning-by-Doing Spillovers in the Semiconductor Industry." *Journal of Political Economy* 102 (6): 1200–1227.

Irwin, Douglas A., and Peter J. Klenow. 1996. "High-Tech R&D Subsidies Estimating the Effects of Sematech." *Journal of International Economics* 40 (3–4): 323–44.

Juhász, Réka, Nathan Lane, and Dani Rodrik. 2024. "The New Economics of Industrial Policy." *Annual Review of Economics*. https://doi.org/10.1146/annurev-economics-081023-024638.

Jung-a, Song, and Eleanor Olcott. 2021. "Global Chip Shortage Spreads to Toasters and Washing Machines." *Financial Times*, April 24. https://www.ft.com/content/9a497f21-e3e4-4c74-8737-07cd13209475.

Kelly, Tim, and Jane Lee. 2022. "IBM Partners with Japan's Rapidus in Bid to Manufacture Advanced Chips." *Reuters*, December 12. https://www.reuters.com/technology/ibm-partners-with-new-japanese-chip-maker-rapidus-make-advanced-chips-2022-12-13/.

Kelly, Tim, and Miho Uranaka. 2023. "Japan Restricts Chipmaking Equipment Exports as It Aligns with US China Curbs." *Reuters*, March 31. https://www.reuters.com/technology/japan-restrict-chipmaking-equipment-exports-aligning-it-with-us-china-curbs-2023-03-31/.

Khan, Saif M., Alexander Mann, and Dahlia Peterson. 2021. "The Semiconductor Supply Chain: Assessing National Competitiveness." Center for Security and Emerging Technology (CSET) Issue Brief, January. https://cset.georgetown.edu/wp-content/uploads/The-Semiconductor-Supply-Chain-Issue-Brief.pdf.

Kim, Alex. 2023. "CHIP on the Shoulder." Wilson Center (blog), October 10. https://www.wilsoncenter.org/blog-post/chip-shoulder.

Kim, S. Ran. 1998. "The Korean System of Innovation and the Semiconductor Industry: A Governance Perspective." *Industrial and Corporate Change* 7 (2): 275–309.

Landler, Mark. 2020. "The Silicon Godfather; The Man Behind Taiwan's Rise in the Chip Industry." *New York Times*, February 1. https://www.nytimes.com/200%2/01/business/the-silicon-godfather-the-man-behind-taiwan-s-rise-in-the-chip-industry.html.

Liu, Quianer, and Tim Bradshaw. 2023. "China Imposes Export Curbs on Chipmaking Metals." *Financial Times*, July 3. https://www.ft.com/content/6dca353c-70d8-4d38-a368-b342a6450d95.

Liu, Quianer, and Cheng Leng. 2023. "China's Economic Malaise Hits Efforts to Raise $41BN Chip Fund." *Financial Times*, September 26. https://www.ft.com/content/521c8ac3-1933-4077-88b9-e9086a0196ca.

Lu, Yiwen. 2023. "Intel Moves Forward With $25 Billion Chip Plant Expansion in Israel." *New York Times*, December 27. https://www.nytimes.com/2023/12/27/world/middleeast/intel-israel-factory-expansion.html.

Martina, Michael, and Stephen Nellis. 2018. "Qualcomm Ends $44 Billion NXP Bid after Failing to Win China Approval." *Reuters*, July 26. https://www.reuters.com/article/idUSKBN1KF18X/.

Micron. 2022. "Micron Announces Historic Investment of up to $100 Billion to Build Megafab in Central New York." Press release, October 4. https://investors.micron.com/news-releases/news-release-details/micron-announces-historic-investment-100-billion-build-megafab.

Miller, Chris. 2022. *Chip War: The Fight for the World's Most Critical Technology*. New York: Scribner.

Miller, Joe, and Martin Arnold. 2021. "Car Chip Shortage Weighs on German Economy." *Financial Times*, July 7. https://www.ft.com/content/3d7d4ad2-63ef-45f0-adee-a4b7abe3d027.

Mundie, Craig, and Paul Otellini. 2017. "Ensuring US Leadership and Innovation in Semiconductors." White House (blog), January 9. https://obamawhitehouse.archives.gov/blog/2017/01/09/ensuring-us-leadership-and-innovation-semiconductors.

Nagao, Rihao. 2023. "Japan and EU to Share Chip Subsidy Info to Disperse Production. Three-Way Exchange with U.S. Aims for Better Supply Chain Distribution." Nikkei Asia, June 29. https://asia.nikkei.com/Business/Tech/Semiconductors/Japan-and-EU-to-share-chip-subsidy-info-to-disperse-production.

Nellis, Stephen, and Chavi Mehta. 2023. "Qualcomm Beats Estimates on China Strength, Shares Rise." *Reuters*, November 1. https://www.reuters.com/technology/qualcomm-forecasts-revenue-profits-above-estimates-2023-11-01/.

National Institute of Standards and Technology (NIST). 2023a. *Notice of Funding Opportunity (NOFO): CHIPS Incentives Program—Commercial Fabrication Facilities*. Washington, DC: Department of Commerce.

National Institute of Standards and Technology. 2023b. "Preventing the Improper Use of CHIPS Act Funding." Proposed rule, Federal Register, March 23. https://www.federalregister.gov/documents/2023/03/23/2023-05869/preventing-the-improper-use-of-chips-act-funding.

National Institute of Standards and Technology. 2023c. "Preventing the Improper Use of CHIPS Act Funding." Rule, Federal Register, September 25. https://www.federalregister.gov/documents/2023/09/25/2023-20471/preventing-the-improper-use-of-chips-act-funding.

Nohara, Yoshiaki. 2023. "In Boost for Chip Ambitions, Japan Inks $1.3 Billion in Subsidies for Micron Plant." *Bloomberg*, October 2. https://www.bloomberg.com/news/articles/2023-10-03/japan-inks-1-3-billion-in-subsidies-for-micron-hiroshima-plant?sref=ATN0rNv3.

Norton Rose Fulbright. 2023. "EU Chips Act: Risks and Opportunities for Businesses." September. https://www.nortonrosefulbright.com/en/knowledge/publications/f176610c/eu-chips-act-risks-and-opportunities-for-businesses.

OECD. 1992. *Globalisation of Industrial Activities, Four Case Studies: Auto Parts, Chemicals, Construction, and Semiconductors*. Paris: OECD.

OECD. 2019. "Measuring Distortions in International Markets: The Semiconductor Value Chain." OECD Trade Policy Paper 234.

O'Keeffe, Kate, and Asa Fitch. 2023. "Blacklisted Chinese Chip Maker Does a Thriving Business with U.S." *Wall Street Journal*, October 5. https://www.wsj.com/politics/national-security/blacklisted-chinese-chip-maker-does-a-thriving-business-with-u-s-f75cca55.

Okuno-Fujiwara, Masahiro. 1991. "Industrial Policy in Japan: A Political Economy View." In *Trade with Japan: Has the Door Opened Wider?*, edited by Paul Krugman, 271–304. Chicago: University of Chicago Press.

Pasquini, Linda. 2023. "Germany Approves Stakes by Bosch, Infineon and NXP in TSMC Chip Plant," *Reuters*, November 7. https://www.reuters.com/technology/german-regulator-approves-stakes-by-bosch-infineon-nxp-tsmc-chip-plant-2023-11-07/.

Prestowitz, Clyde V. 1988. *Trading Places: How We Allowed Japan to Take the Lead*. New York: Basic Books.

Reuters. 2021. "TSMC to Invest $100 Billion over 3 Years to Meet Chip Demand." April 1. https://www.reuters.com/article/idUSKBN2BO3ZJ/.

Reuters. 2023a. "Intel to Invest $1.2 bln in Costa Rica over Next Two Years." August 30. https://www.reuters.com/business/intel-invest-12-bln-costa-rica-over-next-two-years-2023-08-30/.

Reuters. 2023b. "Micron Warns of Bigger Revenue Hit from China Ban." June 16. https://www.reuters.com/technology/micron-says-half-china-headquarter-revenue-risk-due-ban-2023-06-16/.

Roser, Max, Hannah Ritchie, and Edouard Mathieu. 2023. "What is Moore's Law?" Our World in Data, March 28. https://ourworldindata.org/moores-law.

Samsung. 2021. "Samsung Electronics Announces New Advanced Semiconductor Fab Site in Taylor Texas." Press release, November 23. https://semiconductor.samsung.com/us/sas/local-news/

samsung-electronics-announces-new-advanced-semiconductor-fab-site-in-taylor-texas/.

Schuman, Michael. 2022. "Why Biden's Block on Chips to China Is a Big Deal." *The Atlantic*, October 25. https://www.theatlantic.com/international/archive/2022/10/biden-export-control-microchips-china/671848/.

Sheahan, Maria. 2016. "China's Fujian Drops Aixtron Bid after Obama Blocks Deal." *Reuters*, December 8. https://www.reuters.com/article/idUSKBN13X196/.

Shepardson, David. 2022. "U.S. Picks Team to Oversee $52.7 Billion in Semiconductor Funding." *Reuters*, September 20. https://www.reuters.com/technology/us-picks-team-oversee-527-bln-semiconductor-funding-2022-09-20/.

Shepardson, David. 2023. "U.S. Commerce Dept. Names Financial Sector, Government Experts to Chips Team." *Reuters*, February 17. https://www.reuters.com/technology/us-commerce-names-financial-sector-government-experts-chips-team-2023-02-17/.

SIA. 2023a. 2023 Factbook. Washington, DC: Semiconductor Industry Association.

SIA. 2023b. "Global Semiconductor Sales Increase 3.3% in 2022 Despite Second-Half Slowdown." Press release, February 3. https://www.semiconductors.org/global-semiconductor-sales-increase-3-2-in-2022-despite-second-half-slowdown/.

State Council. 2020. "Notice on Promoting the Integrated Circuit Industry and Development in the New Era." July 27. https://www.gov.cn/zhengce/content/2020-08/04/content_5532370.htm.

Swanson, Ana. 2023. "The CHIPS Act Is About More Than Chips: Here's What's in It." *New York Times*, February 28. https://www.nytimes.com/2023/02/28/business/economy/chips-act-childcare.html.

Sullivan, Jake. 2023. "Remarks by National Security Advisor Jake Sullivan on Renewing American Economic Leadership at the Brookings Institution." Speech, Brookings Institution, Washington, DC, April 27. https://www.whitehouse.gov/briefing-room/speeches-remarks/2023/04/27/remarks-by-national-security-advisor-jake-sullivan-on-renewing-american-economic-leadership-at-the-brookings-institution/.

Tilton, John E. 1971. *International Diffusion of Technology: The Case of Semiconductors*. Washington, DC: Brookings Institution.

Thurk, Jeff. 2022. "Outsourcing, Firm Innovation, and Industry Dynamics in the Production of Semiconductors." Unpublished.

Thurow, Lester C. 1992. *Head to Head: The Coming Economic Battle among Japan, Europe, and America*. New York: Warner Books.

TSMC. 2020. "TSMC Announces Intention to Build and Operate an Advanced Semiconductor Fab in the United States." Press release, May 15. https://pr.tsmc.com/english/news/2033.

US Department of Commerce. 2023. "Biden–Harris Administration and BAE Systems, Inc., Announce CHIPS Preliminary Terms to Support Critical U.S. National Security Project in Nashua, New Hampshire." Press release, December 11. https://www.commerce.gov/news/press-releases/2023/12/biden-harris-administration-and-bae-systems-inc-announce-chips.

US Department of Commerce. 2024. "Biden–Harris Administration Announces CHIPS Preliminary Terms with Microchip Technology to Strengthen Supply Chain Resilience for America's Automotive, Defense, and Aerospace Industries." Press release, January 4. https://www.commerce.gov/news/press-releases/2024/01/biden-harris-administration-announces-chips-preliminary-terms-microchip.

US Department of Justice. 2005. "Samsung Agrees to Plead Guilty and to Pay $300 Million Criminal Fine for Role in Price Fixing Conspiracy." Press release, October 13. https://www.justice.gov/archive/atr/public/press_releases/2005/212002.htm.

US Department of Justice. 2019. "Chinese Telecommunications Device Manufacturer and its U.S. Affiliate Indicted for Theft of Trade Secrets, Wire Fraud, and Obstruction of Justice." Press release, January 28. https://www.justice.gov/opa/pr/chinese-telecommunications-device-manufacturer-and-its-us-affiliate-indicted-theft-trade.

US Department of State. 2023a. "Department of State Announces Plans to Implement the CHIPS Act International Technology Security and Innovation Fund." Press release, March 14. https://www.state.gov/department-of-state-announces-plans-to-implement-the-chips-act-international-technology-security-and-innovation-fund/.

US Department of State. 2023b. "New Partnership with Costa Rica to Explore Semiconductor Supply Chain Opportunities." Press release, July 14. https://www.state.gov/new-partnership-with-costa-rica-to-explore-semiconductor-supply-chain-opportunities/.

US Department of State. 2023c. "New Partnership with Panama to Explore Semiconductor Supply Chain Opportunities." Press release, July 20. https://www.state.gov/new-partnership-with-panama-to-explore-semiconductor-supply-chain-opportunities/.

US Department of State. 2023d. "New Partnership with Vietnam to Explore Semiconductor Supply Chain Opportunities." Press release, September 11. https://www.state.gov/new-partnership-with-vietnam-to-explore-semiconductor-supply-chain-opportunities/.

US Geological Survey. 2023a. "Mineral Commodity Summaries: Gallium." https://pubs.usgs.gov/periodicals/mcs2023/mcs2023-gallium.pdf.

US Geological Survey. 2023b. "Mineral Commodity Summaries: Germanium." https://pubs.usgs.gov/periodicals/mcs2023/mcs2023-germanium.pdf.

US International Trade Commission (USITC). 1993. *Industry and Trade Summary: Semiconductors. USITC Publication 2708*. Washington, DC: US International Trade Commission.

US Trade Representative (USTR). 2018. *Findings of the Investigation into China's Acts, Policies, and Practices Related to Technology Transfer, Intellectual Property, and Innovation under Section 301 of the Trade Act of 1974*. Washington, DC: Office of the United States Trade Representative.

Varas, Antonio, and Raj Varadarajan. 2020. *How Restrictions to Trade with China Could End US Leadership in Semiconductors*. Boston: Boston Consulting Group.

Varas, Antonio, Raj Varadarajan, Jimmy Goodrich, and Falan Yinug. 2020. *Government Incentives and US Competitiveness in Semiconductor Manufacturing*. Boston and Washington, DC: Boston Consulting Group and Semiconductor Industry Association.

Varas, Antonio, Raj Varadarajan, Jimmy Goodrich, and Falan Yinug. 2021. *Strengthening the Global Semiconductor Supply Chain in an Uncertain Era*. Boston and Washington, DC: Boston Consulting Group and Semiconductor Industry Association.

Vidalon, Dominique, and Sudip Kar-Gupta. 2023. "France to Provide 2.9 Billion Euros in Aid for New STMicro/GlobalFoundries Factory." *Reuters*, June 5. https://www.reuters.com/markets/europe/france-provide-29-bln-euros-aid-new-stmicroglobalfoundries-factory-2023-06-05/.

Wang, Dan. 2021. "China's Sputnik Moment?" *Foreign Affairs*, July 29. https://www.foreignaffairs.com/united-states/chinas-sputnik-moment.

Wang, Dan. 2023a. "China's Hidden Tech Revolution: How Beijing Threatens U.S. Dominance." *Foreign Affairs* 102 (2): 65–77.

Wang, Dan. 2023b. "Biden Is Beating China on Chips. It May Not Be Enough." *New York Times*, July 16. https://www.nytimes.com/2023/07/16/opinion/biden-china-ai-chips-trade.html.

Wang, Lisa. 2023. "Taiwan 'Chips Act' Sets R&D Spending at NT$6 Billion." *Taipei Times*, May 2. https://www.taipeitimes.com/News/front/archives/2023/05/02/2003798964.

Waters, Richard. 2023. "Nvidia Expects Revenue Boom as AI Drives Chip Demand." *Financial Times*, May 24. https://www.ft.com/content/6b03dee3-ea79-48dd-b1eb-3df4f029d806.

Wu, Debby, and Aggi Cantrill. 2023. "TSMC to Build $11 Billion German Plant with Other Chipmakers." *Bloomberg*, August 8. https://www.bloomberg.com/news/articles/2023-08-08/tsmc-partners-with-infineon-nxp-bosch-to-build-german-chip-fab?sref=ATN0rNv3.

Yu, Yifan, and Ting-Fang Cheng. 2022. "How Intel Plans to Rival TSMC and Samsung as a Chip Supplier." *Nikkei Asia*, November 4. https://asia.nikkei.com/Business/Business-Spotlight/How-Intel-plans-to-rival-TSMC-and-Samsung-as-a-chip-supplier.

Zhu, Julie. 2022. "Exclusive: China Readying $143 Billion Package for Its Chip Firms in Face of U.S. Curbs." *Reuters*, December 13. https://www.reuters.com/technology/china-plans-over-143-bln-push-boost-domestic-chips-compete-with-us-sources-2022-12-13/.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 111–130

Alexander Hamilton's *Report on Manufactures* and Industrial Policy

Richard Sylla

Current proponents of industrial policy in the United States sometimes refer to Alexander Hamilton's December 1791 *Report on the Subject of Manufactures* (hereafter *Manufactures*) as an early endorsement of industrial policy, and a precedent for the sorts of industrial policies they would like to see implemented in our own time. After all, if Hamilton—a leading founder of the United States, a framer of its Constitution, its first Secretary of the Treasury, and the hero of a celebrated Broadway musical—was for industrial policy, how could anyone be against it? In Hamilton's time, of course, many notable thinkers and policy-makers were against industrial policy, just as many are today.

Hamilton did not espouse state-directed economic development, contrary to the views sometimes attributed to him positively (Cohen and DeLong 2016; Parenti 2020) or negatively (DiLorenzo 2009; Hogeland 2024). He favored tariffs as the most practical way of raising government revenue in the 1790s. But he opposed raising those tariffs to truly protective levels, and his Federalist political party suffered in popularity as a result. Hamilton designed his policies to create a sound system of banking and finance for the capital needs of the government and American entrepreneurs; to use infrastructure, innovation, and technology diffusion to speed up economic growth and diversification; and to support industries crucial to

■ *Richard Sylla is Professor Emeritus of Economics, Stern School of Business, New York University, New York City, New York. He is also a Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. His email address is rs5@stern.nyu.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.111.

US national defense in a world dominated by marauding European empires. He was writing at a time when export-driven growth seemed impossible, because of barriers to trade imposed by European empires. His focus, therefore, was on how a mostly agrarian society could benefit from increased domestic demand for its products through a growing manufacturing sector. To those ends, and in those ways, he wanted government to develop a broader economic context that was favorable for the future growth of manufacturing.

Despite the historical significance of *Manufactures* and its possible relevance to current discussions of industrial policy, economists have paid little attention to it. Unless they have an interest in economic history, the history of economic thought, or US history, there perhaps is little reason they should. Those who have studied it are laudatory. Schumpeter (1954, p. 199) says that *Manufactures* ". . . is really 'applied economics' at its best. . . ."[1] More recently, Irwin (2017, p. 80) writes, "This brilliant report ranks among the most important and influential policy documents in US history." McCraw (1994, p. 32), a business historian, considering *Manufactures* in the context of Hamilton's comprehensive program for US economic modernization, concludes, "Hamilton was the first systematic macroeconomic planner in the United States and one of the first in any country."

My goal here is to make Hamilton's *Manufactures* more familiar to economists in general, and especially to those involved in debates over industrial policy. I begin by clarifying the question that Hamilton was addressing at the time: Should the United States even consider becoming a manufacturing nation? I then place *Manufactures* in the context of the comprehensive program of financial and economic modernization that Hamilton advocated, and Congress largely enacted, after he became the first US Secretary of the Treasury in 1789. Next, I survey the content and arguments of *Manufactures*. I point out how policymakers implemented Hamilton's recommendations, initially and over time. Finally, I take up the practical effects of the policies in the context of nineteenth-century US economic growth and development.

Did the US Economy Even Need Manufacturing?

Debates about industrial policy in the 2020s differ greatly from those of the 1790s.[2] Few people today question the legitimacy of modern manufacturing technologies. For more than two centuries, such technologies have contributed mightily to unprecedented economic growth and development. Some term this "the Great

[1] Schumpeter (1954, p. 199) adds in a footnote that Hamilton "was one of those rare practitioners of economic policy who think it worthwhile to acquire more analytic economics than that smattering that does such good service in addressing audiences of a certain type. He knew Smithian economics well—not only Adam Smith himself—so well in fact as to be able to mold it to his own visions of practical possibilities or necessities and to perceive its limitations. All his reports . . . are much more than untutored common sense."

[2] For a nearly encyclopedic account of the pros and cons of industrial policy, including historical country case studies of where it worked and where it did not, see Fasteau and Fletcher (2024).

Enrichment," resulting in real per capita incomes of today's developed economics that are 30 or more times greater than two centuries ago (McCloskey and Carden 2020).

Much of the industrial-policy debate today occurs in mature industrial or even postindustrial economies. One issue is whether governments should intervene to alter the course that established manufacturing industries and other economic sectors might otherwise take. Proponents argue that such interventions will improve overall economic efficiency, and achieve other, perhaps noneconomic, objectives. Opponents deny that governments have any special capability to do this, and claim that attempts to do it are likely to degenerate into corruption and crony capitalism.

In the United States of the 1790s, however, many people—including some leading founders—questioned whether the country should even consider becoming a manufacturing nation. Such opponents of manufacturing could point to the new nation's abundant and mostly unsettled land resources, its high wages, and its lack of capital to justify a mostly agricultural future. The only exception that opponents of American manufacturing might allow involved considerations of national security. Congress, following up on a request from President George Washington in January 1790, asked Hamilton to report on the subject of manufactures, "and particularly to the means of promoting such as will tend to render the United States, independent on foreign nations, for military and other essential supplies" (Hamilton 1791, p. 230). Fresh in the minds of national leaders were the difficulties of equipping American military forces during the War of Independence.

Opponents of American manufacturing could point to Adam Smith's (1776) recently published *Wealth of Nations*, which a number of American leaders had studied, for justification of their position. Smith (1776, pp. 347–48) pointedly argued, for example:

> Were the Americans, either by combination or by any other sort of violence, to stop the importation of European manufactures, and, by thus giving a monopoly of such of their own countrymen as could manufacture the like goods, divert any considerable part of their capital into this employment, they would retard instead of accelerating the further increase in the value of their annual produce, and would obstruct instead of promoting the progress of their country towards real wealth and greatness.

Smith's mention of "combination" and "violence" no doubt referred to American boycotts of British imports in the revolutionary ferment of the early 1770s, when he was writing. But he likely was also thinking about what economists would later term "comparative advantage." The Americans had a lot of land and were short of both labor and capital. With such a combination of resources, manufacturing seemingly made little sense. At other places in *Wealth of Nations*, Smith agreed with the French physiocrat school of economics that agriculture was the most productive economic activity, although he disagreed with them that it was the only productive activity (Smith 1776, pp. 344–45). In any case, Smith's advice to Americans was to stick to agriculture.

A few years later, Thomas Jefferson (1785, pp. 252–53) agreed with Adam Smith that Americans should stick to farming:

> In Europe the lands are either cultivated, or locked up against the cultivator. Manufacture must therefore be resorted to of necessity not of choice, to support the surplus of their people. But we have an immensity of land courting the industry of the husbandman. Is it best then that all our citizens should be employed in its improvement, or that one half should be called off from that to exercise manufactures, and handicraft arts for the other? Those who labour in the earth are the chosen people of god. . . . [G]enerally speaking, the proportion which the aggregate of the other classes of citizens bears in any state to that of its husbandmen, is the proportion of the unsound to its healthy parts. . . . While we have land to labour then, let us never wish to see our citizens occupied at a work-bench, or twirling a distaff. Carpenters, masons, smiths are wanting in husbandry: but, for the general operations of manufacture, let our work-shops remain in Europe.

Facing such authoritative opposition to American manufacturing on both sides of the Atlantic, Hamilton had his work cut out for him. He therefore began *Manufactures* with an extended discussion of why the United States should become a manufacturing nation with a diversified agricultural/manufacturing/commercial economy. He agrees with some of Adam Smith's teachings and disagrees with others. Indeed, *Manufactures* is one of the first extended commentaries on the *Wealth of Nations*, although Hamilton made it far from easy for his readers to see this. He paraphrases Smith without attribution. At one place in *Manufactures*, he directly quotes Smith without naming him. He does mention Smith in the marginalia of the third draft of *Manufactures*, but the text refers to him only as "a judicious writer." Hamilton wrote to persuade Congress, not to make it easy for historians of economic thought to trace the origins of his ideas by including footnotes to his sources.[3]

Hamilton, ahead of his time, saw modern manufacturing technologies as the wave of the future, although they barely existed in the United States—or anywhere except England, and even there only recently. His objective was to convince Congress that going down that road was in the national interest, and to speed up the trip by enacting policy measures to "encourage" (his oft-used term) the development of modern manufacturing.

Given the different contexts of the 1790s and the 2020s, it might be more accurate to view Hamilton's *Manufactures* as less about "industrial policy" as that term is currently understood, and more about "industrialization policy" for an

[3]Those interested in Hamilton's possible sources of information used in writing *Manufactures* should read the extensively introduced and footnoted version in *The Papers of Alexander Hamilton* (Hamilton 1791, pp. 236–340). It is the one cited here. Fortunately, that version is now easily accessible at the website "Founders Online" sponsored by the Library of Congress.

undeveloped, overwhelmingly agricultural economy in a world economy with little modern manufacturing and virtually no prospects for free international trade. In late eighteenth century Great Britain, what much later would be termed "the Industrial Revolution" snuck up unannounced. By contrast, in the United States, Hamilton and a few others foresaw the Industrial Revolution and espoused public policies to accelerate it.

Manufactures in the Context of Hamilton's Plan for US Economic Modernization

By the time Hamilton delivered *Manufactures* to Congress in December 1791, he already had three notable policy successes: establishing a market for federal debt, a national bank with branches around the country, and a bimetallic currency standard. He refers to these steps in *Manufactures* as already working, to allay fears that the United States was ill-prepared for modern manufacturing because, most notably, it lacked the capital that would be required. Hence, *Manufactures* is in part an early report on how Hamilton saw his plans for modernization of the US economy progressing.

Hamilton's first success came in the summer of 1790 when Congress, after much debate and deal-making, enacted the essence of his plan for establishing the credit of the new federal government. The backdrop was the accumulation of national and state debts that resulted from the War of Independence. These debts had largely been unpaid, accumulating arrears of interest, because national and state governments lacked the revenues or the will, or both, to pay them. Hamilton's plan called for federal assumption of state debts, and restructuring the enlarged national debt via voluntary exchanges of old debt at par, but at a reduced rate of interest, for new debt securities. The new debt was "funded," meaning that specific government revenues were pledged to pay the interest and, when possible, the principal. Exchanges of old debt for new began in late 1790. The new federal securities rose rapidly in value, and they were actively traded in organized markets that arose in major cities. This was the birth of US Treasury bond market.

Hamilton's second notable policy success came in February 1791 when Congress enacted his plan for a national bank, the Bank of the United States. Washington's cabinet divided on the issue of the Bank's constitutionality; the president sided with Hamilton's broad construction of constitutional powers. In the corporate charter of the Bank of the United States, written by Hamilton, private stockholders would own 80 percent of the corporation's shares and elect its management; the federal government took a 20 percent stake and had oversight. The Bank would serve as the government's fiscal agent, assist it with managing the national debt, and lend it money. The Bank of the United States could (and did) open branches around the country, and engage in ordinary commercial banking by taking deposits and making loans. The bank's charter served as a model for states to emulate as they proceeded to charter more banks and other corporations (Bodenhorn 2010; Sylla,

Legler, and Wallis 1987). The home office of the Bank of the United States opened in December 1791 in Philadelphia, then the national capital, the same month Hamilton presented *Manufactures* to Congress.

A third policy success came in 1792, after *Manufactures*, when Congress established a US mint along the lines set out in Hamilton's Mint Report of January 1791. In that report, Hamilton defined the new US dollar in terms of silver and gold, establishing a bimetallic monetary base for the country. The report had no immediate bearing on his policies for encouraging manufacturing. In the longer run, however, a common currency area among US states facilitated interstate trade for manufactures and other goods and services.

A Tour of *Manufactures*

Hamilton begins *Manufactures* by making a case for why the United States needed to encourage manufacturing. International trade was far from free in the 1790s. The leading foreign nations—Britain, France, and Spain—each had a population several times larger than that of the United States, which then was about four million. Moreover, each controlled a vast overseas empire and had mercantilist trade policies that in practice restricted or excluded the newly independent Americans from trading the surplus commodities of US farms, forests, and fisheries with either the home country or its colonies. Therefore, Hamilton (1791, p. 230) wrote, many Americans had "an earnest desire, that a more extensive demand for that surplus may be created at home." Domestic manufacturing, he would go on to argue, could provide that demand, and government policies should encourage it.

Refuting Arguments Unfriendly to Encouraging *Manufactures*

To a modern reader, aware of industrialization's contributions to the Great Enrichment of the past two centuries, the early pages of *Manufactures* might seem odd. In them, Hamilton paraphrases, often in passages set off by quotation marks, the economic arguments of Adam Smith and others to the effect that a country such as the United States should stick to agriculture. One argument is straight *laissez faire*: "[I]t can hardly ever be wise in a government to attempt to give a direction to the industry of its citizens. This under the quicksighted guidance of private interest, will if left to itself, infallibly find its way to the most profitable employment, and . . . the public prosperity will be most effectually promoted. To leave industry to itself, therefore, is, in almost every case, the soundest as well as the simplest policy" (Hamilton 1791, p. 232).

Another claim for the primacy of agriculture is that the situation of the United States in 1791—small population, vast land resources, scarce labor, high wages, and a scarcity of capital—meant that "a successful competition with the manufactures of Europe must be regarded as little less than desperate. Extensive manufactures can only be the offspring of a redundant, at least a full population" (Hamilton 1791,

p. 233). In essence, this was the argument of Jefferson, who likely picked it up from Adam Smith.

Hamilton countered that these arguments, while having "certainly respectable pretensions," did not actually govern the conduct of nations. They were based, moreover, on the notion that "[a]griculture is, not only, the most the most productive . . . species of industry" (as Adam Smith contended), but that agriculture was "the only productive species," as the French physiocrats asserted. In both cases, Hamilton (1791, p. 240) deemed the notion "both quaint and superficial. It amounts to this—That in the production of the soil, nature co-operates with man; and the effect of their joint labour must be greater than that of the labour of man alone." Hamilton (1791, pp. 240–241) noted that in manufacturing, nature—say, in the form of water-powered machinery—also cooperated with man:

> It is very conceivable, that the labour of man alone laid out upon a work, requiring great skill and art to bring it to perfection, may be more productive, *in value* [italics in the original], than the labour of nature and man combined, when directed towards more simple operations and objects: And when it is recollected to what an extent the Agency of nature, in the application of the mechanical powers, is made auxiliary to the prosecution of manufactures, the suggestion, which has been noticed, loses even the appearance of plausibility.

Hamilton (1791, pp. 241–42) completes his refutation of the economic superiority of agriculture by noting that labor in agriculture is "periodical and occasional, depending on seasons," while labor in manufacturing is "constant and regular, extending through the year, embracing in some instances night as well as day. . . . And if it may likewise be assumed as a fact, that manufactures open a wider field to exertions of ingenuity than agriculture, it would not be a strained conjecture, that the labour employed in the former, being at once more *constant* [italics in original], more uniform and more ingenious, than that which is employed in the latter, will be found at the same time more productive."

Why Manufacturing Makes a Positive Contribution to Economic Growth

Hamilton (1791, p. 249) next lists and discusses seven reasons why manufacturing establishments would make the total product and income—his terms were "Produce" and "Revenue"—of a country "greater than they could possibly be, without such establishments." Most items on the list are familiar to modern economists. The list:

1. The division of Labour.
2. An extension of the use of Machinery.
3. Additional employment to classes of the community not ordinarily engaged in business.
4. The promoting of emigration from foreign Countries.

5. The furnishing greater scope for the diversity of talents and dispositions which discriminate men from each other.
6. The affording a more ample and various field for enterprize.
7. The creating in some instances a new, and securing in all, a more certain and steady demand for the surplus produce of the soil.

The first item Hamilton takes from Adam Smith and discusses in Smithian terms. To illustrate the second, Hamilton (1791, p. 252) cites "[t]he Cotton Mill invented in England, within the last twenty years. . . . [A]ll the different processes for spining (sic) cotton are performed by means of Machines, which are put in motion by water, and attended chiefly by women and children. . . . And it is an advantage of great moment that the operations of this mill continue with convenience, during the night as well as through the day. . . . To this invention is to be attributed essentially the immense progress, which has been so suddenly made in Great Britain in the various fabrics of Cotton."

Hamilton explains the third item on his list in terms of increased labor force participation. Women and children were underemployed in American agriculture, the country's predominant economic activity. Manufacturing enterprises would provide them with job opportunities and cash incomes. Modern readers might question whether employing children in manufacturing was a good thing. But such thinking is anachronistic; in America and the world of 1791, children often worked at early ages. Manufacturing would also provide "occasional and extra employment to industrious individuals and families" when they were seasonally idle. He likely was referring to farmers and their families in the off-seasons.

Items 4–6 on Hamilton's list are almost self-explanatory. A diversified economy, with agricultural, commercial, and manufacturing sectors, would appeal to a wider range of immigrants, thereby alleviating America's labor shortage: "[T]he results of human exertion may be immensely increased by diversifying its objects" (Hamilton 1791, pp. 255–56). Moreover,

> The spirit of enterprise, useful and prolific as it is, must necessarily be contracted or expanded in proportion to the simplicity or variety of the occupations and productions, which are to be found in a Society. It must be less in a nation of cultivators, than in a nation of cultivators and merchants; less in a nation of cultivators and merchants, than in a nation of cultivators, artificers and merchants.

Item 7 on the list harks back to the beginning of *Manufactures*, where Hamilton indicated that the mercantilist regulations of the European empires made it difficult for the United States to export its surplus agricultural commodities. Mercantilist policies, Hamilton (1791, p. 258; italics in the original) now says, "sacrifice the interests of a mutually beneficial intercourse to the vain project of *selling everything* and *buying nothing*." They made the foreign demand for the products of American agriculture "casual and occasional" rather than "certain or constant." The United

States therefore needed a "an extensive domestic market," Hamilton (1791, p. 258) argued: "To secure such a market, there is no other expedient, than to promote manufacturing establishments. Manufacturers who constitute the most numerous class, after the Cultivators of land, are for that reason the principal consumers of the surplus of their labor."

Hamilton (1791, pp. 262–65) next shows that he knew his Adam Smith, and also knew why Smith's keen theoretical insights were not relevant to the situation faced by the United States:

> If the system of perfect liberty to industry and commerce were the prevailing system of nations—the arguments which dissuade a country in the predicament of the United States, from the zealous pursuit of manufactures would doubtless have great force. . . . But the system which has been mentioned, is far from characterizing the general policy of nations. [The prevalent one has been regulated by an opposite spirit.]
>
> The consequence of it is, that the United States are to a certain extent in the situation of a country precluded from foreign Commerce. They can indeed, without difficulty obtain from abroad the manufactured supplies, of which they are in want; but they experience numerous and very injurious impediments to the emission and vent of their own commodities. . . . The regulations of several countries, with which we have the most extensive intercourse, throw serious obstructions in the way of the principal staples of the United States. . . .
>
> If Europe will not take from us the products of our soil, upon terms consistent with our interest, the natural remedy is to contract as fast as possible our wants of her.

Refuting Other Objections to a Policy of Encouraging *Manufactures*

Hamilton then considers and refutes a number of additional arguments against government encouragement of manufacturing.

The first is a repeat of the *laissez-faire*, Adam Smith doctrine that "[i]ndustry, if left to itself, will naturally find its way to the most useful and profitable employments" without government aid (Hamilton 1791, p. 266). Hamilton answers that people are risk averse and reluctant to launch untried ventures. Also, the "long matured establishments of another country" make it difficult for new domestic ventures to compete on quality and price—what has become known as the infant-industry argument. Most problematic of all for would-be American manufacturers, however, was that other countries extensively subsidized the exports of "the establishments to be imitated." Hamilton (1791, pp. 266–68) hence concludes: "To produce the desireable (sic) changes, as early as may be expedient, may therefore require the incitement and patronage of government."

He next discusses contentions that manufacturing in the young United States could not be successful because of the "scarcity of hands—dearness of labour—want of capital." Hamilton (1791, pp. 269–71) admitted that the first two were obstacles, largely because the availability of cheap and fertile land on the frontier was a magnetic attraction for people in the more fully settled regions of the country. But, he noted, the US economy did have settled regions, and they were already engaged in manufacturing pursuits. He reiterates that women, children, and immigrants would become new sources of labor. His strongest emphasis, however, is machinery as a substitute for labor: "the vast extension given by the late improvements to the employment of Machines, which substitute the Agency of fire and water, has prodigiously lessened the necessity for manual labor."

Machinery, of course is a capital investment, and the United States supposedly had a shortage of capital. Not so, says Hamilton, in moving to a discussion of considerations that "remove all inquietude on the score of a want of Capital." In this part of *Manufactures*, Hamilton points with pride to the successes his previous fiscal and banking reforms were already having. New banks were lending, and foreign investors were transferring capital to the United States by purchasing Treasury bonds, shares in the Bank of the United States, and other corporate securities (Sylla, Wilson, and Wright 2006). Hamilton (1791, pp. 274–76) notes:

> The introduction of Banks . . . has a powerful tendency to extend the active Capital of a Country. Experience of the Utility of these Institutions is multiplying them in the United States. It is probable that they will be established wherever they can exist with advantage. . . .
>
> The aid of foreign Capital may safely . . . be taken into calculation. Its instrumentality has long been experienced in our foreign commerce, and it has begun to be felt in various other modes. Not only our funds [the national debt], but our Agriculture and other internal improvements have been animated by it. It has already in a few instances extended even to our manufactures. . . .
>
> It is at least evident, that in a Country situated like the United States, with an infinite fund of resources yet to be unfolded, every farthing of foreign capital, which is laid out in internal ameliorations, and in industrious establishments of a permanent nature, is a precious acquisition.

Most of Hamilton's discussion of why a supposed lack of capital was not a barrier to manufacturing, however, dealt not with banks and foreign capital, but to his restructuring of the national debt. "It happens," Hamilton (1791, p. 277) says, "that there is a species of Capital actually existing within the United States, which relieves from all inquietude on the score of a want of Capital—This is the funded Debt. . . . Public Funds answer the purpose of Capital, from the estimation in which they are usually held by Monied men; and consequently from the Ease and dispatch with which they can be turned into money."

What Hamilton meant was that the liquid securities markets emerging in major US cities in the early 1790s, together with a growing number of banks, made it possible for holders of public debt to sell their securities easily to gain money to finance other investments, or to collateralize the securities for bank loans for the same purpose. Sylla (1998; relying on Davis 1917) indicates that by 1792, banks accepted the new 6 percent US debt securities at 100 percent of par value as loan collateral.

For these reasons, "The operation of public funds as capital is too obvious to be denied." Of course, some writers, notably Adam Smith with his unfavorable view of public debt, did deny that more public debt could increase the capital of a country, and might even decrease the capital of a country if public debt crowded out private investment. Hamilton (1791, pp. 281–83, italics in original) spends several pages analyzing this issue, concluding:

> [I]t is important to distinguish between an *absolute increase of Capital, or an accession of real wealth, and an artificial increase of Capital*, as an engine of business, or as an instrument of industry and Commerce. In the first sense, a funded debt has no pretensions to being deemed an increase in Capital; in the last, it has pretensions which are not easy to be controverted. Of a similar nature is bank credit and in an inferior degree, every species of private credit. . . .
>
> [A]s far as the nature of the subject admits of it, there appears to be satisfactory ground for a belief, that the public funds operate as a resource of capital to the Citizens of the United States, and, if they are a resource at all, it is an extensive one.

What Hamilton did not directly say, but must have had in mind, is that his public-debt restructuring by December 1791 had greatly increased the market value of the national debt since he took office in September 1789. At the earlier date, the par value of the federal debt was $52.2 million (not including the later federal assumption of state debts), and its market value was $22.5 million. The market value was thus only 43 percent of the par value. By December 1791, the par value of the national debt (now including assumed state debts) was $77.3 million, and its market value risen to $69.6 million. If the roughly $20 million par value of assumed state debts had the same market value discount of 43 percent as the federal debt, its market value would have been $8.1 million, making the sum of federal and state debts have about $31.1 million in market value in 1789. This was far below the $69.6 million at which the market valued them a little more than two years later. (For context, in 1792, the first year the Treasury paid interest on assumed state debts, federal expenditures were $5.1 million, revenues $3.7 million, and the resulting deficit of $1.4 million raised the total national debt at par to $80.4 million. Rough estimates of US GDP at this time place it as around $200 million.)

Hamilton's policies thus had increased the wealth of public creditors by about $38.5 million, in the same way a rise in the stock market or in house values raises

the wealth of those asset owners today.[4] This increase in the capital was potentially available for investment. Studies of Treasury documents, moreover, indicate that the public debt was more widely held than earlier historians had supposed (Wright 2008). Hamilton had good reasons to take pride in the increased access to capital fostered by his policies.

After citing 17 areas of US manufacturing that had "grown up and flourished with a rapidity which surprises" (Hamilton 1791, pp. 283–84), Hamilton goes on to tout the benefits of a diversified agricultural, commercial, and manufacturing economy (Hamilton 1791, pp. 287–92). A diversified economy offers consumers a great variety of goods and services, and it reduces the risks of stagnation that an undiversified, agricultural economy faces when demand for its produce declines.

Finally, in this refutation section of *Manufactures*, Hamilton takes up perceptions of conflicts of interest between agriculturalists and manufacturers in the early years of the country. Some Southerners shared Jefferson's notions of agriculture as a way of life, and disliked Hamilton's plan to use national revenues to encourage manufacturing because most manufacturing was located in the New England and Middle Atlantic states. Northern manufacturers, although Hamilton does not get into this, wanted protection against the competition of imported manufactures and would come to view the rapid expansion of western agriculture as drawing away from them the supply of industrial labor and raising the wages they had to pay.

Hamilton denies any such conflicts of interests, saying that it is "a maxim well established by experience . . . that the aggregate prosperity of manufactures, and the aggregate prosperity of Agriculture are intimately connected." Because most proponents of the supposed conflict were southerners, Hamilton (1791, pp. 293–95) counters their doubts by noting that much manufacturing at the time involved processing the outputs of agriculture, forestry, and mining. Therefore, the farmers of the South and the North would benefit by a growing demand for their inputs by the expansion of manufacturing. A formal economic modeling exercise of the varied interests, conducted two centuries later, concluded that Hamilton was correct on these issues (Passell and Schmundt 1971).

Policy Tools to Encourage *Manufactures*

How did other countries promote their manufactures? Hamilton (1791, pp. 296–311) provides and discusses the pros and cons of a long list of policies:

I Protecting duties—or duties on those foreign articles which are rivals of the domestic ones, intended to be encouraged . . .
II Prohibitions of rival articles or duties equivalent to prohibitions . . .
III Prohibitions of the exportation of the materials of manufacture . . .
IV Pecuniary bounties . . .

[4]The par and market values of the national debt are from a spreadsheet compiled by George Hall, and shared with me by Thomas Sargent. They are described in Hall, Payne, and Sargent (2018).

V Premiums . . .
VI The Exemption of the Materials of manufactures from duty . . .
VII Drawbacks of the duties which are imposed on the Materials of Manufactures . . .
VIII The encouragement of new inventions and discoveries, at home, and of the introduction into the United States of such as may have been made in other countries; particularly those, which relate to machinery . . .
IX Judicious regulations for the inspection of manufactured commodities . . .
X The facilitating of pecuniary remittances from place to place . . . by rendering more easy the purchase of raw materials and the payment for manufactured supplies . . .
XI The facilitating of the transportation of commodities.

In his discussion of these policy tools, Hamilton (1791, pp. 296–97) notes that the United States already had mildly protective (not prohibitive) duties on many imports, which in fact were the federal government's main source of revenue from the 1790s to the 1860s. To promote manufacturing, he argued, such import duties should not fall on raw materials.

Prohibitive duties obviously encouraged domestic manufactures, but were "only fit to be employed when a manufacture, has made such a progress and is in so many hands as to insure a due competition, and an adequate supply on real terms" (Hamilton 1791, p. 297).

Hamilton (1791, pp. 297–98) was skeptical about export prohibitions. He was fond, however, of "bounties," which we call "subsidies," on several grounds. One was that bounties, unlike tariffs, encouraged manufacturing without raising prices to consumers. Another was that unlike high protecting duties, bounties did not tend to create scarcities. But "continuance of bounties on manufactures long established must almost always be of questionable policy" (Hamilton 1791, pp. 300–301). Despite his preference for bounties over protective tariffs, Hamilton noted: "There is a degree of prejudice against bounties from an appearance of giving away the public money . . . and from a supposition that they serve to enrich particular classes, at the expence of the Community." In fact, Congress would reject the bounties Hamilton proposed.

"Premiums"—essentially prizes—were akin to bounties, but more specific than general. They "serve to reward some particular excellence or superiority, some extraordinary exertion of skill. . . . But their effect is to stimulate general effort" (Hamilton 1791, pp. 304–05).

Hamilton thought that, with a few exceptions, it was good policy to exempt raw materials used by manufacturers from import duties. If the United States imposed duties on such materials, a "drawback," in the form of a refund of the duty paid by the domestic manufacturer, would be appropriate for manufactures to be particularly encouraged—but only for infant, not mature, industries.

Hamilton also thought that the encouragement to inventors and authors provided by patents and copyrights ought to extend to "Introducers" of foreign

improvements to the United States. He suggests that the foreign introducer, not necessarily the inventor, of an improvement receive US patent protection, but admits that an authority to do that was questionable. As an alternative, toward the end of *Manufactures* he would suggest (Hamilton 1791, pp. 338–40) the establishment and funding of a Board to encourage and pay for the transfer of important foreign technologies and the migration of workers skilled in them to the United States. Hamilton (1791, pp. 308–09) lamented the "selfish and exclusive policy" of other countries that sought to prevent technological transfers. As a public official, he had to respect other nations' intellectual property and laws. Unofficially, he and other American officials encouraged to a considerable extent the pirating and smuggling of protected foreign technologies (Ben-Atar 2004).

Regulated inspection of American manufactures to weed out shoddy goods would be a form of quality control that would both protect domestic consumers and increase the reputation of US exports (Hamilton 1791, pp. 308–09).

As regarded facilitating pecuniary remittances, a "general circulation of Bank paper" such as the currency notes of the Bank of the United States would aid interstate payments for raw materials and manufactured products. He further suggested that national rules making inland bills of exchange drawn in one state and payable in another negotiable everywhere would be another aid to interstate commerce (Hamilton 1791, pp. 309–10).

To justify the last item on his list of policy tools, transportation improvements, Hamilton approvingly introduces a long direct quotation from Adam Smith, without identifying Smith as the author. Among other things, Smith said these were "the greatest of all improvements" (Hamilton 1791, p. 311). Hamilton favored a national plan of transportation improvements, and direct aid from the federal government to implement it. From the 1790s to the 1860s, constitutional issues and clashing state interests undermined such a national program. During those decades, state and local governments would plan and execute nearly all internal transportation improvements.

Specific Policy Recommendations

In the final pages of *Manufactures*, Hamilton proposes increases and reductions in existing tariffs for some manufactures and raw materials, and bounties for others. The affected manufactures and raw materials included iron, copper, lead, coal, wood, skins, grain, flax and hemp, cotton, wool, silk, glass, gun powder, paper, printed books, and refined sugars and chocolate.[5] Hamilton (1791, pp. 313–14)

[5] In his discussion of cotton, almost as an aside, Hamilton (1791, p. 328) writes: "[I]t may be announced, that a society is forming with a capital which is expected to be extended to at least half a million of dollars; on behalf of which measures are already in train for prosecuting on a large scale, the making and printing of cotton goods." This was the Society for Establishing Useful Manufactures. What Hamilton did not say is that he himself had selected the site in Paterson, New Jersey, where the Passaic River offered waterpower to run machinery. He also authored the Society's corporate charter and worked to ensure its enactment in 1791 by the New Jersey legislature. The Society was a mixed success. Mismanagement prevented it from opening the factories it intended to build in the 1790s as demonstration projects. But the corporation continued to provide sites and power for other entrepreneurs into the twentieth

selected these industries based on five criteria: "the capacity of the Country to furnish the raw material—the degree in which the nature of the manufactures admits of a substitute for manual labour in machinery—the facility of execution—the extensiveness of the uses, to which the article can be applied—its subserviency to other interests, particularly the great one of national defense."

The details of Hamilton's tariff recommendations need not detain us here. Irwin (2004, pp. 812–13) nicely summarizes them in a one-page table, and makes some pertinent points. One is that, despite what many historians said about him, Hamilton was not a protectionist: "The import duties he proposed were quite modest in comparison to what domestic manufacturers would have liked (and in comparison to those imposed later in the nineteenth century)." Duties already in effect in 1791 ranged from 5 to 12.5 percent ad valorem. Hamilton's proposals changed this range from zero to 15 percent. The modest duties caused manufacturers who desired more protection to shift their political support away from Hamilton's Federalist Party and toward Jefferson's Democratic Republican Party, which favored tougher measures to reduce imports from Britain, the leading trading partner of the United States. Hamilton and his Federalist party would pay a political price for their support of moderate tariffs for revenue instead of high tariffs to protect manufacturers.

According to Irwin (2004, pp. 813–14), Hamilton in *Manufactures* "was skeptical of high protective tariffs because they sheltered inefficient and efficient producers alike, led to high prices for consumers, and gave rise to smuggling, which cut into government revenue." And Hamilton very much needed more revenue. Federal revenues, mostly from customs duties, were not sufficient to cover interest payments on the national debt and fund the government's ordinary operations until the last year (1794–1795) of Hamilton's tenure as Treasury Secretary. In the interim, to cover the revenue shortfall, Hamilton had to borrow from domestic and foreign sources (Sylla 2010).

One policy proposal toward the end of *Manufactures* deserves special attention. In discussing iron, the manufacture of which he deemed "entitled to preeminent rank," Hamilton (1791, pp. 314–17) proposed a tariff of 15 percent, his top rate, on imports of firearms and other military weapons. Weapons manufacturers already existed and "only require the stimulus of a certain demand to render them adequate to the supply of the United States." To ensure that demand, he proposed that the federal government make annual purchases of weapons of domestic manufacture, store them in government arsenals, and replace any withdrawals. He further proposed that that the government itself should consider manufacturing weapons, as an exception to the "general rule" that "manufactories on the immediate account of Government are to be avoided." Congress adopted the essence of Hamilton's proposals (as discussed below), which had beneficial long-term effects on US technological development and industrial growth. National security and economic growth continue to be objectives of current industrial policies.

---

century, and Paterson became a major center of American manufacturing (Davis 1917). There is now a Paterson National Historical Park to commemorate these Hamiltonian industrial origins.

Practical Effects of *Manufactures*

Congress received, debated, and enacted the essential provisions of Hamilton's earlier reports on Public Credit in 1790 (Hamilton 1790a, pp. 65–110), the National Bank (Hamilton 1790b, pp. 305–42), and a Mint (Hamilton 1790b, pp. 570–607). In the case of *Manufactures*, it only received the report; there was no debate on it, and no comprehensive enactment of its policy proposals.

There are, however, two examples of quick adoption of Hamilton's recommendations. Within months of receiving *Manufactures*, Congress in 1792 adopted most of Hamilton's recommended tariff modifications. They did so, however, less to encourage manufacturing than to gain revenue to fund increases in military spending after a disastrous rout of a US army by Native American forces on the western frontier (Irwin 2004).

Then in 1794, Congress passed "An Act to provide for the erecting and repairing Arsenals and Magazines, and for other purposes" (Peters 1845–1867, p. 352). The act led to the establishment of arsenals and armories owned and operated by the federal government, as Hamilton had suggested in *Manufactures*. Over subsequent decades, these public enterprises, especially the federal armory at Springfield, Massachusetts, became hotbeds of technological development in machine tools, standardized interchangeable parts, and mass production (Smith 1977). By the middle of the nineteenth century, American armaments-making technology was second to none. British officials came to the United States in the 1850s to study the technology and to purchase it for Britain's armories.

As the nineteenth century unfolded and the United States would pass Britain to become the leading manufacturing nation, policymakers only sometimes chose a Hamiltonian path, and not always in ways Hamilton had suggested. US tariffs first became consciously protective in 1816, and rose to still higher levels of protection amidst political controversy in 1824 and 1828. Before 1816, international trade was severely disrupted by the protracted Napoleonic wars in Europe, the US embargo of 1808 (in which the US passed a law forbidding US imports and exports) and related trade restrictions, and the War of 1812 between the United States and Britain. Import disruptions of that time stimulated domestic manufacturing, especially in mechanized cotton spinning. When peace came in 1815, Britain flooded the United States with manufactured imports, prompting cries for protection, to which Congress responded. Britain clearly intended to squelch nascent US manufacturing by dumping its manufactures on American markets. Henry Lord Brougham, a member of Parliament, wrote, "It was well worthwhile to incur a loss upon the first exportation in order by glut to stifle in the cradle those rising manufactures in the United States which the war has forced into existence contrary to the usual course of nature" (quoted by Higgins 2024, p. 30).

The US government did not implement transportation improvements according to the sort of comprehensive national plan that Hamilton recommended in *Manufactures*, and that Albert Gallatin, Jefferson's Treasury Secretary, would endorse in a lengthy report on roads and canals delivered in 1808. National

politics and presidential vetoes repeatedly defeated proposals for federal involvement in this area. Instead, until the 1860s, state and local governments would do most of the sponsoring and financial supporting of roads, canals, river and harbor clearing, and railroad building. Still, substantial public investments did occur, and the use of federal budget surpluses to pay off the national debt aided state government borrowing to fund them. Domestic transportation costs fell dramatically, extending and linking markets in the ways that both Adam Smith and Hamilton had envisioned.

After 1860, US economic policy became more decidedly Hamiltonian. Congress raised tariffs, first to help finance the Civil War and then to generate surpluses that gradually paid down the war debt during the late nineteenth century. In 1863–1864, it established a national banking system and introduced a uniform national currency backed by the federal government to replace the previous system in which thousands of state-chartered banks issued their own currencies. The US government went well beyond Hamilton's board to promote arts, agriculture, manufacturing, and commerce by establishing both a Department of Agriculture and land-grant colleges to foster research and education in agriculture and the mechanic arts. It encouraged railroad building with grants of federal land along the routes that railway companies developed, and authorized financial subsidies from the Treasury for some railway companies. In the Homestead Act of 1862, which granted free federal land to settlers who would live for a time on the land and develop it, Congress went beyond anything Hamilton had recommended. The act did stimulate immigration, which was one of Hamilton's goals to alleviate labor scarcity.

During what economic historians call the "long nineteenth century" from 1789 to 1915, US industrial production rose dramatically. According to the Davis index (Davis 2004), industrial production expanded steadily over the entire period at an average annual rate of roughly 5 percent, or a doubling roughly every 14 years. Industrial production in 1915 was 455 times what it was in 1790; over same period US population increased 26 times, from 3.9 million to 100.5 million. When World War I began, the United States produced more than one-third of world industrial output. No other country came close. Hamilton could not have asked for more.

Business-cycle interruptions to the growth of industrial production were relatively short. Whenever the Davis index of industrial production reaches a peak, the trough is mostly a year later, sometimes two years, and then expansion resumes. The only exceptions to steady secular growth around 5 percent per year appear to be a slowdown from roughly 1807 into the mid-1820s, and then above-average growth from the mid-1820s to the mid-1850s. The former period featured Jefferson's embargo on US exports, the War of 1812, and the Panic of 1819. The latter period marked the rapid expansion of mechanized cotton textile production, the leading manufacturing industry of the antebellum decades, and the early boom in railroad building. Industrial production grew at roughly 5 percent per year both before and after the Civil War.

Of course, we should not attribute the remarkable and steady expansion of nineteenth-century US industrial output entirely or even mostly to Hamilton's paper

on *Manufactures*. Many other factors were involved, including territorial acquisitions, the spread of banking and capital markets, the discovery and exploitation of natural resources, waves of immigration, widespread land ownership, property rights, the growth of a vast domestic market, and broadly laissez-faire policies encouraged by public policies of the sort Hamilton had recommended. But back in 1791, Hamilton had a prescient vision of how America's economy with governmental backing could shift from an agricultural economy and become a major manufacturing nation. His vision was not far off from the way that future actually unfolded.

■ *For constructive comments on an early draft of this article, the author thanks Jonathan Parker, Hugh Rockoff, Nancy Spannaus, Timothy Taylor, and Heidi Williams.*

References

Ben-Atar, Doron S. 2004. *Trade Secrets: Intellectual Piracy and the Origins of American Industrial Power.* New Haven: Yale University Press.

Bodenhorn, Howard. 2010. "Federal and State Commercial Banking Policy in the Federalist Era and Beyond." In *Founding Choices: American Economic Policy in the 1790s*, edited by Douglas A. Irwin and Richard Sylla, 151–76. Chicago: University of Chicago Press.

Cohen, Stephen S., and J. Bradford DeLong. 2016. *Concrete Economics: The Hamilton Approach to Economic Growth and Policy.* Boston: Harvard Business Review Press.

Davis, Joseph H. 2004. "An Annual Index of U.S. Industrial Production, 1790–1915." *Quarterly Journal of Economics* 119 (4): 1177–1215.

Davis, Joseph Stancliffe. 1917. *Essays on the Earlier History of American Corporations.* Cambridge: Harvard University Press.

DiLorenzo, Thomas J. 2009. *Hamilton's Curse: How Jefferson's Arch Enemy Betrayed the American Revolution—and What It Means for Americans Today.* New York: Three Rivers Press.

Fasteau, Marc, and Ian Fletcher. 2024. *Industrial Policy for the United States: Winning the Competition for Good Jobs and High-Value Industries.* New York: Cambridge University Press.

Hall, George, Jonathan Payne, and Thomas Sargent. 2018. "US Federal Debt 1776–1960: Quantities and Prices." New York University Leonard N. Stern School of Business New York University Working Paper 18-25.

Hamilton, Alexander. 1790a. *The Papers of Alexander Hamilton*. Vol. VI, edited by Harold C. Syrett. New York: Columbia University Press, 1962.

Hamilton, Alexander. 1790b. *The Papers of Alexander Hamilton*. Vol. VII, edited by Harold C. Syrett. New York: Columbia University Press, 1963.

Hamilton, Alexander. 1791. *The Papers of Alexander Hamilton*. Vol. X, edited by Harold C. Syrett. New York: Columbia University Press, 1966.

Higgins, Mark J. 2024. *Investing in U.S. Financial History: Understanding the Past to Forecast the Future.* Austin, TX: Greenleaf Book Group Press.

Hogeland, William. 2024. *The Hamilton's Scheme: An Epic Tale of Money and Power in the American Founding.* New York: Farrar, Straus and Giroux.

Irwin, Douglas A. 2004. "The Aftermath of Hamilton's 'Report on Manufactures.'" *Journal of Economic History* 64 (3): 800–821.

Irwin, Douglas A. 2017. *Clashing Over Commerce: A History of US Trade Policy*. Chicago: University of Chicago Press.

Jefferson, Thomas. 1785. *Notes on the State of Virginia*. Edited with an Introduction and Notes by Robert Pierce Forbes. New Haven: Yale University Press, 2022.

McCloskey, Deirdre Nansen, and Art Carden. 2020. *Leave Me Alone and I'll Make You Rich: How the Bourgeois Deal Enriched the World*. Chicago: University of Chicago Press.

McCraw, Thomas K. 1994. "The Strategic Vision of Alexander Hamilton." *American Scholar* 63 (1): 31–47.

Parenti, Christian. 2020. *Radical Hamilton: Economic Lessons from a Misunderstood Founder*. London and New York: Verso.

Passell, Peter and Maria Schmundt. 1971. "Pre-Civil War Land Policy and the Growth of Manufacturing." *Explorations in Economic History* 9 (1): 35–48.

Peters, Richard, ed. 1845–1867. *The Public Statutes at Large of the United States of America, from the Organization of the Government in 1789, to March 3, 1845*. 8 Vols. Boston: Charles C. Little and James Brown.

Schumpeter, Joseph A. 1954. *History of Economic Analysis*. New York: Oxford University Press.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York: The Modern Library, 1937.

Smith, Merritt Roe. 1977. *Harpers Ferry Armory and the New Technology: The Challenge of Change*. Ithaca: Cornell University Press.

Sylla, Richard. 1998. "U.S. Securities Markets and the Banking System, 1790–1840." *Federal Reserve Bank of St. Louis Review* 80 (3): 83–98.

Sylla, Richard. 2010. "Financial Foundations: Public Credit, the National Bank, and Securities Markets." In *Founding Choices: American Economic Policy in the 1790s*, edited by Douglas A. Irwin and Richard Sylla, 59–88. Chicago: University of Chicago Press.

Sylla, Richard, John B. Legler, and John J. Wallis. 1987. "Banks and State Public Finance in the New Republic, 1790–1860." *Journal of Economic History* 47 (2): 391–403.

Sylla Richard, Jack W. Wilson, and Robert E. Wright. 2006. "Integration of Trans-Atlantic Capital Markets, 1790–1845." *Review of Finance* 10 (4): 613–44.

Wright, Robert E. 2008. *One Nation Under Debt: Hamilton, Jefferson, and the History of What We Owe*. New York: McGraw Hill.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 131–154

Evaluating Behavioral Incentive Compatibility: Insights from Experiments

David Danz, Lise Vesterlund, and Alistair J. Wilson

In a mechanism design framework, the economist acts as an engineer, choosing the incentives and rules of an environment to shape participants' behavior towards the designer's objective. For example, the task may be to select bidding rules to maximize revenue from bidders in an auction or to construct a matching algorithm to efficiently allocate applicants to a limited number of medical residency positions. A core challenge in designing a mechanism is that the specific outcome the designer wants to achieve depends on agents' "types," where these types are private information, unknown to the designer. An agent's type includes any information relevant to their decision and the designer's objective, such as their willingness to pay for an item in an auction or their personal rankings of medical residency programs. Successful implementation of a mechanism hinges on the ability to acquire information on types, but it may not be in the agent's interest to reveal this information. For example, a bidder in an auction may be reluctant to reveal their true willingness to pay for the item if it adversely affects the price they must pay. In designing a mechanism, the economist aims both to provide agents with incentives that make them want to reveal their type, and to implement the designer's ideal objective given their types.

In modeling this problem, Hurwicz (1972, 1973) cleverly treated the agent's decision in response to the mechanism as simultaneously revealing their type and

■ *David Danz is a Research Assistant Professor, Lise Vesterlund is the Andrew W. Mellon Professor of Economics, and Alistair J. Wilson is a Professor of Economics, all at the University of Pittsburgh, Pittsburgh, Pennsylvania. Lise Vesterlund is also a Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. Their email addresses are danz@pitt.edu, vester@pitt.edu, and alistair@pitt.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.131.

securing the designer's intended outcome. That is, the selected mechanism is one that addresses the designer's objective subject to the agents' *incentive compatibility constraints*, where these constraints ensure that the only valid rules and incentives set by the designer are those that make the agents prefer to reveal their type truthfully.

Under the assumptions that agents are cognitively perfect and rational and that they hold certain preferences, theoretical modeling of the incentive compatibility constraint has led to the development of countless mechanisms. However, research is showing that when human decision-makers are faced with these mechanisms, they often fail to reveal their type, suggesting that the mechanisms are not incentive compatible in a behavioral sense. Individuals faced with mechanisms that are not *behaviorally incentive compatible* will not reveal their type, leading the designer to select outcomes that differ from their objective: auction revenue not being maximized with participants underbidding, or the allocation of applicants to residency programs being inefficient (and unstable) because hospital-resident pairs want to break from the given match.[1]

In using and improving mechanisms, it is critical that we determine whether they are behaviorally incentive compatible. Although mechanisms are designed to be used in the field, it is not possible in a field setting to verify that they succeed in eliciting participants' private "types." Experimental studies allow for such verification and have served a critical role in assessing whether mechanisms are behaviorally incentive compatible. The reason is that we in an experimental study directly can induce a participant's type and observe whether the induced type is revealed under the mechanism (referred to as *truthful revelation*). While the laboratory differs from the field, the structure of the incentives is the same, and mechanisms that fail in the lab are expected to similarly fail in the field (for example, Kagel and Roth 2000; Kessler and Vesterlund 2015).

This paper will review the techniques used in experiments to assess behavioral incentive compatibility. The experimental tests discussed have been applied to a wide set of mechanisms, including auctions, centralized clearinghouses, and others. However, to demonstrate these techniques we use as a running example the conceptually simple mechanism of eliciting beliefs from individuals where the designer's objective is one of truth-telling. As an example, we may want to learn how likely people think it is that a specific event occurs, say, that the Federal Reserve decreases interest rates by 50 basis points. To achieve truth-telling, we can elicit this belief by presenting incentives that depend on the actual realization of the event and make

[1] In formal terms, consider a screening problem where we abstract from strategic interactions and try to identify an individual's private type, $\theta \in \Theta$, which captures their preference over a set of outcomes $\mathcal{A}$, where $x \succ_\theta y$ indicates a strict preference for x over y. The designer asks the individual to report a type q, and in trying to get truthful revelation, selects a direct mechanism, a rule outlining an outcome $\phi(q) \in \mathcal{A}$ for every report q. A direct mechanism ϕ is incentive compatible if $\phi(q = \theta) \succ_\theta \phi(q = \theta')$ for every possible alternative report $\theta' \neq \theta$. In a strategic mechanism, the incentive compatibility condition will be based on a truthful report being an expected best response conditional on equilibrium behavior of all other types in a Bayesian implementation; or for all possible reports by the other players in a dominant-strategy implementation.

it in the respondent's best interest to report accurately their subjective assessment over the likelihood that rates are decreased. In the case of belief elicitation, the individual's private type is the belief that they hold over the event, with the designer's objective merely being one of truth-telling. So in this case, the designer's objective and the incentive compatibility constraint coincide.

The advantage of using individual belief elicitations to demonstrate experimental tests of behavioral incentive compatibility is that we can ignore specifics of the designer's objective (which here coincides with truth-telling) and any speculation on the behavior of others (as the elicitation is an individual-decision problem, not a strategic game). As such, we can focus squarely on whether participants under the mechanism see it as in their interest to reveal an induced belief. For example, we can in an experiment directly induce a belief of say 30 percent for the participant—by rolling a ten-sided die and asking participants for reports on the likelihood that a 1, 2, or 3 will appear; or by drawing a ball from an urn with 100 balls, of which 30 are blue, and asking for a report on the likelihood that a drawn ball is blue. After inducing the given type (the belief of 30 percent), we can then assess whether a particular belief elicitation mechanism succeeds in incentivizing reports on the induced belief.

In this paper, we begin by motivating the need for incentive-compatible mechanisms to elicit beliefs. We then use belief elicitations to present the techniques used to explore truthful revelation. First, we review tests centered on evaluating behavior under the mechanism of interest. While these tests can demonstrate failure to reveal the induced type, they do not reveal whether the failure results from the mechanism's incentives or from some other aspect of the mechanism. We therefore refer to these as *indirect* assessments of behavioral incentive compatibility. Tests include evaluations within a mechanism of whether participants reveal an induced type, comparisons between mechanisms to evaluate which comes closer to truthful revelation, as well as assessments of what might cause deviations. Second, we report on more recent *direct* assessments of behavioral incentive compatibility. These assessments directly evaluate the mechanism's incentives by asking whether participants prefer the designed incentive for their type to the other alternatives, and by testing whether full and easily understood information on the incentives increases truthful revelation. Throughout the discussion, we will provide evidence suggesting that although some of the most-used belief-elicitation mechanisms are theoretically incentive compatible, because of either failed modeling of the individual's preferences or cognitive abilities they are not behaviorally incentive compatible. Indeed, the incentives used are shown to distort reports, and researchers will often fare better if instead of explaining the mechanism or the incentives to the participants, they just tell them "you will maximize your expected earnings if you give your best estimate."

Why Elicit Beliefs with Mechanisms?

Getting information on people's beliefs is important for assessing collective expectations and for understanding human behavior. In many situations,

researchers will be interested in understanding the extent to which beliefs affect the choices that people make (Manski 2004). Do differences in college attendance result from differences in aptitude or from differences in the expected return from education? Do workers differ in their propensity to apply for promotion because of differences in risk aversion or because of differences in perceptions of how talented they think they are? Is the fact that some people have a greater reluctance to bargain driven by a concern for their counterpart, by the belief that bargaining will result in backlash, or by a belief that they are "not good" at it? Assessing and controlling for beliefs helps us understand behavior and formulate effective policy interventions.

In these and other settings, why not just ask people about their beliefs? Indeed, surveys about beliefs are a common technique used by social scientists. For example, participants could be asked in a survey to report whether they agree with the statement that their relative performance on a test will be in the top quarter of their cohort, perhaps using a five-point scale ranging from "strongly disagree" to "strongly agree." While easy to understand, the reports given may mean different things to different people. One person's "*disagree*" could be another's "*strongly disagree*." To put things on a common scale, we may instead ask participants to report the likelihood that they are ranked in the top quarter of the performance distribution.

But while we might fine-tune the questions we ask, it is harder to encourage the honest and reflective answers we are hoping for. Participants may have a sense that it is likely that they are in the top performance quarter, but find it is difficult to determine how likely. It takes effort to provide a probabilistic assessment of an event occurring: effort to understand the question through this quantitative lens, and perhaps effort to not brag and tell others that you are certain you are in the top quarter (Ewers and Zimmerman 2015), or to be humble and report that you are unlikely to be top-ranked (Thoma 2016).

To encourage truthful reporting, economists have resorted to paying participants. These payments differ from common flat-fee payments for completing a survey because the aim is not one of compensating for time spent, but instead to provide incentives for accuracy of the provided information. Economists have focused on mechanisms that present participants with incentives that make it in their interest to report their beliefs truthfully.[2] An incentive-compatible belief elicitation is structured to reward consideration, to increase accuracy, and to reduce noise in the response.

To see how the incentives selected for a mechanism can achieve this goal, consider the case of the "quadratic-scoring rule" (Brier 1950), one of the earliest deployed elicitation mechanisms (initially developed to assess the accuracy of weather forecasts). Suppose we want to elicit an individual's probabilistic belief $q \in [0,1]$ over a binary event E (say, being in the top performance quarter). Under the quadratic-scoring rule, the individual's monetary reward $\pi(q) \in [\$0, \$X]$ depends

[2] Incentive compatible rules have been shown to outperform incompatible ones (Nelson and Bessler 1989; Palfrey and Wang 2009; Schotter and Trevino 2014) and these in turn dominate unincentivized elicitation (Gächter and Renner 2010; Wang 2011; Trautmann and van de Kuilen 2015).

on their stated belief q, and on the (squared) prediction error based on the realized event E:

$$\pi(q) = \begin{cases} \$X \cdot \left[1 - (1-q)^2\right], & \text{if event } E \text{ occurs,} \\ \$X \cdot \left[1 - q^2\right], & \text{otherwise.} \end{cases}$$

As a numerical example, suppose that someone believes they have an 80 percent chance of scoring in the top quarter on a test. If they report a belief of $q = 0.8$ their payoff will be:

$$\pi(q) = \begin{cases} \$10 \cdot \left[1 - (1-0.8)^2\right], & \text{if event } E \text{ occurs,} \\ \$10 \cdot \left[1 - 0.8^2\right], & \text{otherwise.} \end{cases}$$

That is, this person receives \$9.60 if they actually are in the top quarter, but only \$3.60 if they are not. Given the true belief that there is an 80 percent chance of being in the top quarter, the person expects an 80 percent chance of the high payment and a 20 percent chance of the low payment, yielding an expected payoff of reporting $q = 0.8$ of $0.8(\$9.60) + 0.2(\$3.60) = \$7.68 + \$0.72 = \$8.40$.

Central to the quadratic-scoring rule is that participants who maximize expected payoffs have an incentive to report their prediction accurately. For example, suppose that instead of reporting their true belief of $\theta = 0.8$, they report $q = 0.6$ on being in the top quarter. Given the incentives under the quadratic-scoring rule, a reported belief of 0.6 leads to a payoff of \$8.40 if they are in the top quarter and \$6.40 if they are not. While the participant may report any q they wish, their actual belief of $\theta = 0.8$ that they are in the top quarter is fixed, and so their expected payoff of making this incorrect prediction is $0.8(\$8.40) + 0.2(\$6.40) = \$6.72 + \$1.28 = \$8.00$. As a result, their expected payoff is lower under a report of 0.6 than if they had reported their true belief of 0.8. Reporting a higher belief of say $q = 1.00$ is also disadvantageous. Here the payoff would be \$10 when in the top quarter and \$0 when not in the top quarter, and so the expected payoff is $0.8(\$10) + 0.2(\$0) = \$8.00$, again lower than reporting the actual belief.

As this example illustrates, individuals who want to maximize their expected earnings will prefer to report their true belief θ, because any other report lowers their expected earnings.[3] To put it another way, participants of type θ prefer the incentives meant for them, over those intended for other types.

While the quadratic-scoring rule is theoretically incentive compatible for agents aiming to maximize their *expected* earnings, truthful revelation depends on

[3] More generally, given an actual belief of θ that E occurs, the participant's expected payoff when reporting q is given by:

$$E_\theta \pi(q) = \$X \cdot \left[\theta \cdot \left[1 - (1-q)^2\right] + (1-\theta) \cdot \left(1 - q^2\right)\right].$$

By deriving a first- and second-order condition over the reported value q, we confirm that the unique maximizer is to truthfully report $q^*(\theta) = \theta$.

how individuals respond to the presented incentives. They may make mistakes when attempting to calculate their expected earnings or apply behavioral rules-of-thumb when faced with such problems, or they may not make choices to maximize their expected earnings. For example, individuals who are risk averse over the stakes will be drawn to report a more conservative belief, closer to the center $(q = 1/2)$, to get payoffs that vary less with the realized event. Indeed, concerns that the quadratic-scoring rule is not incentive compatible for risk-averse individuals (Winkler and Murphy 1970), and experimental evidence that it may not be behaviorally incentive compatible (for example, Offerman et al. 2009), has led to the development of belief elicitations that are incentive compatible for arbitrary risk preferences (for example, Hossain and Okui 2013; Mobius et al. 2022). Next, we discuss the experimental techniques that have been used to assess whether a mechanism is behaviorally incentive compatible.

Indirect Assessments of Behavioral Incentive Compatibility

We begin by reviewing the experimental tests that assess truthful reporting under the mechanism. That is, we provide participants with information on the likelihood of an event to induce the participants' belief θ that the event occurs and assess if, when presented with the incentives under a mechanism, reports on their type, q, correspond to the induced type, θ.

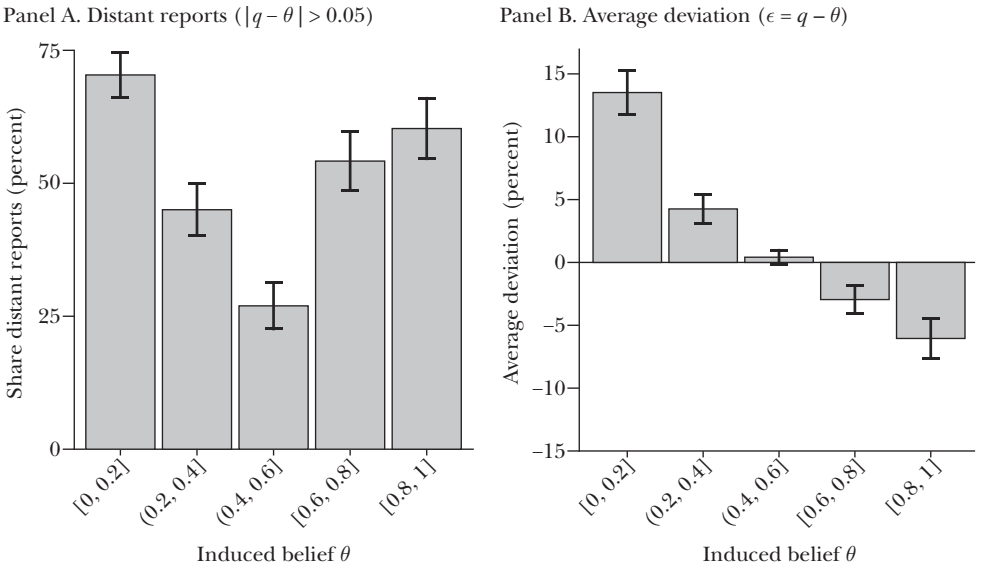
While informative on truthful revelation under the mechanism, these tests do not isolate the effect of incentives from a particular mechanism or directly evaluate preferences over the incentives within that mechanism. Hence, we refer to these tests as *indirect* assessments of behavioral incentive compatibility. They include performance evaluations within a particular mechanism and across mechanisms to determine which comes closer to truthful revelation, as well as assessments of what might cause deviations.

For the purposes of this paper, we will focus on the elicitation of simple induced beliefs, where probabilities are straightforward to see and can be understood with virtually no computational effort, like probabilities based on rolling a die or drawing from an urn. There are of course many studies that compare belief elicitations when induced beliefs require greater computation (for a review, see Benjamin 2019), or when subjective beliefs are elicited (over the behavior of other players in a strategic game, as in Nyarko and Schotter 2002). Our focus on simpler settings rules out confusion that could arise from determining the induced belief and makes it easy to determine whether the reported beliefs differ from the induced ones. If we find that a mechanism fails in a simple setting, we should not expect it to fare better when eliciting more complex beliefs.

Assessing Truthful Revelation Within and Across Mechanisms

Behavioral incentive compatibility is often assessed within a mechanism by simply checking how often reports under the mechanism correspond to the

Figure 1
Reporting Behavior in the Quadratic Scoring Rule



Source: Figures based on the published data from elicitations using the quadratic scoring rule in Offerman et al. (2009); Hossain and Okui (2013); Erkal, Gangadharan, and Koh (2020); and Danz, Vesterlund, and Wilson (2022). Total sample size is 426 participants and 3,213 total decisions.
Note: The figure shows the fraction of distant reports (panel A) and the direction of deviations (panel B) by induced belief (binned into intervals).

induced belief. To demonstrate, we report on studies examining reports under the quadratic-scoring rule, pooling more than 3,000 decisions from Offerman et al. (2009), Hossain and Okui (2013), Erkal, Gangadharan, and Koh (2020), and Danz, Vesterlund, and Wilson (2022).

In panel A of Figure 1, we show by ranges of the induced belief, θ, the fraction of reports, q, that were more than 5 percentage points from θ. We refer to these as "distant reports." For example, the first bar shows that when the induced belief is a number in the range of 0 to 0.2, a full 70 percent of reports deviated by more than 5 percentage points from the induced belief. Across all induced beliefs, 49 percent of reports deviated by more than 5 percentage points and only 43 percent of reports were exactly equal to the induced belief. Furthermore, we see a systematic decrease in the frequency of distant reports when the induced belief is closer to the center, with it being smallest in the center range from 0.4 to 0.6. For noncentered induced beliefs (outside of the 0.4 to 0.6 range) the majority of distant reports pull toward the center and 10 percent claim an exactly centered belief of $q = 0.5$. Evidence of center-biased reporting is also seen in panel B of Figure 1 where the average deviation from the induced belief tends to be positive when the induced belief is less than one-half and negative when the induced belief is more than one-half. In addition,

the average size of the deviation is largest when the induced belief is large or small, because centered reports are farther from these values.

As an assessment of the performance of the quadratic-scoring rule, panels A and B of Figure 1 demonstrate that participants within the mechanism largely fail to report the induced belief. That is, the mechanism does not appear to be behaviorally incentive compatible. Particularly concerning is that deviations from the induced beliefs are large and systematic. Econometrically, center-biased reporting will bias the underlying estimates if we use the reported beliefs q in place of the true beliefs θ, as either an explained or explanatory variable in a regression.

Another popular experimental technique for assessing behavioral incentive compatibility is to compare the performance of different mechanisms to determine which comes closer to truthful revelation. For example, this "horserace" methodology has been used by Hossain and Okui (2013), Erkal, Gangadharan, and Koh (2020), and Danz, Vesterlund, and Wilson (2022) to compare the classic quadratic-scoring rule and the binarized version of the quadratic-scoring rule. The latter was designed to be incentive compatible for individuals irrespective of their risk preferences, and thus address the concern that risk aversion may cause center-biased reporting (Hossain and Okui 2013).

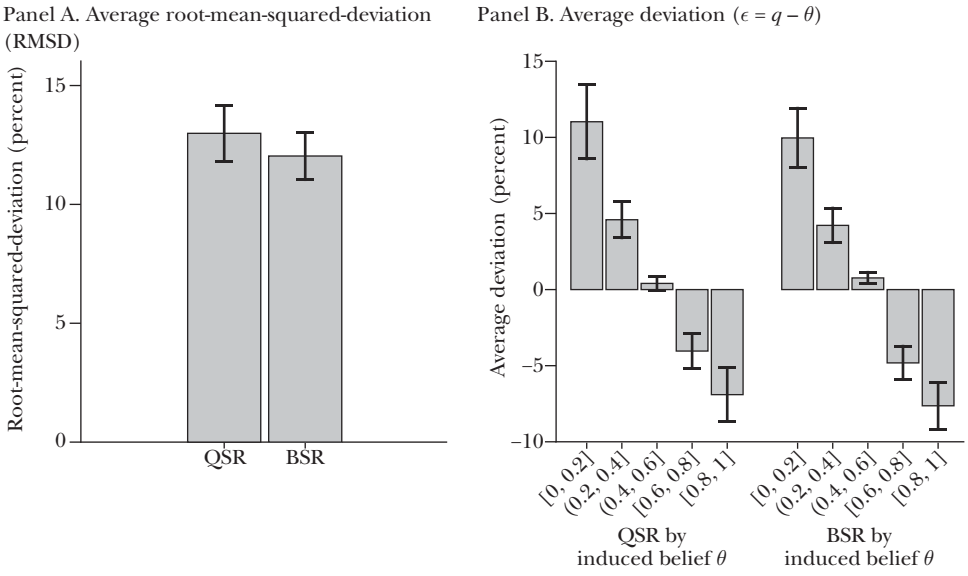
In both the classic and the binarized quadratic-scoring rule, the participant's payment depends on their (squared) prediction error. While payment is decreasing in the prediction error for the classic quadratic-scoring rule, payment for the binarized-scoring rule is a percentage chance of winning a fixed monetary prize, say \$10, and a larger prediction error instead decreases the chance that the participant wins the prize. Specifically, participants under the binarized-scoring rule are incentivized by a state-contingent lottery pair, where a reported belief of q on a binary event E is compensated with a $1 - (1 - q)^2$ chance of winning \$10 if the event occurs; and a $1 - q^2$ chance of winning \$10 if the event does not occur. Thus, if a participant believes and reports that there is an 80 percent chance of an event happening, then the chance of winning \$10 is 96 percent if the event occurs and 36 percent if the event does not occur, where the chance of winning the prize is maximized when the true belief is reported. While the classic quadratic-scoring rule is theoretically incentive compatible for risk-neutral individuals, the binarized version of the scoring rule is theoretically incentive compatible for arbitrary risk preferences.

A common measure of performance success used in horseraces between mechanisms is the square root of the sum of the squares of the deviations between the reported belief and the belief induced by the researcher (specifically, the root-mean-squared-deviation is RMSD $= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(q_i - \theta_i)^2}$). Pooling results from Hossain and Okui (2013), Erkal, Gangadharan, and Koh (2020), and Danz, Vesterlund, and Wilson (2022), we can compare the root-mean-squared-deviation under the classic and binarized quadratic-scoring rule.

The results for the pooled data are shown in panel A of Figure 2. Revealing that while there is substantial deviation from the induced belief under both elicitations, the average root-mean-squared-deviation is smaller in the binarized (BSR) than in

Figure 2

A Comparison of the Quadratic Scoring Rule and the Binary Scoring Rule



Source: Figures based on the published data from binary scoring rule (BSR) and quadratic scoring rule (QSR) elicitations in Hossain and Okui (2013); Erkal, Gangadharan, and Koh (2020); and Danz, Vesterlund, and Wilson (2022).

Note: All data use the Hossain and Okui definition of "betweenness" to exclude participants with reports far from the induced belief, in the opposite half of the probability space. Total sample size is 391 participants and 2,554 decisions. For panel A, a nonlinear test of the difference in root of the squared-deviation, using paper-fixed effects, is significantly different ($p = 0.046$).

the classic quadratic-scoring rule (QSR), suggesting a higher frequency of truthful revelation under the former. In panel B of Figure 2, we further explore the average difference between the reported and induced beliefs in the two mechanisms. Despite a lower spread in the reports around the induced belief in the binarized-scoring rule, the data surprisingly indicate comparable average deviations and similar deviation patterns under the two mechanisms. Both elicitations show evidence of pull-to-center reporting, with positive deviations when the induced belief θ is less than 0.5 and negative deviations when it exceeds it. Overall, reports under both elicitations differ from the induced beliefs and do so in a manner that is likely to affect econometric inference from the elicited beliefs. While risk aversion only should affect deviations under the quadratic-scoring rule, we see center-biased reporting under both mechanisms, suggesting that neither mechanism is behaviorally incentive compatible.

Why Do Individuals Fail to Reveal the Induced Belief? Explanations and Remedies

In efforts to design better mechanisms, it is critical that we understand why a mechanism fails. While our results make clear that something in the classic and

binarized quadratic-scoring rule is malfunctioning, it is not clear what. Experimental techniques have been essential in exploring why individuals do not reveal their types. We offer a few examples to demonstrate the designs used for uncovering possible explanations. For more detail, see the excellent and comprehensive reviews by (Schlag, Tremewan, and van der Weele 2015; Schotter and Treviño 2014; Charness, Gneezy, and Rasocha 2021).

Initial assessments of what drives false reports were focused on understanding whether risk aversion affected deviations under the quadratic-scoring rule. Later investigations have moved to explore a broader set of causes and mechanisms. Three classic experimental-design techniques have been used to shed light on what drives deviations: (1) *design-by-correlation*, where an external measure of a potential driver is used to assess its correlation with the behavior of interest; (2) *design-by-manipulation*, where treatment variation is introduced that will attenuate/exacerbate the effect the driver has on the behavior of interest; or (3) *design-by-subtraction*, where a treatment removes the potential role for the driver of interest entirely, holding everything else constant.

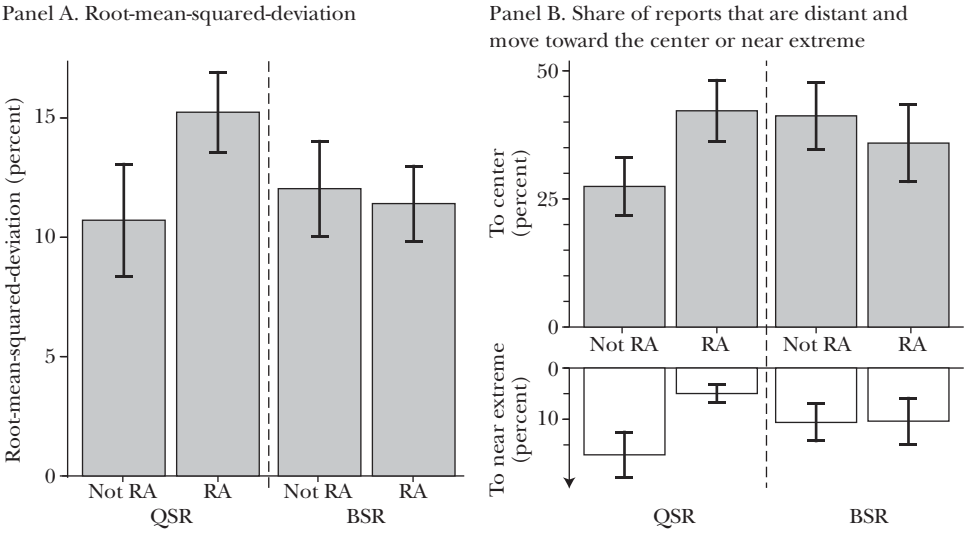
What would these design techniques look like in the context of evaluating whether risk aversion is causing reports to differ from the induced beliefs under the classic quadratic-scoring rule? Design-by-correlation would entail separately eliciting a measure of the participant's risk preference and determining whether it correlates with report deviations. In contrast, design-by-manipulation would explore treatment variations where risk aversion is predicted to further distort the deviations in particular ways, for example by comparing reports when we do and do not give participants an additional stake in the event (and a theoretical motive for a risk-averse individual to hedge). Finally, design-by-subtraction would introduce a treatment, where holding everything else constant the potential for risk aversion is removed.

Design-by-correlation hinges on securing an accurate external measure of the driver of interest (in this case, risk aversion) and a measure that is uncorrelated with other factors that may influence the behavior of interest (for example, confusion). Design-by-correlation is seen as the weaker of the three designs because it does not identify a causal relationship and because inference hinges on the quality of the external measure. Nonetheless it can offer insight. For example, we can elicit risk preferences by presenting participants with a lottery (say a 50 percent chance of winning $10) along with a list of certain payments ($1 to $10 in dollar increments) and ask that participants select the certain payments they prefer to the lottery. A participant indicating that they would prefer certain payments of $4 or more to the lottery would be categorized as risk averse, while a participant, who prefers the lottery unless the certain payment exceeds $6, would be categorized as risk-seeking. The correlation between risk aversion and misreporting can then serve as an indicator for whether center-biased reporting in the quadratic-scoring rule results from it not being incentive compatible for risk-averse individuals.

The data from Hossain and Okui (2013), Erkal, Gangadharan, and Koh (2020), and Danz, Vesterlund, and Wilson (2022) make possible a design-by-correlation

Figure 3

Distortion on Noncentered Induced Beliefs, by Elicitation and Risk Preference



Source: Figures based on the published data from binarized (BSR) and quadratic (QSR) scoring-rule elicitations in Hossain and Okui (2013); Erkal, Gangadharan, and Koh (2020); and Danz, Vesterlund, and Wilson (2022).

Note: All data use the Hossain and Okui definition of "betweenness" to exclude participants with reports far from the induced belief, in the opposite half of the probability space, and include only noncentered beliefs (outside of the 0.4 to 0.6 range). The sample includes 389 participants and 1,851 decisions. Inferentially, in panel A the differences in the RMSD between risk-averse (RA) and not risk-averse (not RA, so either risk-neutral/loving) participants is significant in the QSR (p = 0.046) but not in the BSR ($p = 0.625$). Similarly, in panel B there are significant differences in both movement directions across risk-preference for the QSR ($p = 0.001$ both center and near extreme) but not for the BSR ($p = 0.314$ and $p = 0.936$).

evaluation of the role played by risk aversion in reports under the standard and binarized quadratic-scoring rules. Focusing on noncentered induced beliefs where risk aversion is predicted to cause a distortion (induced beliefs θ outside of the central 0.4 to 0.6 range), we can assess if deviations in reports for risk-averse (RA) respondents are different from those who are not risk-averse (not RA, so risk-loving or risk-neutral).

The first two bars of Figure 3, panel A, show that under the quadratic-scoring rule the root-mean-squared-deviation is greater for the risk-averse participants, revealing a positive correlation between risk aversion and the size of the deviations. The next two bars show under the binarized-scoring rule no correlation between risk preferences and deviations. Taken in combination, the results are consistent with risk-aversion driving deviations under the quadratic-scoring rule.

We can use the same techniques to examine the interaction between risk attitudes and the direction of the deviations. Focusing on noncentered induced beliefs, panel B of Figure 3 shows the direction of the deviation by the belief elicitation

and the participants' risk preferences. The figure shows the direction of the distant reports, moving either towards the center (gray bars) or towards the near extreme (white bars).

The first two sets of bars show for the quadratic-scoring rule the predicted correlation with risk aversion: for risk-averse participants, 42 percent of reports are distant and move toward the center, while for not-risk-averse participants only 27 percent of reports are distant and distorted towards the center (and consistent with risk-seeking preferences, a significantly larger proportion make distant reports toward the near extreme). The next set of two bars show for the binarized-scoring rule that the participants' risk preferences do not correlate with the share of distant reports, neither toward the center nor the near extreme. Instead, independent of risk aversion we find that approximately 40 percent of reports are distant and towards the center and 10 percent are distant and towards the near extreme. In short, assessing the correlation between reported beliefs and participants' risk preferences suggests that risk preferences contribute to the rate of false reports under the quadratic-scoring rule.

Design-by-correlation has also been used to understand the effects of bounded rationality on distortions in belief reports. Burfurd and Wilkening (2022) use a measure of probabilistic sophistication and show that this measure of bounded rationality correlates with larger deviations. Enke and Graeber (2023) examine behavior in a belief-updating task with a shifting prior probability using a binarized-scoring rule. Using a measure of cognitive uncertainty, they assess the impact of bounded rationality on reporting and show that much of the non-Bayesian updating behavior is driven by cognitively-uncertain participants.

For an example of using design-by-manipulation to explain the deviations from induced beliefs, Armantier and Treich (2013) introduce experimental variation over: (1) the size of the incentives used in the quadratic-scoring rule (the maximal prize amount $\$X$), (2) the extent to which the participant has a financial stake in the event being elicited (a separate bonus payment if the elicited event happens), and (3) whether the participant could make a bet on the event being elicited, separate from the elicitation incentives. Relative to a control, these treatment manipulations are predicted to affect reports by risk-averse participants, but to have no effect on reports by those who are risk-neutral. For example, an increase in the size of the incentives should have no impact on reports by risk-neutral participants, while it should make centered reports relatively more attractive for risk-averse participants. Paying a bonus if the event E occurs makes it more attractive for risk-averse participants to report a lower belief, as the bonus decreases the ratio of marginal utilities for the payoff when the event occurs relative to the payoff when the event does not occur. Consistent with risk aversion impacting deviations, they find increased distortions in the reports for all three treatments, leading to the conclusion that risk-aversion contributes to the deviations seen under the quadratic-scoring rule.

Design-by-manipulation has also been used to explore other drivers of deviations. For example, Offerman and Palley (2016) use a manipulation of the classic quadratic-scoring rule. Specifically, they modify the payments to reduce the

distortions from loss aversion, where the core treatment variation increases payoffs in the unlikely state where relative losses occur. Consistent with loss aversion affecting deviations, they show that treatment variation reduces false reports (measured by the root-mean-squared-deviation and by the fraction of centered reports).

In an example of design-by-subtraction to explore drivers of reported deviations, Benoît, Dubra, and Romagnoli (2022) assess the role of participants' preference for events they control.[4] They use an elicitation over the respondent's confidence that they are above the median for performance on a task. However, the mechanism used (a mechanism called the probabilistic BDM, which we discuss further below) makes use of two payment arms: one with an exogenous lottery, and one with a lottery based on their performance. A posited channel for false reports is that participants prefer incentives based on realizations under their control, and so distort their beliefs upward. In a clever design-by-subtraction, Benoît, Dubra, and Romagnoli (2022) remove this feature by replacing the exogenous lottery arm with an equivalent incentive that is based on the respondent's performance. As such, the treatment holds constant the incentives, but removes the control motive. The comparison provides evidence that a preference for control is driving false reports, as reports exhibiting self-confidence decrease substantially in the treatment without the control motive.

To summarize, a range of experimental tests and designs have been used to explore why participants under a theoretically incentive-compatible mechanism fail to report their induced type truthfully. In these studies, much of the experimental focus has been the distortive effects of risk aversion under the quadratic-scoring rule. The literature has responded to these findings in one of two ways. One approach involves patching up the misfunctioning mechanism, by collecting additional behavioral measures and applying a correction to the reports. For example, Offerman et al. (2009) gather additional data on preferences and construct corrections to the reports for both risk preferences and ambiguity.[5] The other approach involves updating the mechanism to remove the distortions, as in developing elicitations that are incentive compatible for risk-averse individuals (for example, Hossain and Okui 2013; Benoît, Dubra, and Romagnoli 2022; Mobius et al. 2022).

[4] While it is tempting to see a comparison of the classic and binarized versions of the quadratic-scoring rule as design-by-subtraction, here we are not holding everything constant except risk aversion, as the entire incentive structure is also changing. While design-by-subtraction is seen as the gold standard for experimental design, it is also one of the more challenging design methods, when the driver we wish to identify is more abstract.

[5] While adding supplemental type information such as risk preferences has proved useful for belief elicitations, in more-general mechanisms the designer's goal will typically depend on these type features too. As such, we cannot use supplemental individual assessments of, say, risk and loss aversion to correct the reported types in auctions or other mechanism, as bids will depend on these features as well as the valuations. Uncovering the "true" type will impact the designer's action or interact with the bids of others. So these additional elements of type must be directly accounted for within the mechanism. Revelation and implementation require that the mechanism is incentive compatible for any outcome-relevant type (for example, information on risk preferences).

Direct Tests of Behavioral Incentive Compatibility

Indirect tests of behavioral incentive compatibility indicate when a mechanism malfunctions, but they do not tell us whether failure results from the mechanisms' incentives. To assess whether a mechanism is behaviorally incentive compatible, recent assessments instead look directly at how participants respond to the incentives of a mechanism and ask whether participants perceive them as intended (Danz, Vesterlund, and Wilson 2022).

We discuss two direct tests of behavioral incentive compatibility. The first, a powerful *incentives-only test*, presents participants with a pure choice over the incentives available under the mechanism and evaluates whether most participants select the presumed maximizer. The second, an *info/no-info test*, uses design-by-subtraction to evaluate whether participants are more likely to reveal their induced type truthfully when provided with clear quantitative information on the incentives.

Incentives-Only Test

The *incentives-only test* strips the mechanism of its belief-elicitation framing and presents participants with a choice over the available incentives, asking them to choose their preferred event-contingent payoffs. For example, participants are informed that their earnings depend on whether a red ball is drawn from an urn with red and blue balls where the share of red balls corresponds to an induced belief of θ. The test presents the incentives under the mechanism as pairs of event-contingent payoffs—a payoff if the ball is red, a payoff if the ball is blue—where each pair corresponds to the incentives from a report of q in the mechanism being tested.

Table 1 provides an example of an incentives-only test of the binarized-scoring rule. The eleven options (A through K) correspond to the event-contingent payoffs from each implied report q on the chance of a red ball being drawn, ranging from 0 to 100 percent in 10 percent increments. For example, suppose that participants are informed that the chance of drawing a red ball is $\theta = 0.2$ and are asked to select their preferred event-contingent payoff pair. For participants selecting choice A, the chance of winning \$8 is 0 percent if the selected ball is red and 100 percent if the ball is blue, so a 20 percent chance of \$0 and an 80 percent chance of \$8. Selecting choice B, the chance of winning \$8 is 19 percent if the selected ball is red, and 99 percent if the ball is blue, and so on. For the objective probability of $\theta = 0.2$ on red, participants will maximize their chance of winning \$8 if they select option C, where, as seen in the right-most column (not visible to participants), selecting C corresponds to reporting a belief of $q = 0.2$.

The incentives-only test shows whether participants see the intended (truthfully revealing) choice as maximizing—that is, whether they make a choice corresponding to $q = \theta$. While truthful revelation is predicted for a rational expected-utility-maximizing agent, deviations may result because of cognitive limitations or nonstandard preferences, and because deviations from the intended choice are relatively inexpensive. To see this, consider again the case where there is a $\theta = 0.2$ chance of drawing a red ball. With the theorized maximizing C choice,

Table 1
Incentives-Only Test: Payoffs Available under the Binarized-Scoring Rule

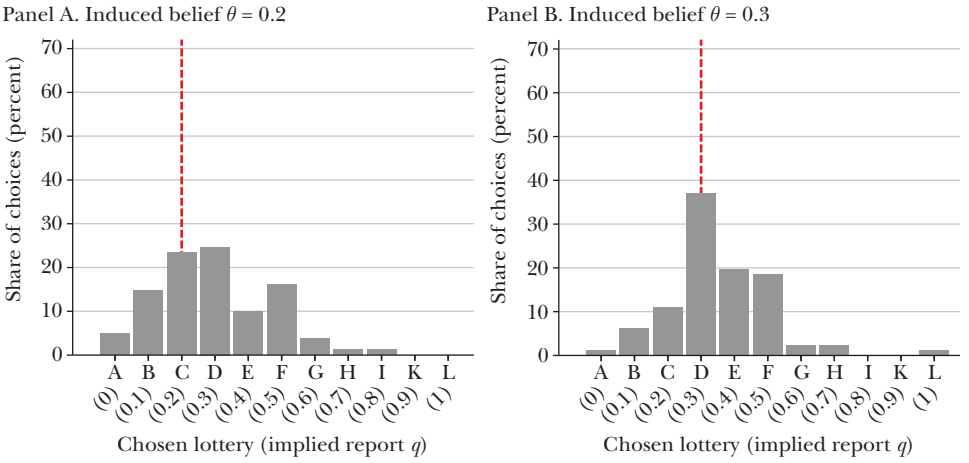
	Binarized scoring rule (BSR)		
	Chance of $8 prize by event		
Lottery option	*Red ball (Prob. θ)*	*Blue ball (Prob. 1 – θ)*	*Implied report q*
A	0%	100%	0.0
B	19%	99%	0.1
C	36%	96%	0.2
D	51%	91%	0.3
E	64%	84%	0.4
F	75%	75%	0.5
G	84%	64%	0.6
H	91%	51%	0.7
I	96%	36%	0.8
J	99%	19%	0.9
K	100%	0%	1.0

Source: Authors' creation.
Note: Participants are shown the menu of options under the binarized-scoring rule (BSR) and are asked to select their preferred option of event-contingent payoffs conditional on a θ chance that the ball is red. With the theorized maximizer under each elicitation being the option corresponding to $q = \theta$. The implied report q column (which is not shown to participants) indicates the report in the BSR to which this lottery incentive is matched.

the chance of winning $8 is 36 percent when the ball is red and 96 percent when blue. This compound lottery yields an 84 percent chance of winning $8, the largest total chance over the available options. However, a choice such as *D* (corresponding to a more-conservative report of $q = 0.3$) increases the chance of winning by 15 percentage points on red (from 36 percent to 51 percent) while decreasing the chance of winning by only 5 percentage points on blue (96 percent to 91 percent). By design, moving from choice *C* to *D* decreases the overall chance of winning, but note that the decrease is a mere one percentage point. The inexpensive deviation to *D* may therefore tempt individuals who prefer smaller differences in the chance of winning across the binary event outcome.

Figure 4 illustrates the results from an incentives-only test of the binarized-scoring rule for induced probabilities on a red ball of $\theta = 0.2$ or 0.3, respectively. Most participants choose event-contingent payoff options that differ from the assumed maximizer under the mechanism (shown by the vertical dashed line), showing directly that the incentives from the binarized-scoring rule are not behaviorally incentive compatible. Further, the test demonstrates the expected direction of deviations under the mechanism, in this case showing preferences for lottery pairs toward the center choice of F, consistent with the center-biased reporting seen in Figure 2, panel B, and Figure 3, panel B.

Figure 4
Chosen Options in the Incentives-Only Test of the Binarized-Scoring Rule



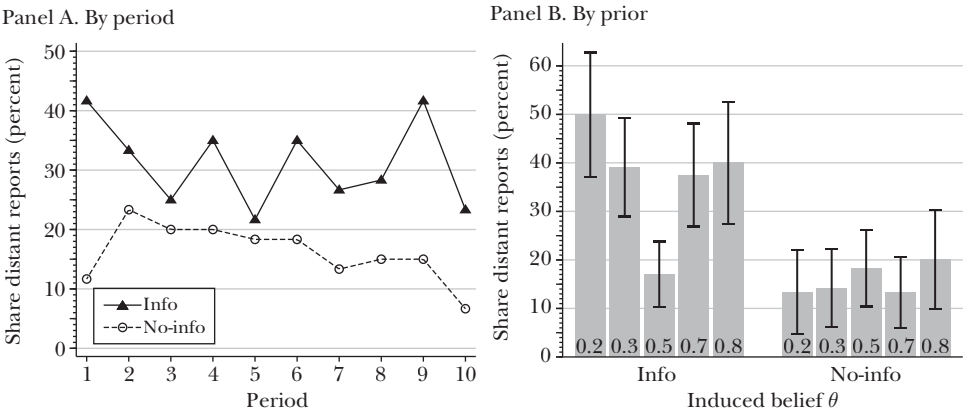
Source: Figure based on the published data from Danz, Vesterlund, and Wilson (2022, Figure 9).
Note: Figure shows distribution of participants' chosen lottery for induced beliefs of $\theta = 0.2$ (panel A) and $\theta = 0.3$ (panel B). The x-axis shows the lottery options (A–K) with corresponding implied belief reports (not shown to participants; 0–1).

Info/No-Info Test

With incentive-compatible belief elicitation, respondents should want to submit their most accurate belief after seeing the incentives. An *info/no-info test* can be used to assess how reports change when participants are given information on the incentives. Holding everything else constant, the test assesses as a minimal criterion for behavioral incentive compatibility whether knowing the offered incentives increases the likelihood that a respondent reveals their type.

The test uses two treatments: an *info* treatment with transparent quantitative information on the incentives, and a *no-info* treatment without the quantitative information on incentives. All other features are held constant. Participants in both treatments are given summary statements on the qualitative consequences of truthful reporting and the size of the stakes involved, $\$X$. The only difference is that participants in the info treatment also receive information on the precise quantitative incentives associated with any report under the mechanism. For example, participants in the no-info treatment for the binarized-scoring rule are only informed that "[t]he payment rule is designed so that you can secure the largest chance of winning the prize by reporting your most-accurate guess." Participants in the info treatment also received (1) a concise verbal description of how prize realizations were determined; (2) were shown the exact incentive for the provisionally selected belief report at the time of choice, and (3) were given feedback on the event outcomes and realized incentives at the end of each period.

Figure 5
Fraction Distant Reports in Info/No-Info Test of the Binarized Scoring Rule



Source: Figure based on the published data from Danz, Vesterlund, and Wilson (2022, Figure 4).
Note: Figure shows fraction of distant reports in the info and no-info treatments over time (panel A) and by induced belief (panel B). Distant reports are belief reports deviating by more than five percentage points from the induced belief.

Figure 5 illustrates the results from the info/no-info test of the binarized-scoring rule. The experiment was conducted over ten periods. At the start of each period, a simple belief was induced (based on probabilities of certain outcomes with a ten-sided die-roll), with the possibilities including 0.2, 0.3, 0.5, 0.7, and 0.8. Panel A of Figure 5 shows the rate of distant reports (those more than 5 percentage points from the induced belief) under the info and no-info treatments. Disturbingly, the rate of distant reports is substantially higher in the info than in the no-info treatment in every period of the experiment, revealing that participants are less likely to report the induced belief when they are presented with information on the quantitative incentives. Further, panel B of Figure 5 shows the rate of distant reports by treatment and for each induced belief. As evidence that incentives are distorting accurate reporting, we see that the rate of distant reports is independent of the induced belief in the no-info treatment (right-hand bars), but varies with the induced belief in the info treatment (left-hand bars), with distant reports being more likely for noncentered induced beliefs than for a centered belief of $\theta = 0.5$. Importantly, there is no evidence that risk aversion is the culprit for deviations under the info treatment, both because risk aversion theoretically should not play a role under the binarized-scoring rule, and because separately measured risk attitudes do not predict the likelihood of distant reports.

The no-info treatment demonstrates that participants have a reasonable understanding of the task at hand—as they report the induced beliefs at high rates in the absence of quantitative information on the incentives. Paradoxically, information on the incentives causes individuals to deviate from reporting their

true type, demonstrating that the binarized-scoring rule is not behaviorally incentive compatible.

Other Applications of Direct Tests

Direct tests of incentive compatibility have been applied to several other belief elicitation mechanisms. For example, Danz, Vesterlund, and Wilson (2024) find that results for the quadratic-scoring rule are similar to those for the binarized-scoring rule. An incentives-only test of the quadratic-scoring rule shows that the majority of participants prefer payoffs that differ from the intended maximizer, and that many prefer the incentives consistent with center-biased reporting, where there are smaller differences in event-contingent payoffs. An info/no-info test of the quadratic scoring rule shows that information on the quantitative incentives *increases* distant reports, a difference that is maintained throughout the experiment. Further, mirroring the results from the binarized-scoring rule, distant reports under the classic quadratic-scoring rule are only sensitive to the induced belief in the info treatment, and are far more likely for noncentered induced beliefs. That is, direct tests of the incentives reveal that the classic quadratic-scoring rule is not behaviorally incentive compatible, and that the incentives directly contribute to the false reports seen under the mechanism.

Danz, Vesterlund, and Wilson (2024) also explore the behavioral incentive compatibility of the *probabilistic Becker-DeGroot-Marschak mechanism* (Becker, DeGroot, and Marschak 1964; Karni 2009; Mobius et al. 2022; see also Smith 1961; Grether 1980), an increasingly popular elicitation. Similar to the binarized-scoring rule, the incentives are designed to be incentive compatible for arbitrary risk preferences and ensure that truthful revelation maximizes the chance of winning a fixed prize. Under the probabilistic Becker-DeGroot-Marschak (p-BDM) mechanism, the participant reports a belief q for, say, the share of red balls in the urn out of a total of 100. The payment depends on the reported belief, the event realization, and a randomly drawn number $z \in [0,1]$. If z is higher than the reported number q, the participant receives $\$X$ with probability z. If the draw z is less than the estimated value q, then the participant receives $\$X$ if the event E occurs. That is, for a reported belief q of event E, the participant receives $\$X$ with probability $q + \left(1 - q^2\right)/2$ if the event occurs and with probability $\left(1 - q^2\right)/2$ if the event does not occur. While truthfully revealing the induced belief maximizes the chance of winning, note that the offered incentives differ markedly from those under the binarized-scoring rule. From Table 1, under the binarized-scoring rule an event-independent probability of winning (of 75 percent) can be ensured by a centered report of $q = 0.5$. In contrast, under the probabilistic Becker-DeGroot-Marschak mechanism, an event-independent probability of winning (50 percent) can be ensured by an extreme report of $q = 0.0$.

Danz, Vesterlund, and Wilson (2024) show in an incentives-only test of the probabilistic Becker-DeGroot-Marschak mechanism that the vast majority of participants prefer choices that differ from the intended maximizer, indeed 69 percent of participants opt for the event-independent choice corresponding to reporting $q = 0.0$. Results from the info/no-info test further confirm that the probabilistic

Becker-DeGroot-Marschak mechanism is not behaviorally incentive compatible. Distant reports are more likely when participants are informed of the incentives under the mechanism, and consistent with the incentives-only test, reports are pulled toward $q = 0.0$. For example, at an induced belief of $\theta = 0.2$ only 7 percent of reports are both distant and towards zero in the no-info treatment. In contrast, this figure jumps to 21 percent of reports in the info-treatment (with no differences in the fraction of distant reports in the other direction).

To summarize, choices made under the incentives-only test for three commonly used belief elicitations reveal that the *majority* of participants *do not* prefer the theorized maximizing choice. Further, info/no-info tests show that providing participants with quantitative information on their incentives substantially increases the rate of false reports. That is, the incentives commonly used to encourage truthful revelation do not make it in the participant's "best interest to reveal their type," implying failures of behavioral incentive compatibility.

Conclusion

Economists have developed a range of mechanisms that are theoretically incentive compatible to provide participants with incentives to reveal their private type. Experimental economics has played a critical role in determining whether mechanisms are also behaviorally incentive compatible. The experimenter's ability to manipulate and induce an individual's type make it possible to determine whether the developed mechanism encourages truthful revelation. In reviewing the experimental techniques developed to assess behavioral incentive compatibility, we focus on the simple case of individual belief elicitation, showing both how indirect assessments can be performed within the mechanism, and how direct assessment can be done by directly evaluating the mechanism's incentives.

Applying the different experimental techniques to assess belief elicitations paints a dismal picture of the extent to which these encourage truthful revelation. Danz, Vesterlund, and Wilson (2024) show for the most-used belief elicitation mechanisms (the classic and binarized quadratic-scoring rule and the probabilistic Becker-DeGroot-Marschak rule) that participants largely prefer payoffs different from the intended maximizer under the mechanism, and that information on the incentives *increases* the rate of false reports.

The high rate of false reports has serious implications when using beliefs elicited under the mechanism. As an example, Danz, Vesterlund, and Wilson (2022) replicate the well-known Niederle and Vesterlund (2007) study on gender and competition. The original finding of Niederle and Vesterlund was that, conditional on performance, men enter competitions more than women, but that part of this difference was driven by men being more confident than women. Using an info/no-info comparison across the binarized scoring rule, Danz, Vesterlund, and Wilson (2022) elicit beliefs on relative performance for men and women. The no-info treatment replicates the prior finding that women are less confident about winning

a competition than men, and that controlling for beliefs reduces the gender gap in preferences for competition. In contrast, for the info treatment, the results do not uncover a gender gap in confidence and controlling for beliefs does not help explain the gender gap in preferences for competition. Providing clear information on the quantitative incentives shifts reported beliefs and changes inference. Both the original study and the no-info treatment lead to a conclusion that differences in confidence between men and women are important, and contribute to the gender gap in competition. In contrast, for the info treatment, the gender gap in competition is solely explained by preferences. These results outline the large ramifications from using an elicitation mechanism that is not behaviorally incentive compatible. Inferences drawn from biased reports will attenuate estimated treatment responses when beliefs are used as a dependent (left-hand-side) variable and bias all estimates when used as an explanatory (right-hand-side) variable.

While we have focused on the case of belief elicitation, indirect assessments of behavioral incentive compatibility have been used to evaluate a broad set of mechanisms, including auctions, centralized clearing houses, and so on (for example, Kagel, Harstad, and Levin 1987; Coppinger, Smith, and Titus 1980; Kagel et al. 1989; Chen and Sonmez 2006; Roth 2017). However, direct assessments can also be extended to such settings, offering simple diagnostic tests directly targeted at the mechanism incentives. Info/no-info tests can be used to determine whether clear information on the incentives increases truthful revelation, while the incentives-only test can be used to convert the effective incentives into stark decision problems by holding constant the theorized behavior of other participants and directly evaluating whether individuals prefer the assumed maximizer.

For example, Danz, Vesterlund, and Wilson (2024) use the pure-incentives test to assess the "deferred acceptance" mechanism that Boston, New York, and other cities use to assign students to schools and that is used nationally to match newly graduated doctors to residency programs. Stripping away the mechanism and the strategic features, which typically require many participants to submit rankings of their potential options, they find the vast majority of participants prefer the outcome associated with truthfully revealing their ranking. That is, the incentives under deferred acceptance are behaviorally incentive compatible, and failures in truthfully revealing preference rankings must result from other aspects of the mechanism. This insight is particularly helpful in light of the evidence that individuals, when faced with the mechanism, fail to reveal their type (as in Echenique et al. 2016; Dreyfuss, Heffetz, and Rabin 2022; Rees-Jones 2018, and this symposium). Results from the pure-incentives test demonstrate that these failures are not driven by the incentives per se, but by other aspects of the algorithm.

Where static mechanisms might fail, behavioral research has opened up other design channels for improving mechanism performance. For example, dynamic framings in which types are revealed through a sequence of simpler, starker decisions, can make the dominant choice more obvious and increase truthful reporting (along the lines of Li 2017, and this symposium). For example, Hao and Houser (2017) demonstrate a substantial increase in truthful reporting when they reframe

the probabilistic Becker-DeGroot-Marschak mechanism as a "clock auction"—that is, an auction with rounds of bidding where in each round participants reveal whether their belief is greater than the current clock value, rather than a declarative mechanism requiring a one-time report on q (see also Chapman et al. 2018; for impact of dynamic framing and more-careful instructions on deviations, see Healy 2017; Holt and Smith 2016).

Another approach—perhaps counterintuitive—is to provide less information on the mechanism's incentives. In the domain of belief elicitations, evidence of failed behavioral incentive compatibility has largely resulted in hiding the mechanism's incentives and instead providing participants with a summary statement of the incentives. For example, we may simply inform participants that truthful revelation maximizes the chance of winning an $8 prize (where statements on truthful revelation being in the participant's "interest" are deceptive given the pure incentives test). While this approach is tempting, we caution against it. If we are to incentivize truthful revelation, we recommend instead that the incentives provided be revised to encourage rather than discourage revelation. As part of this, it may be necessary to consider coarser mechanisms where simple and stark incentives are provided to secure truthful revelation. While this can reduce the precision of the provided reports, it may serve to reduce the hidden distortions in them, too. In developing and exploring new mechanisms, however, it is critical that attention be given to whether new candidates are behaviorally incentive compatible, and tests must be conducted to ensure that individuals see it as in their interest to truthfully reveal their type.

■ *We are grateful to Shengwu Li, Jonathan Parker, Nina Pavcnik, Alex Rees-Jones, Timothy Taylor, and Heidi Williams for invaluable comments and suggestions. We are also grateful to Garett Kohno, Priyoma Mustafi, and Brandon Williams for excellent research assistance with this paper.*

References

Armantier, Olivier, and Nicolas Treich. 2013. "Eliciting Beliefs: Proper Scoring Rules, Incentives, Stakes and Hedging." *European Economic Review* 62: 17–40.

Benjamin, Daniel J. 2019. "Errors in Probabilistic Reasoning and Judgment Biases." In *Handbook of Behavioral Economics: Applications and Foundations*, Vol. 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 69–186. Amsterdam: North-Holland.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Science* 9 (3): 226–32.

Benoıt, Jean-Pierre, Juan Dubra, and Giorgia Romagnoli. 2022. "Belief Elicitation When More than Money Matters: Controlling for 'Control.'" *American Economic Journal: Microeconomics* 14 (3): 837–88.

Brier, Glenn W. 1950. "Verification of Forecasts Expressed in Terms of Probability." *Monthly Weather Review* 78 (1): 1–3.

Burfurd, Ingrid, and Tom Wilkening. 2022. "Cognitive Heterogeneity and Complex Belief Elicitation." *Experimental Economics* 25 (2): 557–92.

Chapman, Jonathan, Erik Snowberg, Stephanie Wang, and Colin Camerer. 2018. "Loss Attitudes in the US Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)." NBER Working Paper 25072.

Charness, Gary, Uri Gneezy, and Vlastimil Rasocha. 2021. "Experimental Methods: Eliciting Beliefs." *Journal of Economic Behavior and Organization* 189: 234–56.

Charness, Gary, and Dan Levin. 2005. "When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect." *American Economic Review* 95 (4): 1300–1309.

Chen, Yan, and Tayfun Sönmez. 2006. "School Choice: An Experimental Study." *Journal of Economic Theory* 127 (1): 202–31.

Coppinger, Vicki M., Vernon L. Smith, and Jon A. Titus. 1980. "Incentives and Behavior in English, Dutch and Sealed-Bid Auctions." *Economic Inquiry* 18 (1): 1–22.

Cox, James C., Bruce Roberson, and Vernon L. Smith. 1982. "Theory and Behavior of Single Object Auctions." *Research in Experimental Economics* 2: 1–43.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2022. "Belief Elicitation and Behavioral Incentive Compatibility." *American Economic Review* 112 (9): 2851–83.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2024. "The Incentives-Only Test: Assessing Behavioral Incentive Compatibility in Mechanisms." Unpublished.

Danz, David, Lise Vesterlund, and Alistair J. Wilson. 2024. *Data and Code for: "Evaluating Behavioral Incentive Compatibility: Insights from Experiments."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E209063V1.

Dreyfuss, Bnaya, Ori Heffetz, and Matthew Rabin. 2022. "Expectations-Based Loss Aversion May Help Explain Seemingly Dominated Choices in Strategy-Proof Mechanisms." *American Economic Journal: Microeconomics* 14 (4): 515–55.

Echenique, Federico, Alistair J. Wilson, and Leeat Yariv. 2016. "Clearinghouses for Two-Sided Matching: An Experimental Study." *Quantitative Economics* 7 (2): 449–82.

Erkal, Nisvan, Lata Gangadharan, and Boon Han Koh. 2020. "Replication: Belief Elicitation with Quadratic and Binarized Scoring Rules." *Journal of Economic Psychology* 81: 102315.

Enke, Benjamin, and Thomas Graeber. 2023. "Cognitive Uncertainty." *Quarterly Journal of Economics* 138 (4): 2021–67.

Ewers, Mara, and Florian Zimmermann. 2015. "Image and Misreporting." *Journal of the European Economic Association* 13 (2): 363–80.

Gächter, Simon, and Elke Renner. 2010. "The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments." *Experimental Economics* 13 (3): 364–77.

Grether, David M. 1980. "Bayes Rule as a Descriptive Model: The Representativeness Heuristic." *Quarterly Journal of Economics* 95 (3): 537–57.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. "Designing Information Provision Experiments." *Journal of Economic Literature* 61 (1): 3–40.

Hakimov, Rustamdjan, and Dorothea Kübler. 2021. "Experiments on Centralized School Choice and College Admissions: A Survey." *Experimental Economics* 24 (2): 434–88.

Hao, Li, and Daniel Houser. 2012. "Belief Elicitation in the Presence of Naïve Respondents: An Experimental Study." *Journal of Risk and Uncertainty* 44: 161–80.

Harrison, Glenn W., Jimmy Martínez-Correa, and J. Todd Swarthout. 2014. "Eliciting Subjective Probabilities with Binary Lotteries." *Journal of Economic Behavior and Organization* 101: 128–40.

Healy, Paul J. 2017. "Epistemic Experiments: Utilities, Beliefs, and Irrational Play." Unpublished.

Hollard, Guillaume, Sébastien Massoni, and Jean-Christophe Vergnaud. 2016. "In Search of Good Probability Assessors: An Experimental Comparison of Elicitation Rules for Confidence Judgments." *Theory and Decision* 80: 363–87.

Holt, Charles A., and Angela M. Smith. 2016. "Belief Elicitation with a Synchronized Lottery Choice Menu That Is Invariant to Risk Attitudes." *American Economic Journal: Microeconomics* 8 (1): 110–39.

Hossain, Tanjim, and Ryo Okui. 2013. "The Binarized Scoring Rule." *Review of Economic Studies* 80 (3): 984–1001.

Hurwicz, Leonid. 1972. "On Informationally Decentralized Systems." In *Decision and Organization: A*

Volume in Honor of J. Marschak, edited by C. B. McGuire and Roy Radner, 297–334. Amsterdam: North-Holland.

Hurwicz, Leonid. 1973. "The Design of Mechanisms for Resource Allocation." *American Economic Review* 63 (2): 1–30.

Kagel, John H., Ronald M. Harstad, and Dan Levin. 1987. "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." *Econometrica* 55 (6): 1275–1304.

Kagel, John H., and Dan Levin. 1993. "Independent Private Value Auctions: Bidder Behaviour in First-, Second- and Third-Price Auctions with Varying Numbers of Bidders." *Economic Journal* 103 (419): 868–79.

Kagel, John H., Dan Levin, Raymond C. Battalio, and Donald J. Meyer. 1989. "First-Price Common Value Auctions: Bidder Behavior and the 'Winner's Curse.'" *Economic Inquiry* 27 (2): 241–58.

Kagel, John H., and Alvin E. Roth. 2000. "The Dynamics of Reorganization in Matching Markets: A Laboratory Experiment Motivated by a Natural Experiment." *Quarterly Journal of Economics* 115 (1): 201–35.

Karni, Edi. 2009. "A Mechanism for Eliciting Probabilities." *Econometrica* 77 (2): 603–06.

Kessler, Judd B., and Lise Vesterlund. 2015. "The External Validity of Laboratory Experiments: The Misleading Emphasis on Quantitative Effect." In *Handbook of Experimental Economic Methodology*, edited by Guillaume R. Fréchette and Andrew Schotter, 391–406. Oxford: Oxford University Press.

Köszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121 (4): 1133–65.

Li, Shengwu. 2017. "Obviously Strategy-Proof Mechanisms." *American Economic Review* 107 (11): 3257–87.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–76.

Möbius, Markus M., Muriel Niederle, Paul Niehaus, and Tanya S. Rosenblat. 2022. "Managing Self-Confidence: Theory and Experimental Evidence." *Management Science* 68 (11): 7793–7817.

Nelson, Robert G., and David A. Bessler. 1989. "Subjective Probabilities and Scoring Rules: Experimental Evidence." *American Journal of Agricultural Economics* 71 (2): 363–69.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122 (3): 1067–1101.

Nyarko, Yaw, and Andrew Schotter. 2002. "An Experimental Study of Belief Learning Using Elicited Beliefs." *Econometrica* 70 (3): 971–1005.

Offerman, Theo, and Asa B. Palley. 2016. "Lossed in Translation: An Off-the-Shelf Method to Recover Probabilistic Beliefs from Loss-Averse Agents." *Experimental Economics* 19: 1–30.

Offerman, Theo, Joep Sonnemans, Gijs Van de Kuilen, and Peter P. Wakker. 2009. "A Truth Serum for Non-Bayesians: Correcting Proper Scoring Rules for Risk Attitudes." *Review of Economic Studies* 76 (4): 1461–89.

Palfrey, Thomas R., and Stephanie W. Wang. 2009. "On Eliciting Beliefs in Strategic Games." *Journal of Economic Behavior and Organization* 71 (2): 98–109.

Rees-Jones, Alex. 2018. "Suboptimal Behavior in Strategy-Proof Mechanisms: Evidence from the Residency Match." *Games and Economic Behavior* 108: 317–30.

Roth, Alvin E. 2017. "Experiments in Market Design." In *Handbook of Experimental Economics*, Vol. 2, edited by John H. Kagel and Alvin E. Roth, 290–346. Princeton: Princeton University Press.

Schlag, Karl H., James Tremewan, and Joël van der Weele. 2015. "A Penny for Your Thoughts: A Survey of Methods for Eliciting Beliefs." *Experimental Economics* 18: 457–90.

Schotter, Andrew, and Isabel Treviño. 2014. "Belief Elicitation in the Laboratory." *Annual Review of Economics* 6: 103–28.

Smith, Cedric A. B. 1961. "Consistency in Statistical Inference and Decision." *Journal of the Royal Statistical Society: Series B (Methodological)* 23 (1): 1–25.

Thoma, Carmen. 2016. "Under- versus Overconfidence: An Experiment on How Others Perceive a Biased Self-Assessment." *Experimental Economics* 19 (1): 218–39.

Trautmann, Stefan T., and Gijs van de Kuilen. 2015. "Belief Elicitation: A Horse Race among Truth Serums." *Economic Journal* 125 (589): 2116–35.

Winkler, Robert L., and Allan H. Murphy. 1970. "Nonlinear Utility and the Probability Score." *Journal of Applied Meteorology and Climatology* 9 (1): 143–48.

Wang, Stephanie W. 2011. "Incentive Effects: The Case of Belief Elicitation from Individuals in Groups." *Economics Letters* 111 (1): 30–33.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 155–174

Behavioral Incentive Compatibility and Empirically Informed Welfare Analysis: An Introductory Guide

Alex Rees-Jones

Consider an economist seeking to compare two economic situations and assess which is better for society. To pursue this goal, she first specifies what "being better for society" means. Formally, this entails quantifying the overall good or social welfare attained given the different possible allocations that could arise. She then forecasts the patterns of behavior that each situation will generate. By evaluating social welfare at the allocations that follow from these forecasted patterns of behavior, this economist now has what she needs to compare the social value of one option versus the other.

The approach just described succinctly captures economists' dominant paradigm for welfare analysis. Stated at this level of generality, it in some ways appears simple and straightforward: one just needs to specify how welfare will be defined and measured and then forecast individual behavior. Of course, forecasting the behavior of humans is challenging. At the same time, an enormous amount of economic research has been conducted with the explicit purpose of informing this stage of the modeling process, providing extensive foundations from which to build.

When generating the needed forecasts of individual behavior, the favored approach in economics is to assume that behavior will satisfy *incentive compatibility*. Put most simply, this means that researchers assume that individuals behave in the

■ *Alex Rees-Jones is an Associate Professor of Business Economics and Public Policy, Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania. He is also a Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. His email address is alre@wharton.upenn.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.155.

manner that best pursues their own interests. This approach is favored for good reason: it is extremely powerful. The assumption that individuals choose optimally immediately makes available all of the standard economic tools of "revealed preference." These tools provide well-developed means of estimating models of individual welfare, and those models can be used to determine what individuals will then choose. This framework thus provides the needed forecasts of the behavior that will be chosen in different situations.

Despite the power of this approach, reliance on incentive compatibility has a clear lack of appeal in certain settings. For many economic questions, individuals' optimization failures are central to the debate and thus cannot be ignored. For other economic questions, individuals' optimization failures might not be central to the existing debate, but incidental misoptimization may change the conclusion of that debate. In such cases, reducing our reliance on incentive compatibility may help us better analyze the economic environment and guide us towards better policies.

Motivated by these considerations, researchers increasingly conduct analysis that may be characterized as relying on *behavioral incentive compatibility*. Under this approach, behavior is forecasted not by assuming that individuals maximize welfare. Instead, the researcher attempts to model both the individuals' welfare and also the forces that guide them towards unwise decisions. These forces at times include psychological factors, incorrect beliefs about aspects of the decision problem, or preferences for things that are judged as normatively irrelevant and thus excluded from a standard welfare consideration. Despite the addition of these factors, the spirit of this exercise is extremely close to that driven by standard incentive compatibility. Just as in the standard case, this approach is based on assuming that individuals' decisions are compatible with their pursuit of incentives. This approach merely embeds some imperfection in their means of that pursuit, often drawing from work in behavioral economics.

In this article, I aim to introduce readers to empirically-informed welfare analysis based on behavioral incentive compatibility and to provide guidance for how to pursue a project involving such analysis. My interest in doing this comes from my experience having written several papers of this variety, actively engaged with this literature through most of its recent rise in prominence, and advised a number of students in their pursuit of this style of project. Having watched the literature evolve through that lens, two things stand out to me.

First, the potential value of this approach no longer needs to be taken on faith, but instead can be inferred from existing literature. Projects are being executed that address important economic questions, do so up to high standards of rigor, and ultimately have influence in diverse literatures. At least in some fields, I believe the approach has demonstrably grown beyond being "something popular with behavioral economists" and into something used, when appropriate, by standard members of the field.

Second, despite that success, there is an unfortunate hurdle that I believe has persistently slowed progress. Different fields have different core behavioral concerns, playing out in potentially very different economic environments. This naturally

contributes to a sense that solutions and approaches must be context-specific. Behavioral economics is often criticized for providing too many ad hoc theories instead of a unified framework that can immediately be brought to new settings; I believe this contributes to a common sensation that welfare analysis informed by behavioral economics would also be ad hoc. Yet, looking across successful examples of this type of research, it appears that the common practices for pursuing these projects are ultimately very similar across the subfields that have adopted them, and that there is an underappreciated degree of commonality in the template that is followed. I will seek to make the main elements of it clear, in the hopes of helping to make the pursuit of these projects less daunting.

Three Examples of Welfare Analysis with Behavioral Incentive Compatibility

To begin, I present three examples of projects where welfare assessments hinge critically on applications of behavioral incentive compatibility. These projects make concrete some of the issues just discussed—for example, they illustrate some types of analysis that can benefit from this approach; they illustrate how application of the approach can significantly change the conclusions we reach; and they illustrate that this approach has found traction across a range of fields with very different methods, settings, and interests.

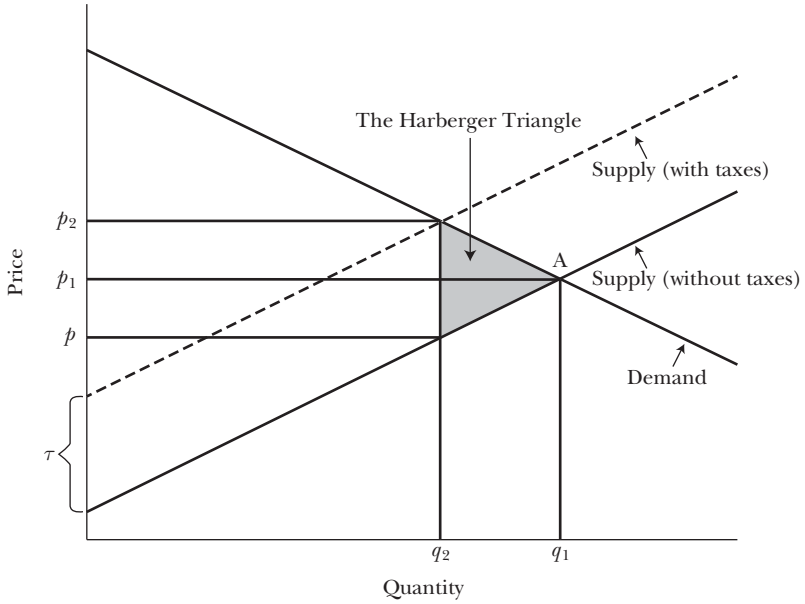
Behavioral Incentive Compatibility and Sales Taxes

We first review the pioneering work of Chetty, Looney, and Kroft (2009), in which the authors consider a classic question: How to calculate the welfare losses from sales taxes?

A common approach to answering this question is to use the "Harberger triangle" approximation (Harberger 1964), which fundamentally relies on an appeal to incentive compatibility. To illustrate, consider a standard supply and demand framework as represented in Figure 1. If we assume that all purchase decisions are incentive compatible, the demand curve serves two important functions in welfare analysis.[1] First, the demand curve provides a direct measure of consumer welfare. The difference between the willingness to pay encoded in the demand curve and the amount actually paid (that is, "consumer surplus") is a natural money-metric measure of the consumer benefits arising from the trade. Second, the demand curve allows us to infer what purchases will be made in counterfactual situations, such as when considering a new tax to be imposed in a previously untaxed market. Figure 1 illustrates the case where the new tax, τ, is imposed on the supply side and thus shifts the supply curve upward. This raises the equilibrium tax-inclusive price

[1] The assumption that all sales decisions are rational serves an analogous role for forecasting welfare and behavior of the supply side. I focus on the demand side here as it is the focus of Chetty, Looney, and Kroft (2009).

Figure 1
The Harberger Triangle



Source: Reproduced from Hines (1999).
Note: This figure presents a standard demonstration of Harberger triangle analysis. In this demonstration, a tax of size τ is introduced on the supply side of the market. The Harberger triangle is represented in the shaded region and captures the lost surplus from the trades that were eliminated by the post-tax price increase.

from p_1 to p_2, which rationally dissuades consumers who have willingness to pay between p_1 and p_2 from purchasing the good. The consumer surplus lost by these dissuaded consumers, along with the producer surplus lost by the producers who no longer trade with those consumers, is the "excess burden" or welfare loss from the imposition of this tax. It is represented in the shaded triangle in Figure 1.

Harberger-style analysis has been used extensively in economics and is clearly valuable. However, Chetty, Looney, and Kroft consider a specific reason why this analysis might be incomplete and why there might be room for improvement: consumers may not optimally react to taxes when the taxes are not salient. Consider, for example, a consumer who selects which groceries to purchase based on their price tags. In many parts of the world, these price tags would report the amount of money that must be paid to take ownership of the goods. In a US store, by contrast, price tags typically exclude sales taxes, which are then later imposed at the register. This labelling can naturally be expected to lead to mistakes if some consumers do not know sales tax rates, or do not know that some groceries are taxable and others are not, or know that there are taxable and untaxable groceries but do not know which are which, or know all of this but forget to attend to it, or remember

to attend to all of this but make mistakes in calculations, or do some rounding along the way, or correctly process everything but only notice changes in taxes slowly, or are able to correctly process everything but deem doing so not worth their time, and so on.

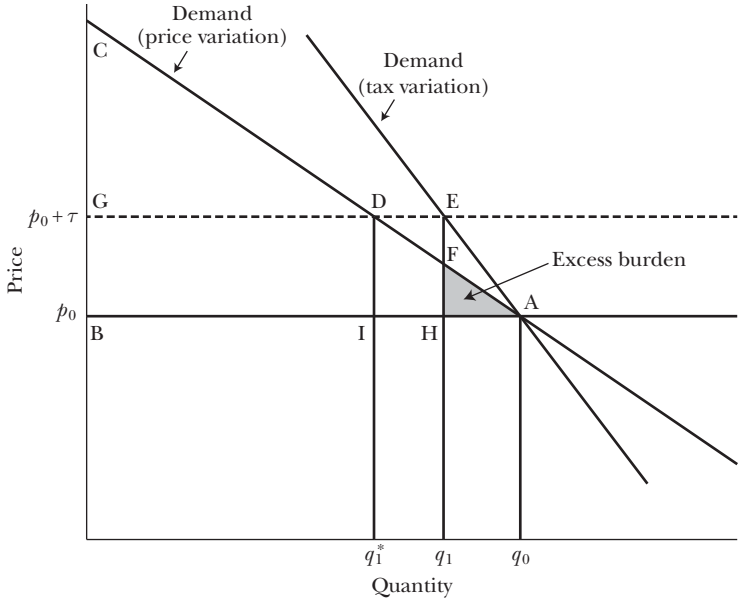
If consumers are failing to attend correctly to sales taxes collected at the register, what are the consequences for behavior? A natural consequence would be insensitivity, or lack of elasticity, to the sales tax amount, even in cases where the consumer would be sensitive to exactly comparable changes in the price advertised on the price tag. Chetty, Looney, and Kroft provide two empirical exercises—one field experiment and one observational study—that each directly demonstrate underreaction to taxes and allow for estimation of a parameter that governs the resulting reduction of elasticity.

Bringing these findings and observations together, Chetty, Looney, and Kroft demonstrate how nonsalient taxes can be accommodated in Harberger triangle calculations. Their modification may be understood as replacing the prior reliance on incentive compatibility with reliance on behavioral incentive compatibility. Formally, they model individuals as still making nearly rational decisions by assessing if the value of a good exceeds its price. However, they assume that only a portion of the tax is accounted for when price is calculated, rendering the overall decision rational except for a price misperception. The resulting demand can be estimated by examining purchase decisions as taxes vary, and may be used to forecast how demand behavior will change as taxes are changed. However, unlike in the standard case, this demand curve no longer is assumed to reveal welfare, because it is influenced by mistakes. A demand curve that is not influenced by mistakes can be estimated by examining how demand responds to variation in posted prices, which Chetty, Looney, and Kroft assume are processed correctly.

Figure 2 summarizes this analysis. Consider a market that would be at point A in the absence of a tax. From this point, Chetty, Looney, and Kroft draw two demand curves. The steep one represents demand arising from (nonsalient) tax variation. The shallower demand curve represents demand arising from price variation. The difference between these demand curves reflects the empirical finding that the quantity demanded will change more in response to a change in salient price than a comparable change in nonsalient tax. With these two curves graphed, Chetty, Looney, and Kroft then assess surplus or welfare by making use of the demand curve arising from price variation while still assessing predicted behavioral changes from the demand curve arising from tax variation.

To illustrate the calculation of the Harberger triangle, imagine a small tax τ were imposed on the economy in equilibrium A. Assume this economy faces a flat supply curve, as Chetty, Looney, and Kroft assume to focus attention on the demand side. In this case, if consumers responded to the new tax using the welfare-relevant demand curve (as arises from price variation), the new equilibrium would occur at point D. The standard Harberger triangle would be AID. The key observation of Chetty, Looney, and Kroft is that this calculation would overestimate the demand response of consumers by not accounting for their propensity to underreact to the

Figure 2
A Harberger Triangle When Taxes Are Non-Salient



Source: Figure 4 from Chetty, Looney, and Kroft (2009) (with modifications to labeling).
Note: This figure presents analysis of the consequences of introducing a tax of size τ to a market in which consumers react to tax salience. In this analysis, consumers are assumed to respond optimally to variation in posted prices, and thus the demand relationship from price variation can be used to infer welfare. Consumers are not assumed to respond optimally to variation in nonsalient taxes, leading a separate demand relationship to arise from tax variation.

nonsalient tax. Taking that into account would require using the other demand curve, which suggests equilibrium would occur at point E. If the demand curve with tax variation were believed to be welfare-relevant, this would suggest that the Harberger triangle should be AHE. However, the demand curve with tax variation is not welfare-relevant, and the demand curve for price variation must be used for the calculation of surplus. As a result, triangle AHF provides the desired estimate of excess burden: it uses the demand curve from tax variation to determine the quantity demanded in equilibrium q_1, but assesses lost consumer surplus by integrating the price-based demand curve between that new quantity demanded and the initial quantity demanded q_0.

In applying these results, this analysis yields a perhaps surprising conclusion. Despite the common intuition that failure to optimize is harmful, this analysis shows that welfare losses stemming from taxes are *reduced* when the taxes are nonsalient.

This arises because some surplus-reducing but individually-rational decisions to quit purchasing the good are not made.[2]

Behavioral Incentive Compatibility and Health Insurance

A second application of behavioral incentive compatibility for welfare analysis can be found in the work of Handel (2013). In this paper, Handel assesses a classic topic in health insurance markets: how they operate in the presence of adverse selection.

To illustrate the issue of adverse selection, consider a population of risk-neutral individuals buying a health insurance plan. If purchased, this plan will cover all healthcare costs. The operator of the health insurance plan knows the average cost of providing healthcare and offers a plan at that cost. Further assume that individuals decide whether to purchase the plan in an incentive-compatible way: they buy the plan if their expected costs of healthcare are higher than the cost. This understandable behavior leads to an unfortunate market dynamic. The insurer will soon find that, while the plan was priced appropriately for the average person in the population, the plan is not priced appropriately for the individuals who purchased the plan. These customers have been selected for the adverse trait of having higher-than-average health costs. The insurer must raise prices to prevent operating at a loss. This leads to further selection by "pricing out" even more customers, perhaps making further price increases necessary. Repeated rounds of this repricing can lead a large fraction of the populace to rationally remain uninsured due to the unavailability of an acceptably priced insurance product—a phenomenon often referred to as a "death spiral."

Adverse selection is a phenomenon that is driven by an unfortunate pattern of incentive-compatible behavior. Handel is partially motivated by the observation that, for health insurance choice, idealized incentive compatibility might fail. To illustrate again with an example, consider a worker trying to pick the best health plan of the five offered by her employer. These plans often have differences in deductibles, copays, coinsurance rates, coverage, and more. Thus, at a minimum, determining the optimal choice requires some understanding of how these provisions operate and interact. Furthermore, the consequences of these various provisions must be assessed across a large number of health situations this employee might face. The employee must assess her optimal plan if she is healthy all year, her optimal plan if she develops a specific rare illness, her optimal plan in a large number of other health contingencies, and the likelihoods of all of these different contingencies. Given these challenges, one could imagine that this worker might avoid making a serious attempt to determine the optimal plan due to its perceived futility, or that

[2]Follow-up papers have demonstrated that this simple conclusion might not hold when the method of welfare analysis is enriched in certain ways (for example, see Goldin and Homonoff 2013; Reck 2016; Taubinsky and Rees-Jones 2018). However, even these follow-up papers adopt the behavioral incentive compatibility approach when assessing welfare and merely debate the specification of some model components.

she might incorrectly select the optimal plan even if she tries to determine it. In such cases, this worker will of course be individually worse off. But how does the aggregation of these mistakes affect the market as a whole, and adverse selection itself?

To study these questions, Handel studies imperfect choice of health care plans in a large US firm, where workers must make decisions somewhat like the example just considered. In his data, he directly documents a low propensity to change plans across time and provides some suggestion that this market is insufficiently active relative to a rational benchmark. But much more strikingly, he studies a case where a plan became formally dominated—that is, another plan became preferable to this plan no matter the health contingency that would arise. Remaining in the dominated plan is not incentive compatible, and yet many employees failed to abandon it. These features together provide a compelling demonstration of some degree of consumer inertia.

To model inertia, Handel augments an otherwise-rational model of insurance demand to include an "as-if" switching cost. In standard models in this environment, individuals will switch to a new insurance plan if the plan offers infinitesimally better terms than its best competitor. In Handel's model, consumers act as if they will only switch from their plan when the returns to doing so are sufficiently large. Handel's estimated model suggests that benefits of switching plans must be valued above approximately $2,000 to motivate a switch. Of course, switching plans does entail some time and effort, so some degree of switching cost can be rationalized. But it is hard to rationalize a switching cost that is so large. This supports treating an individual's reliance on this switching cost as a mistake, and supports the treatment of the switching cost as an element of estimated utility that should be excluded from welfare. Use of the model in this way serves as the centerpiece of Handel's application of behavioral incentive compatibility.

Handel uses this estimated model to assess the welfare effects of consumer misoptimization in this market. To do so, he evaluates the effect of reducing inertia by scaling down switching costs. As a baseline analysis, Handel considers this change while holding plan pricing fixed (and thus preventing the consequences of adverse selection from playing out). In this analysis, reducing inertia leads to improved sorting of individuals to their individually rational policies, resulting in a substantial improvement to consumer welfare. This accords with common intuitions that helping individuals avoid mistakes helps their welfare, all else equal, perhaps suggesting that "nudges" to combat inertia would be useful.

However, a quite different conclusion arises once the impacts of adverse selection are reintroduced to the model. When plans are allowed to endogenously reprice their products as consumer demand changes, Handel finds that reducing inertia exacerbates adverse selection. As individuals sort to new plans once inertia is reduced, some plans are effectively removed from the market due to losing their lower-cost customers who previously stayed in the plan due to inertia. As such individuals are lost, prices rise, leading to further re-sorting. The end result is a microcosm of a death spiral that drives substantial declines in overall welfare.

A simple takeaway from this paper—only assessable through application of behavioral incentive compatibility—is that inertia and consumer misoptimization may at times play an important role in keeping health insurance markets functional in the presence of potentially debilitating adverse selection. In markets with adverse selection, "the problem" is generally that individuals make their optimal choices based on private information (in this case about health costs). If a behavioral force like inertia prevents them from doing so, this can at times be helpful for overall welfare, even if the behavioral forces come with welfare losses of their own.

Behavioral Incentive Compatibility in School-Choice Market Design

A third example of behavioral incentive compatibility in welfare analysis can be found in the work of Kapor, Neilson, and Zimmerman (2020). This paper assesses some much-studied questions in market design: how should we assign students to schools, and should we favor the "immediate acceptance" or the "deferred acceptance" algorithm?

When determining students' assignments to schools, a large and growing number of school systems use a formal centralized matching system. In such a system, both students and schools are asked to submit their preferences for assignments. For a student, this could be an indication of their favorite school to attend, their second favorite school, and so on until their last acceptable school. For a school, this could be an indication of their favorite student to admit, their second favorite, and so on until their last acceptable student. Schools additionally report how many seats they have available. Once this information is submitted, the school district can use it to determine a desirable way to assign students to schools.

Incentive compatibility plays a crucial role in assessing these procedures. Typical analysis assumes that students rank schools while rationally responding to any strategic incentives introduced by the procedure. This practice is clearly important because students can often face strong incentives *not* to report their true preferences.

To illustrate this potential for incentives to misreport preferences, imagine that assignments are determined by the following procedure. First, the school district tries to assign each student to the school she ranked first. If the school said that the student is unacceptable, or if the school is already filled to capacity with other applicants that the school prefers, the student is not assigned a seat. Otherwise, the student gets a seat at the school. Those assignments are treated as final and each schools' capacity is updated to reflect the seats that have been removed from the market. In the next step, the school district repeats this procedure, now trying to match students who did not match to their first-choice school to the remaining seats at their second-choice school. The procedure continues iterating in this way, moving down the students' preference lists, until all students are matched or every student has attempted to match at every school that they ranked.

The procedure just described is famous within the school-choice literature for producing unfortunate incentives, and is called the immediate acceptance mechanism or the Boston mechanism. To illustrate the incentive problems, consider two

schools, A and B. Both are very popular: they can fill all of their seats with students who ranked them first. Now consider a student who prefers A to B. Fortunately, school B ranks this student very highly. Unfortunately, school A does not. In such a case, this student could be matched to school B if she ranks it first. However, if she ranks school A first and school B second, she will not match to school A in the first stage of the procedure and will have no remaining seats available at school B in the second stage. This student thus faces clear incentives not to list school A: doing so would cost her the chance to study at school B. Generalizing beyond this simple example, this procedure offers strong incentives not to rank options where a match is unlikely, and generally punishes sincere participants to the benefit of the strategically savvy (Pathak and Sönmez 2008).

Avoiding this incentive problem is one of several reasons why economists have favored the use of the deferred acceptance mechanism of Gale and Shapley (1962). This mechanism may be understood as a modification to immediate acceptance that does not remove filled seats after each round, but instead allows more-preferred new applicants to displace previous matches. This eliminates the problem discussed in the example above, where a desired applicant is only considered after seats have been irrevocably claimed by students who ranked the school higher. Under deferred acceptance, these claims are no longer irrevocable. This mechanism structure results in deferred acceptance being strategy-proof: regardless of the behavior of other market participants, students can do no better than truthfully reporting their preferences (Dubins and Freedman 1981; Roth 1982). For this reason and others, deferred acceptance has largely served as the tool of choice for school-choice market designers in recent decades. (For much fuller detail on the use of these mechanisms for school assignment, a useful starting point is Abdulkadiroğlu and Sönmez 2003.)

The contrast between these two mechanisms may suggest that the choice between them is obvious: use of deferred acceptance, where students can report their preferences truthfully, seems wise compared to use of immediate acceptance, where strategic behavior is necessary and sincerity is punished. One provocative counterpoint to this comparison comes from Abdulkadiroğlu, Che, and Yasuda (2011), who note that immediate acceptance can, under some conditions, extract cardinal preference information that can lead to a higher-welfare match. This could potentially lead a market designer to prefer immediate acceptance despite its incentive properties.

To illustrate the issue, consider two students vying for two positions at schools, again labeled A and B. Say the two students both rank position A over position B. Despite that symmetry in rankings, there can be significant asymmetry in the welfare consequences of assignments. For example, if one student has essentially the same welfare at each school, whereas the other student is vastly better off at school A than school B, there could be strong welfare motives for saving the seat at A for the student who benefits from it more. The operation of deferred acceptance has no feature that pushes for this outcome. By contrast, the optimal reporting strategy for immediate acceptance is a function of cardinal utility differences and can at times lead to welfare gains by guiding assignments with that information.

The discussion up until now explains the state of the literature at the time Kapor, Neilson, and Zimmerman entered. To summarize, in this literature, deferred acceptance was broadly preferred to immediate acceptance as a means of matching students to schools. However, some theoretical considerations suggested that immediate acceptance might have welfare benefits. The models that lead to these conclusions rely on students optimally strategizing about their preference submission, taking into account their probabilities of matching to different schools. But what if students and their families don't know these probabilities, or have systematically biased beliefs? This motivates Kapor, Neilson, and Zimmerman's central question: are the theoretical benefits of immediate acceptance "worth it" when failures of probability estimates are taken into account?

Kapor, Neilson, and Zimmerman address this question using data from the New Haven Public School System. During the window of study, New Haven based school assignments on a procedure that was essentially identical to immediate acceptance. Kapor, Neilson, and Zimmerman secured access to administrative data, thus giving them access to the reported preferences that are used by the algorithm to determine the match. Such data are extremely valuable for the pursuit of a standard study of a school choice mechanism. Despite being valuable, they are insufficient for Kapor, Neilson, and Zimmerman's purposes, because they do not directly reveal the (possibly incorrect) beliefs about admissions probabilities that families hold. To address this data need, they also fielded a survey among participants in this match. While the survey served several purposes, its key function was to elicit families' beliefs about admissions probabilities with different possible preference submissions. They use these data to document substantial inaccuracy in families' probabilistic beliefs.

These findings illustrate a potential need to import a behavioral incentive compatibility notion into welfare inferences for this setting. To estimate preferences and assess welfare in a setting like this, the current standard approach is to assume that the preferences that were submitted maximize expected utility (as in Agarwal and Somaini 2018). This provides revealed-preference valuations of the different schools, which may be used to measure the welfare of a given assignment. Kapor, Neilson, and Zimmerman instead assume that the rank-ordered lists that were submitted maximize expected utility *conditional on the model of incorrect perceptions of match probabilities.*

Assessing total welfare with both approaches, a striking pattern emerges. When relying on standard incentive compatibility, Kapor, Neilson, and Zimmerman find that immediate acceptance outperforms deferred acceptance. This, viewed in isolation, would be a provocative finding: the widespread preference for deferred acceptance on the grounds of its avoidance of strategic incentives might be reducing welfare. This provocative finding is immediately reversed when considering the analysis based on behavioral incentive compatibility: once analysis accounts for families' difficulty in assessing admissions probabilities, the benefits of immediate acceptance decline. Deferred acceptance then preserves its status as the favored mechanism. This serves as an example of a case where reliance on standard incentive compatibility might lead to an unwise policy decision, and one that would be avoided by taking into account additional behavioral considerations.

Guidance for Welfare Analysis Based on Behavioral Incentive Compatibility

The three examples just considered contain welfare analyses of quite different economic questions drawn from quite different economic fields. Despite the different foundations of each of these analyses, there are clear similarities in their manner of execution. While I have discussed only three examples, I believe this similarity to be reflective of a broader phenomenon. In my observation, successful welfare analyses using the behavioral incentive compatibility approach tend to draw upon a relatively small set of tricks and techniques to make this potentially very complicated exercise manageable. In this section, I aim to provide general guidance on the execution of this approach that makes these techniques clear. To do so, I walk step-by-step through the stages that a researcher must complete in order to execute this approach and draw attention to common solutions to the problems that arise at each stage.

Specifying the Model of Welfare

While this organization is not universally the case, many papers relegate their welfare analysis to a short, final section that is presented as a way of interpreting earlier estimates. As a means of efficient scientific communication, I believe this practice often makes sense. However, this structure of writing can lead one to infer that, during the research process, the development of welfare analysis begins after the empirics are largely completed. While this ordering sometimes works, I do not recommend it. These analyses normally involve a model that is comparatively complex. Empirics that are not tailored to the model's requirements will often fail to provide everything that is needed. What is worse, one may determine late in the process that some needed pieces cannot be generated even with modifications to one's empirics.

Given these concerns, I strongly recommend writing out one's desired model of welfare as early as possible in a project so that it might inform the design of the empirical strategy (which might itself then point to necessary changes to the model). In simplest terms, specifying this model will involve providing a precise means of evaluating the social welfare arising from a given allocation and a precise means of forecasting the allocation that will arise from individuals' behavior. After specifying both the welfare criterion and the behavioral model, the research can then turn to estimating the behavioral model.

To begin this process, the first step is specifying a welfare criterion; that is, one must specify how to assess if a situation is better or worse. In common economic applications, this is often done by summing the costs and benefits as in cost/benefit analysis, summing the surplus from trades as in supply/demand analysis, or by summing some measures of individuals' welfare as in utilitarian analysis.

When ranking alternatives using a welfare criterion, a researcher is codifying their moral values. Quite inconveniently for economists, not all humans share the same moral values, and concordantly not all researchers agree on what constitutes

good welfare analysis. Some might be happy to measure welfare with the sum of surpluses as in the Harberger triangle analysis, while others might balk at ignoring *who* gets the surplus (say, the rich or the poor?). Some might prefer to proceed with a sum of utility functions that reflect a declining marginal utility from wealth, while others might balk at the different treatments individuals get in such an approach. Disagreements like these, and many more, provide a large amount of material for debate on essentially any welfare analysis one could write.

The subjectivity inherent in welfare analysis means that deploying it can be contentious, but it need not always be so. In some subfields, or in some topic areas, the need for welfare analysis has been sufficiently strong that researchers have had to engage with it often. And in doing so, they have often developed strong norms on how such analyses should be conducted and have developed extensive literatures to support such decisions. If one wants to assess tax policy, for example, there are extremely well-developed frameworks available that the research community demonstrably will tolerate. If one wants to assess a topic that does not have an existing playbook for welfare analysis, tolerance is not guaranteed.

This leads to one important recommendation for the process of project development: assess early on whether the project requires *just* innovation in the way behavior is modeled or whether it also requires innovation on standard welfare analysis. One could proceed in either case, but it is important to be clear-eyed that simultaneously innovating on two fronts is substantially more difficult than "merely" innovating on one. Battles on multiple fronts should be initiated with great caution and only with a compelling need. This advice is supported when examining our leading examples. In each of these papers, the analysis was carefully designed to look "normal" to members of the relevant literatures if the isolated behavioral element were removed. In each case, I believe the wisdom of the paper might not have been as widely appreciated if this decision had not been made.

Specifying the Model of Behavior

With a welfare function in hand, we may now perform welfare comparisons as long as we know the inputs to the welfare function that arise in each studied situation. In traditional economic analysis, these inputs are often the allocation of goods, which is assumed to be influenced by the choices of individuals pursuing their rational incentives.

The defining characteristic of welfare analysis based on behavioral incentive compatibility is that allocations are assumed to be influenced by the choices of individuals pursuing their incentives while also being affected by behavioral economic forces. The boundaries of what constitutes "behavioral economic forces" are somewhat nebulous, but I personally interpret this very broadly. Clearly within the boundaries are issues that draw directly on cognitive or social psychology; issues related to biased or imperfect forecasting of probabilities or states; issues that relate to social preferences; issues that relate to nonexponential time discounting; and issues that relate to imperfect cognition, perception, or attention. In our three focal examples, some behavioral economic forces were:

(1) a tendency to underreact to nonsalient taxes, which could occur if individuals forget to attend to them, (2) a tendency to fail to change health insurance plans when it is financially advantageous to do so, which could occur if individuals fail to attend to their insurance or find doing so psychologically aversive, and (3) a tendency to incorrectly assess one's probability of acceptance at a school, which could arise from a wide variety of the failures of probabilistic reasoning or information frictions.

Because there are so many ways for decision making to be imperfect, there are an enormous number of possible models that could be deployed within the behavioral incentive compatibility approach. Despite the idiosyncrasy in models that this causes, there are some important regularities in how the models are developed. I highlight two regularities: (1) *using simple models relative to behavioral-economic norms* and (2) *making defensible normative judgements.*

Simple models relative to behavioral-economic norms. When studying imperfections in decision-making, there are often multiple possible underlying forces that could generate the behavior of interest. Modeling the full details of these competing forces can be critical in a study oriented towards best understanding the root cause of the phenomenon. Such a model can illustrate what is necessary to identify separately one force from another, and if estimated it could provide a comparatively detailed and accurate means of predicting behavior. But while there are clearly circumstances where a detailed and process-focused modeling approach is appropriate, proceeding in this way is rarely ideal for pursuing welfare analysis. Some distinctions that are extremely consequential in, say, a study of psychology are not consequential for welfare. In the common situation where tractability is a problem, a researcher faces strong incentives to remove such distinctions from at least the basic version of the model under study.

To illustrate, consider again the underreaction to sales taxes studied by Chetty, Looney, and Kroft (2009). As discussed earlier, there are many reasons why this underreaction could arise, and these reasons might be active at the same time. To repeat a few: perhaps some individuals do not know that the sales tax applies to the item considered, and perhaps some individuals decide not to take the moment to consider the sales tax because they deem it not worth their time, and perhaps some individuals wish to attend to sales taxes but persistently forget to do so. A model that fully incorporated the nuances of these different causes of the behavior would be challenging to identify empirically and would complicate theoretical analysis. However, Chetty, Looney, and Kroft argue that they do not need to model each of these distinctions fully, because the welfare-relevant consequence of any of these stories is a wedge between "true" price elasticity and the analogous elasticity in the presence of nonsalient taxes. Thus, Chetty, Looney, and Kroft work with a maximally simple model of this phenomenon: elasticity is scaled down by a single parameter when price variation is coming from a nonsalient tax. If Chetty, Looney, and Kroft's goal were to fully understand the determinants of this inelasticity, or to determine how to design interventions to combat it, this modeling decision would be limiting. But given that their goal was to incorporate the consequences of nonsalience into

Harberger triangle analysis, this simplification instead makes progress possible where it would not be otherwise.

This value of simplification is also on clear display in the other two example papers. Handel (2013) studies individuals failing to change their insurance plan when it is financially advantageous to do so. Many failures of decision making could lead to this behavior, and yet Handel restricts these forces to operate through a single "as-if" cost-of-change parameter. Kapor, Neilson, and Zimmerman (2020) study families applying to schools in an imperfect way due to their inaccurate assessments of their probability of admission. These inaccurate assessments of probabilities could arise for many reasons and may have many causes, and yet Kapor, Neilson, and Zimmerman work with a simple model and explicitly discuss some issues excluded for tractability. I believe the fact that all three of these papers work with simplified behavioral models reflects a broader regularity: researchers attempting empirical welfare analysis based on behavioral incentive compatibility face a challenging enough task that they often cannot proceed without some degree of model simplification.

The advice to work with a simple model that is tailored to welfare analysis may not feel useful to a researcher who currently has a complex model in hand. In such a situation, how can the complex model be improved? One systematic way to pursue this question is to attempt a sufficient statistics approach, in which the researcher considers the desired welfare analyses and determines, in those formulas, the minimal amount of information that needs to be measured. In some cases, one can find that not all model primitives need to be estimated—a common example is finding that a local elasticity is sufficient for analysis rather than needing to know the further parameters of a utility function. This approach has long been used to facilitate welfare analysis with standard, fully rational economic models. I believe the realization that this approach works quite well for behavioral economic models is one of the factors contributing to the recent surge of work applying behavioral incentive compatibility. For more guidance on the sufficient statistics approach, see Chetty (2009).

Defensible normative judgments. By assuming that individuals pursue behavioral incentives that are different than those encoded in the welfare function, the researcher is assuming that individuals pursue goals that should not be objectively valued. Modern economists have been wary of taking this type of paternalistic stance, and for good reason. Social planners acting on paternalistic motives have at times been mistaken, misguided, or evil, and this has generated a basis to view such analysis as dangerous. What's more, there is an off-putting hubris inherent in paternalistic policy analysis: who is the researcher to say, confidently, that they know what is best for others? These concerns are among the factors that have pushed economists to be so firmly wedded to revealed-preference approaches. And as a result of that training, most economists will only abandon the presumption of welfare-maximizing behavior after being confronted with a quite strong case.

This status quo means that a researcher must make a very strong case for her behavioral incentive compatibility assumptions. In the best-case scenario, this will involve (1) a strong conceptual case for why imperfect decision making might occur, (2) a strong rationale for why pursuit of this imperfection should not be weighted

by the social planner, and (3) a strong empirical demonstration that supports the conceptual case. All three of the running examples were written in accordance with this advice. They each consider a relatively simple decision error that seems natural to many readers. The behaviors they consider are relatively unambiguously "errors" that are difficult to attribute to unusual preferences. And each paper contains clear empirical "smoking gun" evidence that its hypothesized imperfection is active. I believe their ability to deliver on these three requirements was critical to the success of these papers, and these characteristics are common among similar successful cases.

Compared to these examples, researchers face a more challenging situation if they cannot compellingly demonstrate the presence of the hypothesized behavioral channel, or cannot compellingly resolve its welfare-relevance. However, even in those cases, possible paths forward are available.

When the behavioral channel is in doubt, the welfare exercise can still be pursued contingently: *if* individuals behave in this way, *then* these welfare results follow.[3] This path may be of limited interest if few readers accept the "if" clause, but at least it allows for communication of results to those that accept that clause.

When the welfare-relevance of the behavior is unclear, welfare analysis can often become quite challenging to pursue. This problem has plagued some of the most common models in behavioral economics. To illustrate, consider the phenomenon of loss aversion that is famously incorporated into prospect theory (Kahneman and Tversky 1979; for a review in this journal, see Barberis 2013). Loss aversion is modeled as a tendency for individuals to value marginal reductions of a loss discretely more than they value marginal increases of a gain, thus making the assessment of the same absolute amount differ depending on whether it is framed as a loss or a gain. Despite the very large amount of research on loss aversion, there remains active disagreement as to whether it reflects a welfare-relevant preference or a mistake in reasoning. This disagreement has been a hinderance to individuals who seek to conduct welfare analysis with prospect theory (including me). Encouragingly, recent papers have provided useful guidance on how to best proceed in the presence of such modeling uncertainty. The core idea of these papers is to parameterize welfare-relevance and consider a range of values for the relevant parameter. With this framework, one can characterize welfare under the assumption that the behavioral component is zero percent welfare-relevant, 100 percent welfare-relevant, and everything in between. Presenting results in this way allows a reader to assess the conclusions that align with their beliefs on welfare relevance and allows the researcher to clearly communicate when claims are sensitive or insensitive to these assumptions. For development of this approach, see Goldin and Reck (2022) or Reck and Seibold (2023).

[3] Of course, all welfare analysis is contingent on its behavioral assumptions, but it is common (and reasonable) to emphasize this contingency to different degrees depending on the degree of confidence in those assumptions.

Estimating the Model for Welfare Analysis

By completing the steps in the previous sections, a researcher has laid out the key objects necessary to conduct a welfare analysis. We now turn to the question: how might these objects be estimated? As before, the great variety of settings and behaviors that could be modeled preclude a complete answer to this question. However, again, there are clear commonalities in successful approaches.

A useful paradigm for approaching this problem appears in Bernheim and Rangel (2009). They suggest partitioning observed decisions into those that are *suspect* or *nonsuspect*—that is, suspected of being influenced by forces that stop choice from revealing welfare-relevant preferences, or not suspected of doing so. With such a partition in hand, one can then estimate the welfare-relevant parameters (say, of a demand function or of individual utility functions) from the nonsuspect data using standard revealed-preference methods. The parameters of the model of behavioral incentive compatibility can be estimated by applying the same methods to the suspect data. Chetty, Looney, and Kroft (2009) serves as an excellent example of this approach: responses to posted prices are treated as nonsuspect, whereas responses to taxes collected at the register are treated as suspect. This partitioning generates the two different demand curves plotted in Figure 2.

When one has data on both suspect and nonsuspect choices, the framework just described serves as the default template for an empirical approach. This framework is often unavailable, however, due to the researcher determining that *all* observed decisions are suspect. In this situation, the common path forward is to seek additional data that identify the necessary features of the behavioral model. In principle, this exercise could be conducted with many forms of outside data and could even rely on estimated parameters from prior papers. However, the most common version of this approach involves designing and deploying a survey that is precisely tailored to provide the necessary missing information. This approach is well demonstrated by Kapor, Neilson, and Zimmerman (2020), who could not directly infer their probability misperceptions of interest from administrative data on New Haven school choice and thus conducted a survey that directly elicited families' beliefs about match probabilities. With such additional data in hand, estimating a model of probability misperceptions is much more straightforward.

Economists' use of tailored surveys has grown rapidly in recent years. This has both caused, and been caused by, major reductions in the logistical difficulties of deploying such a study. Researchers now have access to both user-friendly platforms for distributing surveys online and means to target the deployment of such surveys directly to the desired study participants. As a result, this tool has greater usefulness, and more and more papers are responding by using a tailored survey to fill a critical gap in field data. In such projects, the design of the survey is a key stage where creativity can be extremely rewarded: pairing the right type of data through these means can make progress possible where it would otherwise be inconceivable.

Performing Welfare Analysis

Once a researcher has estimated all the necessary model components, how should she then proceed with welfare analysis? My recommendation on this question is perhaps disappointingly uncomplicated. With the model in hand, the researcher should now directly attempt to understand the consequences of the economic decisions she set out to study. This could entail comparing welfare before or after a policy change, or comparing welfare across several different economic regimes, or comparing welfare across different values of a policy parameter to inform how it should be determined. Except for having generated the estimated models in different ways, a researcher may largely proceed as she would have if she deployed standard methods.

That said, when pursuing this welfare analysis, it is important to remember that our behaviorally informed models are still imperfect approximations. An immediate implication is that these approximations may fail if we use them to forecast behavior or welfare outside the range of situations used to estimate them. To illustrate, the results of Chetty, Looney, and Kroft (2009) suggest that sales taxes are often ignored when purchasing comparatively cheap items in a grocery store, but this finding might not hold when taxes are much higher or if the goods considered are more expensive.[4] These issues make it important to think critically about the boundaries of safe application of one's estimated model. However, it is also worth remembering that this requirement is in no way new and in no way special. The concerns above are essentially an application of the Lucas (1976) critique—that is, the concern that parameter estimates can change when the underlying policy regime changes. This critique has plagued economists regardless of their reliance on behavioral incentive compatibility.

In the course of conducting this welfare analysis, a researcher will normally wish to establish the exact role that the behavioral incentive compatibility assumption is playing. The answer to this question will of course vary across contexts, but existing research suggests a theme. Across the three studies we examined, we see clearly that individual mistakes do more than merely hurt the people who make them. This recurring finding has served as a counterpoint to the historical tendency of behavioral economists to focus their attention on the individual consequences of these mistakes. It appears that, in some cases, the consequences of behavioral influence on the broader market can be of even greater importance. In our three examples, these broader consequences included lowering the total welfare costs of taxation, preserving an insurance market that would otherwise have been significantly harmed by adverse selection, and disrupting the ability to infer utility from choices to a degree that influences which school-choice mechanism we recommend. In all three of these cases, standard welfare analysis is oriented around studying the consequences to distortions in behavior arising from optimal response to incentives. When incentive

[4] See Taubinsky and Rees-Jones (2018) for supporting evidence.

compatibility is replaced with behavioral incentive compatibility, the manner in which these distortions play out changes, thus driving the differences in the approaches.

Conclusion

The profession's tolerance of imperfectly rational "behavioral" assumptions in welfare economics has changed dramatically in recent history. Prior to the turn of the millennium, behavioral economists largely avoided engagement with full, technical welfare analysis. And indeed, such engagement would rarely have been welcomed. In the span of merely a decade or two, analysis of this variety has gone from being extremely rare to quite common, with notable examples of this analysis serving focal roles in several literatures.

As these analyses have propagated, so too has evidence on how to best pursue them. This paper has summarized commonalities in how these papers are executed and offered guidance on bringing this approach to new problems. As our focal examples illustrate, this body of work has begun to achieve the long sought-after goal of integrating behavioral economics into our most fundamental economic analyses. As the path for such research becomes more deeply trodden, I hope more and more researchers will choose to follow it.

References

Abdulkadiroğlu, Atila, Yeon-Koo Che, and Yosuke Yasuda. 2011. "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered." *American Economic Review*, 101(1): 399–410.

Abdulkadiroğlu, Atila, and Tayfun Sönmez. 2003. "School Choice: A Mechanism Design Approach." *American Economic Review*, 93: 729–47.

Agarwal, Nikhil and Paulo Somaini. 2018. "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism." *Econometrica*, 86: 391–444.

Barberis, Nicholas C. 2013. "Thirty Years of Prospect Theory in Economics: A Review and Assessment." *Journal of Economic Perspectives*, 27: 1 (Winter), 173–96.

Bernheim, B. Douglas and Antonio Rangel. 2009. "Beyond Revealed Preferences: Choice-Theoretic Foundations for Behavioral Welfare Analysis." *The Quarterly Journal of Economics*, 124(1): 51–104.

Chetty, Raj. 2009. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods." *Annual Review of Economics*, 1: 451–88.

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–77.

Dubins, Lester E. and David A. Freedman. 1981. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly*, 88: 485–94.

Gale, David and Lloyd S. Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly*, 69: 9–15.

Goldin, Jacob and Tatiana Homonoff. 2013. "Smoke Gets in your Eyes: Cigarette Tax Salience and

Regressivity." American Economic Journal: Economic Policy, 5(1): 302–36.
Goldin, Jacob and Daniel Reck. 2022. "Optimal Defaults with Normative Ambiguity." *Review of Economics and Statistics*, 104(1): 17–33.
Handel, Benjamin R. 2013. "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts." *American Economic Review*, 103 (7): 2643–82.
Harberger, Arnold C. 1964. "The Measurement of Waste." *American Economic Review*, 54(3): 58-76.
Hines, James R. 1999. "Three Sides of Harberger Triangles." *Journal of Economic Perspectives*, 13(2): 167–88.
Kapor, Adam J., Christopher A. Neilson, and Seth D. Zimmerman. 2020. "Heterogeneous Beliefs and School Choice Mechanisms." *American Economic Review*, 110, 1274–1315.
Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47: 263–91.
Lucas, Robert. 1976. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets*, Vol. 1, edited by Karl Brunner and Allan H. Meltzer, 19–46. Amsterdam: Elsevier.
Pathak, Parag A. and Tayfun Sönmez. 2008 "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review*, 98(4): 1636-1652.
Reck, Daniel. 2016. "Taxes and Mistakes: What's in a Sufficient Statistic?" SSRN Working Paper 2268617.
Reck, Daniel and Arthur Siebold. 2023. "The Welfare Economics of Reference Dependence." NBER Working Paper 31381.
Roth, Alvin E. 1982. "The Economics of Matching: Stability and Incentives." *Mathematics of Operations Research*, 7: 617–28.
Taubinsky, Dmitry and Alex Rees-Jones. 2018. "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment." *Review of Economic Studies*, 85(4): 2462–96.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 175–192

Designing Simple Mechanisms

Shengwu Li

Economists today design rules for real-world marketplaces. For example, economists in the 1990s designed auctions for radio spectrum broadcast licenses and reformed the National Resident Matching Program (NRMP), a clearinghouse that matches graduates of American medical schools to their first jobs (Milgrom 2000; Roth 2002). The practice soon found myriad applications; for instance, economists have designed systems to coordinate electricity generation, to match students to public schools, to swap donated goods between food banks, and to sell loans from the Bank of England to financial institutions (Wilson 2002; Abdulkadiroğlu et al. 2005; Prendergast 2017, 2022; Klemperer 2010).

In these mechanisms, participants convey information about their preferences, which affects who gets what and (sometimes) how much they pay. Participants will naturally try to figure out whether it makes sense to provide information in accordance with their true preferences, or whether it might prove advantageous to strategize—that is, whether the mechanism is "incentive-compatible." If a mechanism is simple to play, then it will be easier for participants to see whether it is in their interest to reveal their true preferences.

But what makes a mechanism simple? When is it easy for participants to see that a mechanism is incentive-compatible? I will start here by explaining how and why economists came to ask these questions. Then I will discuss three recent answers, which will capture different aspects of what makes a mechanism simple.

■ *Shengwu Li is Professor of Economics, Harvard University, Cambridge, Massachusetts. His email address is shengwu_li@fas.harvard.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.175.

Taking Mechanisms Literally

Early work in mechanism design treated mechanisms as metaphors. For instance, Hurwicz (1973) describes a "resource allocation mechanism" as follows:

> Simplifying to the utmost, we may imagine each agent having in front of him a console with one or more dials to set; the selection of dial settings by all agents determines uniquely the flow of goods and services (trade vector) between every pair of agents and also each agent's production (input-output vector), his "trade with nature."

Hurwicz is describing an abstract representation of richer real-world institutions. The dials stand for purchases that the agents could make, bids that they could place, bargaining strategies that they could adopt, and so on. Later work converged on a standard abstraction, in the form of the "revelation principle": Under some conditions, it is without loss of generality to restrict attention to truthful behavior in incentive-compatible revelation mechanisms. In such mechanisms, each agent is asked to report their private information directly, and it is in their best interest to report truthfully (Gibbard 1973; Dasgupta, Hammond, and Maskin 1979; Myerson 1979). A mechanism is *dominant-strategy incentive-compatible*, or "strategy-proof,' if reporting truthfully is always a best response, regardless of the behavior of other players.

On the metaphorical interpretation of a mechanism, it does not matter whether participants understand that the mechanism is incentive-compatible. After all, people do not interact with the mechanism per se; instead, the mechanism represents aspects of their everyday economic transactions. People may be proficient at arranging those transactions—making purchases in shops, haggling over contracts—and yet not recognize that the mechanism is incentive-compatible when confronted with its abstract form.

The practice of designing real-world markets led economists to focus on a more literal interpretation of mechanisms, namely that mechanism design involves "explicitly analyzing the consequences of trading rules that presumably are really common knowledge" (Wilson 1987). On this view, the rules of the mechanism are not shorthand for a vague decentralized process; rather, they capture actual rules of the institution as understood by participants. For instance, the rules for a sealed-bid auction specify what bids are permitted and how the winner and the payments are determined.

When we regard the mechanism as capturing the rules of a formal process, such as an auction or a school choice system, it matters whether participants understand that the mechanism is incentive-compatible. This is a pressing question for market design, for several reasons. First, market designers often consider novel design proposals that are unfamiliar to the intended participants—or worse, may be deceptively familiar, leading participants to adopt heuristics that are unsuited to the new rules. Second, in applications such as school choice or

certain high-stakes auctions, some people participate in the mechanism exactly once, and thus cannot learn from experience. Third, even if the designer asserts that the mechanism is incentive-compatible, participants may distrust that claim. Moreover, that distrust may be justified: before the intervention of economists, the National Resident Matching Program (mistakenly) claimed that its algorithm made it incentive-compatible for medical students to report their true preferences, even though this was false (Williams 1995).[1] Similarly, when Google started selling Internet advertising via auction, it incorrectly claimed that its auction format was strategy-proof, asserting that its "unique auction model uses Nobel Prize-winning economic theory to eliminate [. . .] that feeling you've paid too much" (Edelman, Ostrovsky, and Schwarz 2007). (This presumably referred to the 1996 prize to William Vickrey and James Mirrlees, for work that we will shortly discuss.) To the best of my knowledge, both claims were genuine mistakes rather than intentional deception.

Participants in real-world mechanisms do not always recognize that the mechanism is incentive-compatible. The modern algorithm for the National Resident Matching Program is (for all practical purposes) incentive-compatible,[2] and the program prominently advises that applicants should rank jobs in their preferred order. Nonetheless, 17 percent of applicants claimed, when surveyed, to have submitted a nontruthful rank-order list (Rees-Jones 2018). When Rees-Jones and Skowronek (2018) recruited medical students who recently participated in the National Resident Matching Program to take part in an incentivized experiment with the same algorithm, they found that 23 percent of subjects submitted nontruthful rank-order lists in the experiment. As further evidence, in strategy-proof mechanisms that match students to degree programs, some students make unambiguous mistakes. They rank a program without a scholarship above the same program with a scholarship, even though the scholarship is worth thousands of euros. This behavior has been found in matching mechanisms for graduate psychology degrees in Israel (Hassidim, Romm, and Shorrer 2021) and for undergraduate degrees in Hungary (Shorrer and Sóvágó 2023).

Consequently, it is often not enough that a mechanism is theoretically incentive-compatible. Participants have to see for themselves that it is incentive-compatible; the mechanism has to be simple. Simplicity bypasses the need for participants to trust the mechanism designer. Simplicity eases the cognitive cost of participation, which is just as real as other costs. Simplicity can level the playing field, in the sense of Pathak and Sönmez (2008), by preventing unfair outcomes caused by unequal

[1] At the time, the National Resident Matching Program ran the hospital-proposing deferred acceptance algorithm, which is strategy-proof for the hospitals listing their preferred candidates, but not incentive-compatible for the applicants themselves.

[2] The program now uses the applicant-proposing deferred acceptance algorithm, modified to account for couples. The standard deferred acceptance algorithm is strategy-proof for the proposing side (Dubins and Freedman 1981), but accounting for couples slightly undermines this property. Computational experiments indicate that the probability that an applicant can profitably deviate under the modern algorithm is about 1 in 10,000 (Roth and Peranson 1999).

strategic sophistication. Finally, simplicity allows the designer to rely more confidently on the predictions of classical game theory, because that analysis depends on participants responding predictably and correctly to incentives.[3]

In the rest of this article, I explain three challenges that participants may need to overcome, in order to recognize incentive-compatibility: (1) thinking contingently about unobserved moves by other players; (2) planning for their own future moves; and (3) reasoning about other players' beliefs. In each of these cases, I will explain formal criteria that capture the difficulty in question, and discuss mechanisms that alleviate that difficulty.

These three challenges are not exhaustive of the ways that mechanisms can be simple or complex. This article is not a comprehensive survey, and it is shaped by my idiosyncratic tastes and the limits of my expertise. I have omitted excellent papers for the sake of brevity.

Thinking Contingently about Unobserved Moves by Other Players

Some mechanisms require participants to reason case-by-case about other players' moves, in order to see that the mechanism is incentive-compatible, while others do not.

In a seminal paper, Vickrey (1961) invented the *second-price sealed-bid auction*,[4] in which all participants simultaneously place bids, then the highest bidder wins and pays the second-highest bid. Vickrey argued that the second-price auction is desirable because it is simple: "Each bidder can confine his efforts and attention to an appraisal of the value the article would have in his own hands, at a considerable saving in mental strain and possibly in out-of-pocket expense."

The second-price sealed-bid auction is strategy-proof, so it is in your own best interest to bid your true value. One can see this by analogy. In a dynamic *ascending auction*, the price starts low and gradually rises. At each moment, each bidder decides whether to keep bidding or to quit irrevocably. When just one bidder is left, that bidder clinches the object at the current price. It is obvious that you should keep bidding if the price is below your true value, and should quit if the price is above. Vickrey pointed out that the second-price auction is "logically isomorphic to" the ascending auction. In the ascending auction, each bidder essentially chooses when to quit. The bidder with the highest quit-price wins, and pays the second-highest quit-price. Quit-prices in the ascending auction are equivalent to bids in the second-price auction, so it is a dominant strategy to place a bid equal to your value.

[3] Li and Dworczak (2024) make a formal critique of this last justification. If participants are strategically unsophisticated, then instead of using a simple mechanism, the designer might profitably adopt complex mechanisms that confuse the participants.

[4] It is more accurate to say that Vickrey reinvented the second-price auction. Second-price auctions were used to sell postage stamps to collectors as early as 1893 (Lucking-Reiley 2000).

Notice that Vickrey's argument that the second-price sealed-bid auction is incentive-compatible works by first reasoning about a dynamic mechanism—the ascending auction—and then arguing that two mechanisms are equivalent. But do real participants treat these two auctions as equivalent?

For auctions in the wild, it can be hard to tell whether participants are bidding truthfully, because we usually do not know their value for the object. To overcome this obstacle, Kagel, Harstad, and Levin (1987) had lab subjects bid in auctions for a virtual prize, which had a different value for each bidder. All subjects knew their own value, in dollars, for the virtual prize. The winning bidder received cash equal to their value minus the price they paid. This method enables the experimenter to observe whether subjects are bidding truthfully. Kagell, Harstad, and Levin (1987) randomized whether subjects participated in second-price sealed-bid auctions or in dynamic ascending auctions. Subjects played their assigned auction format up to 30 times in a row, with their values drawn randomly each time. Again, second-price auctions and ascending auctions are equivalent in theory; it is a dominant strategy to choose a bid (or quit-price) equal to your value. Thus, we might expect that in both formats subjects will bid truthfully, and that the clearing price will always be equal to the second-highest value.

The theoretical prediction of truthful bidding was largely borne out in ascending auctions. Subjects rapidly figured out that truthful bidding is optimal, and in 76 percent of auctions the clearing price was at the second-highest value. By contrast, in second-price sealed-bid auctions, subjects did not behave as theory predicted. Subjects often submitted bids far from their values, and the clearing price was at the second-highest value in no more than 20 percent of auctions. Even after 30 rounds of play, prices remained far from the theoretical prediction.

In summary, Kagell, Harstad, and Levin (1987) found that lab subjects rapidly converge on truthful bidding in ascending auctions, but make large and persistent mistakes in second-price auctions. This result is well-replicated (for examples, see Kagel and Levin 1993; McCabe, Rassenti, and Smith 1990; Harstad 2000; Li 2017; Breitmoser and Schweighofer-Kodritsch 2022). Ascending auctions and second-price auctions are not equivalent in practice.

What did our theory miss? Why is it easy for people to see that the ascending auction is strategy-proof? Why is it hard to see this for the second-price auction?

Imagine a participant who is bidding in a second-price sealed-bid auction. Let us assume that this participant is not a trained economist, and thus does not realize that the second-price auction is "logically isomorphic to" the ascending auction.[5] Suppose that the participant values the object at $40. To see that bidding $40 is better than deviating to a different bid, the participant must make a case-by-case comparison, calculating payoffs for each profile of opponent bids. In particular, the participant must understand that if a truthful bid of $40 would win at a price

[5] In a lab experiment, Breitmoser and Schweighofer-Kodritsch (2022) found that subjects are more likely to bid truthfully in second-price auctions when the mechanism is framed in a way that makes the isomorphism salient. But why does the ascending auction framing make it easier to see the dominant strategy?

of $30 (because the highest other bid is $30), then deviating to bid $50 would also lead to a win at $30. And if a truthful bid of $40 would lose, then a bid of $50 could only win at prices above $40. If the participant does not keep track of each contingency, then they might be tempted to bid above their true value in order to raise the chance of winning. Of course, this strategy would be a mistake, because boosting their bid to $50 raises their chance of winning only when the highest other bid is between $40 and $50. In that case, the price is so high that winning is undesirable.

Now imagine a participant who is bidding in a dynamic ascending auction, and values the object at $40. The price is now $30. If the participant continues to bid truthfully, then the participant will either win at a price between $30 and $40, or will lose and have payoff zero. Thus, every possible payoff from bidding truthfully is at least zero. On the other hand, if the participant deviates to quit right now, then their payoff is zero for sure.[6] The worst-case payoff from truthful bidding is at least as good as the best-case payoff from deviating, so the participant need not keep track of different contingencies to see that truthful bidding is optimal.

A symmetric argument establishes that the participant should quit when the price exceeds their value. Suppose that the participant values the object at $40 and the price rises to $41. If the participant quits, then their payoff is zero for sure. If instead the participant keeps bidding, then their payoff can never be positive, because now they can only win at prices above $41.

In summary, we can distinguish dynamic ascending auctions and second-price sealed-bid auctions in this way: To play optimally in a second-price auction, you need to compare strategies case-by-case, whereas to play optimally in an ascending auction, you need not make that comparison.

In Li (2017), I generalized this idea to any extensive-form mechanism. A strategy is *obviously dominant* if, for any deviating strategy, at any information set where the two strategies choose different actions for the first time, the worst-case payoff under the dominant strategy is at least as good as the best-case payoff under the deviation. (When the worst case from the dominant strategy is preferred to the best case from the deviation, a case-by-case comparison is not needed.) A mechanism is *obviously strategy-proof* if the truthful strategy is obviously dominant.

Dynamic ascending auctions are obviously strategy-proof. In contrast, at the moment you submit your bid in a second-price sealed-bid auction, bidding your value can lead to a zero payoff (if you lose), whereas bidding above your value can yield a positive payoff (depending on the second-price bid). Thus, truthful bidding is not an obviously dominant strategy, and second-price auctions are not obviously strategy-proof.

Every obviously strategy-proof mechanism is strategy-proof. In fact, a stronger claim holds: Every obviously strategy-proof mechanism is *weakly group strategy-proof*, meaning that no coalition of players can jointly deviate and all strictly benefit. We will

[6]We have ignored the possibility of ties. Formally, we are modeling the ascending auction as a mechanism in which we cycle between the active bidders in some fixed order, asking each to raise their bid by a dollar or to quit irrevocably. When just one bidder remains, that bidder wins and pays their last bid.

prove the contrapositive. Suppose the mechanism is not weakly group strategy-proof. Then there exists a strategy profile and a deviating coalition, such that the deviation is strictly profitable for every member. Along the resulting path-of-play, consider the first coalition member to deviate from the truthful strategy. That player strictly benefits, so at that information set, one possible outcome from the deviating strategy is strictly better than one possible outcome from the truthful strategy. It follows that the truthful strategy is not obviously dominant, which completes the proof.

Obvious strategy-proofness characterizes the ascending auction. Suppose that we restrict attention to auctions that are efficient (allocating the object to a bidder with maximal value) and in which only the winning bidder makes payments. In that class, an extensive-form mechanism is obviously strategy-proof if and only if it is an ascending auction (Li and Lo 2023, pp. 576–82). Under weaker restrictions, one can characterize obviously strategy-proof mechanisms as "personal-clock auctions," a larger class that includes ascending auctions, descending-price reverse auctions, and some hybrid formats (Li 2017). Roughly, in a personal-clock auction, each participant chooses between clinching a sure thing (for example, leaving empty-handed) and a tentative alternative with a price that gets steadily worse for the participant (for example, bidding to buy the object at a price that increases each round). For another example, consider a reverse auction in which the bidder is a seller, facing a gradually descending price. At each point, the seller can either quit and keep their object (the sure thing) or offer to sell at the current price (the tentative alternative).

In a practical application of these issues, the US Federal Communications Commission (FCC) auctioned off $19.8 billion of radio spectrum rights in 2017. It included a reverse auction that paid television stations to relinquish their over-the-air broadcast rights, so that this spectrum could be repurposed for wireless Internet access. The Federal Communications Commission's reverse auction had to deal with complicated feasibility constraints. To check whether an allocation is feasible, one must solve a graph coloring problem for a graph with up to 3,000 nodes and 2.7 million edges; the nodes represent television stations and the edges represent interference constraints. Stations that share an edge cannot be assigned to the same channel. This problem is computationally hard (Karp 1972). Consequentially, the reverse auction used a suite of state-of-the-art feasibility-checking algorithms to determine the allocation and payments (Leyton-Brown, Milgrom, and Segal 2017).

The reverse auction proceeded, roughly, as follows: The Federal Communications Commission makes an opening offer to each television station. Each television station can either quit or continue. If it quits, then it remains on the air, retaining its broadcast rights with zero net payment. If it continues, then either it sells its rights at the current price, or the FCC makes a lower offer and the process repeats. It is an obviously dominant strategy for each television station to bid truthfully, quitting once the price drops below its value for remaining on the air. This reverse auction belongs to the class of deferred-acceptance clock auctions studied by Milgrom and Segal (2020), which have useful implications for privacy and budget-balance.

I have not described how the reverse auction determines the offers or when it decides to close a sale. Those details depend on the complicated feasibility-checking algorithms, and most bidders lacked the computational horsepower to check whether the algorithms were running as intended. But even without those details, you can see that the reverse auction is incentive-compatible. This is no accident. A mechanism is obviously strategy-proof if and only if truth-telling can be seen to be dominant from a certain kind of partial description. Such a description specifies the sequences of queries that one participant might receive, the answers that are allowed, and how queries and answers map to possible outcomes (Li 2017, Theorem 1). Notice that this description omits how the outcome depends on unobserved opponent moves; in this sense, obviously strategy-proof mechanisms formalize the idea of a dominant strategy that can be recognized without contingent reasoning.

Psychologically, it is harder to account for hypothetical contingencies than to reason about observed events (Esponda and Vespa 2014). Dynamic mechanisms can help participants to avoid mistakes, by resolving uncertainty about what other players have done.[7]

The assumptions we make about preferences affect the structure of obviously strategy-proof mechanisms. Recent work has explored beyond auctions, characterizing obviously strategy-proof mechanisms for settings such as two-sided matching, social choice, and object allocation without transfers (Arribillaga, Massó, and Neme 2020, 2023; Ashlagi and Gonczarowski 2018; Bade 2019; Bade and Gonczarowski 2017; Mandal and Roy 2022; Thomas 2020; Troyan 2019). Other work has built tools to study obviously strategy-proof mechanisms, providing a "revelation principle" (Mackenzie 2020) and algorithms to construct them (Golowich and Li 2021).

Planning for Future Moves

Some mechanisms require participants to plan far in advance for their own future moves, in order to see that the mechanism is incentive-compatible, while other mechanisms do not require such planning.

Suppose that we have a set of indivisible goods and a set of agents, and we want to allocate one good to each agent, without money transfers. One natural solution is to order the agents randomly, give the first agent their favorite good, then give the second agent their favorite good among those that remain, and so on iteratively. This algorithm results in efficient allocation: any other allocation that makes one agent better-off must also make another agent worse-off (Bogomolnaia and Moulin 2001).

[7]This insight connects to a larger literature on contingent reasoning (Martínez-Marquina, Niederle, and Vespa 2019; Cohen and Li 2022; Esponda and Vespa 2023), reviewed by Niederle and Vespa (2023). Glazer and Rubinstein (1996) formalize another sense in which dynamic games can be easier to understand than their static counterparts.

To carry out the above algorithm, we need to ask agents about their preferences. We could approach agents one-by-one in random order, asking each to pick their favorite good among those that remain. The resulting *dynamic random priority mechanism* is obviously strategy-proof. In contrast, we could ask each agent to submit rank-order lists of the goods, and then process those lists according to the algorithm. The resulting *static random priority mechanism* is strategy-proof. However, it is not obviously strategy-proof. Because of the randomness of the order in which the rank-order lists are chosen, it is possible that if you rank goods truthfully, you might receive your third-favorite, whereas if you deviate to rank your second-favorite good at the top, then you might receive your second-favorite good, which you strictly prefer.

In practice, dynamic random priority results in higher rates of truthful play than its static equivalent. In an incentivized laboratory experiment with four players and four goods, dynamic random priority resulted in the dominant-strategy outcome in 93 percent of games, whereas static random priority resulted in the dominant-strategy outcome in only 64 percent of games (Li 2017, p. 3282).[8]

Ascending auctions and dynamic random priority are both intuitively simple. But some games with obviously dominant strategies are not intuitively simple. For example, take any chess position such that White can force a win, and consider the subgame that starts from that position. If White plays the win-forcing strategy, then by definition the worst-case outcome is that White wins. If White ever deviates from that strategy, then the best-case outcome is that White wins. Thus, the win-forcing strategy is obviously dominant in the subgame. However, real chess players often fail to play the win-forcing strategy, even from positions where such a strategy has been found by computers (Anderson, Kleinberg, and Mullainathan 2017). (One of the unanswered questions of chess is whether White can force a win from the opening position. Chess great Bobby Fischer once opined, "I think it's almost definite that the game is a draw theoretically.")

What makes dynamic random priority simple, and chess complicated? Intuitively, to identify a win-forcing strategy in chess, one has to plan far in advance, consider different future contingencies, and then backward-induct to judge the merits of each current move. By contrast, when playing dynamic random priority, each player chooses exactly once, so forward planning is not required. Conventional game theory elides this distinction, because a strategy is viewed as a complete contingent plan, specifying not only what one does right now, but also what one will do at all future contingencies.

One way that a mechanism can be simple is that it might require only limited forward planning. Pycia and Troyan (2023) formalized the idea of limited forward planning. They consider games of perfect recall, that is, the information sets are such that each participant "remembers" all the past information sets that they encountered and all the past moves they took. However, at each information set, the active player can only foresee some of their other information sets. Thus,

[8] Dreyfuss, Heffetz, and Rabin (2022) and Meisner and von Wangenheim (2023) argue that this disparity might be due to loss aversion rather than strategic mistakes.

each player forms a *partial strategic plan* based on their own information set, that specifies moves at foreseeable information sets. This plan is simply dominant if the worst-case outcome from following the plan is at least as good as the best-case from deviating to another plan that chooses differently at that information set. Specifically, the worst-case and the best-case are with respect to the actions of other players and also the actions that the player takes at unforeseeable information sets.

This approach to modeling limited forward planning results in a variety of incentive criteria, which depend on how we specify the foreseeable information sets. At one extreme, suppose that all information sets are foreseeable; in this case, simple dominance is equivalent to obvious dominance. At the other extreme, suppose that only the present information set is foreseeable, and let us call the resulting criterion *strong obvious dominance*. For example, in dynamic random priority, picking your favorite object is strongly obviously dominant. Each player is called to play exactly once, so it does not matter that future information sets are unforeseeable. Strong obvious dominance may seem too demanding a criterion, but there is a large literature studying the welfare and revenue guarantees of posted-price mechanisms, which have strongly obviously dominant strategies (Lucier 2017).

However, one can imagine an intermediate case, where the player can foresee some, but not all, of the future information sets. Consider a bidder who values the object at \$40. The bidder is participating in an ascending auction, with the price rising by \$1 in each step. The price starts at \$1, and the bidder reasons, "If I quit now my payoff is \$0 for sure. But if I agree to bid \$1 and plan to quit at \$2, then my payoff could be \$40 − \$1 = \$39, and is at least \$0. So I'll keep going for now." Then the price rises to \$2, and the bidder thinks, "If I quit now my payoff is \$0 for sure. But if I keep bidding, planning to quit at \$3, then my payoff could be \$40 − \$2 = \$38, and is at least \$0." Thus, by looking just one step ahead at each point, the bidder is led to behave in a way that reproduces the optimal strategy. Notice that the bidder is making a collection of partial strategic plans, and these plans are not consistent with each other. When the price is \$1, they plan to quit at \$2. But when the price reaches \$2, the bidder revises that plan.

To formalize this idea, let the foreseeable information include both the current information set (which includes all the player's earlier moves) and every information set that is "one step ahead." Observe that the ascending auction is one-step simple, meaning that truthful bidding can be induced by a collection of partial strategic plans (one for each information set), each of which is simply dominant when the player looks just one step ahead. Just a little bit of foresight is enough to play optimally in an ascending auction.

Let us return to the allocation of indivisible goods without transfers. As we saw, dynamic random priority is obviously strategy-proof. But there are other obviously strategy-proof mechanisms in this setting; Pycia and Troyan (2023) characterized those mechanisms, and some of them are not intuitively simple, instead requiring participants to plan far into the future in order to see that the mechanism is incentive-compatible. They showed that strengthening the simplicity requirement

from obvious strategy-proofness to one-step simplicity rules out the counterintuitive mechanisms.

It can take detailed contingent thinking to see that a static mechanism, such as the second-price sealed-bid auction, is incentive-compatible. Dynamic mechanisms, such as the ascending auction, can mitigate this difficulty by paring down the contingencies that participants must consider. But dynamic mechanisms can raise new difficulties, in some cases requiring participants to plan in advance—even far in advance—to understand that a given strategy is optimal. The concept of partial strategic plans formally introduced by Pycia and Troyan (2023) provides a way to study this kind of complexity.[9]

Reasoning about Other Players' Beliefs

The study of simplicity is not only about refining strategy-proofness. Even mechanisms without dominant strategies can vary in how simple they are. In particular, some mechanisms require participants to reason about other players' beliefs, in order to see that the mechanism is incentive-compatible, while other mechanisms do not.

Consider bilateral trade with transfers; there is a seller who can produce an indivisible object, and a prospective buyer. The seller has a privately-known cost C, and the buyer has a privately-known value V. If a sale is made at some price p, then the buyer's payoff is $V - p$ and the seller's payoff is $p - C$. Otherwise, both payoffs are 0.

One natural mechanism for bilateral trade is the double auction, studied by Chatterjee and Samuelson (1983): The seller and buyer simultaneously submit offers, s and b respectively. They transact if and only if the offer s from the seller is greater than or equal to the offer b from the buyer. The resulting price is a weighted average of s and b, that is, $\alpha s + (1 - \alpha) b$ for parameter $\alpha \in [0, 1]$. Suppose for the moment that $\alpha = 0.5$, so the mechanism splits the difference between the two offers. In this case, neither side has a dominant strategy. The seller's utility-maximizing offer depends on the buyer's offer b; if the seller knew the buyer's offer b, then the seller would set $s = b$ if the seller's cost satisfies $C \leq b$, and set $s > b$ otherwise.

Suppose it is common knowledge that both players are rational, that is, they both make offers that maximize expected utility. Because both offers are made at the same time, the seller does not know the buyer's offer b. Instead, the seller has to form beliefs about the distribution of b. Because the seller knows that the buyer is rational, the seller knows that the buyer's choice of b depends on the buyer's value V and on the buyer's beliefs about the seller's offer s. And that, in turn, depends on the buyer's beliefs about the seller's cost C, and the buyer's beliefs about the seller's beliefs about b. Thus, the seller's offer depends on a

[9] Jéhiel (1995) and Jéhiel and Samet (2007) also study limited forward planning in games, but it is an open question whether one can adapt those ideas for use in mechanism design.

second-order belief; what the seller believes that the buyer believes about C, as well as a third-order belief; what the seller believes that the buyer believes that the seller believes about V, and so on.

This many-level reasoning process seems fantastical, and in practice will be beyond the capacity of many participants. If participants are inexperienced, it seems unlikely that they will find the equilibrium of the double auction with parameter $\alpha = 0.5$ from first principles. However, some mechanisms do not require participants to reason about higher-order beliefs. Suppose that we take the double auction, and instead set the split-the-difference parameter α equal to 1. Then the seller is effectively making a take-it-or-leave-it offer of s to the buyer, and the transaction occurs at a price of s if and only if $s \leq b$. In this case, it is a dominant strategy for the buyer to set $b = V$, so the buyer does not need to think about the seller's strategy. If the seller knows that the buyer will set $b = V$, then the seller's problem reduces to choosing an offer s to maximize $(s - C)$ multiplied by the probability that $s \leq V$. Observe that the solution depends on the seller's cost C and on the seller's belief about the buyer's value V, but not on any higher-order beliefs. The seller can calculate an optimal offer using only first-order beliefs, although the seller does not have a dominant strategy.

Börgers and Li (2019) proposed a new incentive criterion, picking out mechanisms in which participants can "play well" using just their first-order beliefs. The proposal is as follows: Let us fix, for each player, a set of possible utility functions (which can be thought of as the different values that each buyer might have for an object). Each player will then form a strategy, which means choosing an action based on their utility function; for example, in the double auction, the actions are the feasible offers. A strategy is *undominated* if there is no other strategy that is always at least as good, and sometimes strictly better. A *first-order belief* is a distribution over the utility functions of the other players; for instance, in a double auction, the seller might have a first-order belief that the buyer's value is uniformly distributed on all integers between 0 and 10. A strategy is robust with respect to that first-order belief if, assuming the other players' utilities are indeed distributed according to that belief, the strategy is a best response to every profile of undominated strategies for the other players. A mechanism is *strategically simple* if for every player and every first-order belief, there exists a robust strategy.

Every strategy-proof mechanism is strategically simple, but the reverse does not hold true. For example, the double auction with split-the-difference parameter $\alpha = 1$ is strategically simple, but not strategy-proof. For the buyer, offering $b = V$ is a dominant strategy, and hence a robust strategy. Moreover, offering $b = V$ is the buyer's only undominated strategy, so the seller has a robust strategy which depends only on the seller's cost C and on the seller's beliefs about the distribution of V (that is, the seller's first-order belief). By contrast, the double auction with $\alpha = 0.5$ is not strategically simple. The buyer has many undominated strategies, and for nontrivial seller beliefs there is no robust seller strategy.

Börgers and Li (2019) formally proved that in the bilateral trade setting, all the strategically simple mechanisms have a special structure: The "leading" player either chooses an offer p from a set of permitted prices, or declines trade. If an offer of p

is made, then the "following" player either accepts (and trade occurs at price p) or declines, in which case no trade occurs. This class includes the double auction with $\alpha = 0$ and with $\alpha = 1$, as well as variations that restrict the set of permitted prices.

More generally, Börgers and Li (2019) found that, under a richness condition, the strategically simple mechanisms can be characterized as "local dictatorships." Roughly, this means that if we fix a profile of utility functions, and consider the restricted mechanism that includes only actions that are undominated *at that profile*, then in the restricted mechanism only one of the player's actions determines the outcome. The meaning of what they call "local dictatorships" is subtle, and includes mechanisms that are not "dictatorial" in the colloquial sense. One example of a local dictatorship is a voting mechanism in which each committee member submits a list ranking all alternatives and (at the same time) the chairperson selects two alternatives for a head-to-head vote. From the selected pair of outcomes, the final outcome is the alternative that is ranked higher by a majority of the committee (suppose that all players have strict preferences over the alternatives). In this situation, each committee member has only a single undominated action, namely, to submit their truthful rank-order list. Thus, in the restricted mechanism, the chairperson's action determines the outcome.

Conclusion

Mechanisms can be complicated in many ways, challenging players to think contingently about the moves of other players, to plan for their own future moves, and to reason about other players' beliefs. Thus, there are multiple ways to define what makes a mechanism "simple." The right criterion depends on context; after all, participants in designed mechanisms include schoolchildren, doctors, lumberjacks, fishermen, and telecommunications firms advised by game theorists (Athey and Levin 2001; Marszalec, Teytelboym, and Laksá 2020; Bulow, Levin, and Milgrom 2009).

Why have formal simplicity criteria at all? After all, we know simplicity when we see it. If we use intuition to judge formal criteria, why not use intuition instead of formal criteria? I see two considerations that weigh against the intuitive approach. First, simplicity is not the only goal—there are often others, such as efficiency, fairness, or revenue. Unless we put simplicity on equal mathematical footing, we cannot study these trade-offs systematically. Without formal criteria for simplicity, we may fixate on other desiderata that are well-formalized.

Second, our intuitive judgements about simplicity are distorted by economic training. Experts in mechanism design suffer from the curse of knowledge; it is hard for them to adopt the perspective of people who do not know game theory (Camerer, Loewenstein, and Weber 1989). In order to produce scientific knowledge about simplicity, we must construct theories that can be tested with data. There is a pressing need for experiments, both to test current theories about simplicity and to discover empirical regularities that future theories might explain. In testing theories, bear in mind that some tasks might be easy if done in isolation, but difficult

when they are part of a broader strategic context. Kagel and Levin (2016) survey experiments on auctions, and Hakimov and Kübler (2021) survey experiments on matching mechanisms.

Much remains to be done. The criteria for simplicity discussed here are not exhaustive; there are other dimensions on which mechanisms can be simple or complex. There could be better ways to formalize the same dimensions, or ways that are more tractable, track human behavior more closely, or have firmer cognitive foundations. Moreover, simplicity comes in degrees, and it would be useful to compare the relative simplicity of different mechanisms, or to find ways to say that one mechanism is "as simple as possible" given other constraints. Nagel and Saitto (2023) recently made progress in this direction.

Finally, another important direction for the simplicity literature is to study how to describe and explain mechanisms, to help participants see for themselves that the mechanism is incentive-compatible (Gonczarowski, Heffetz, and Thomas 2023). Real-world mechanisms, such as spectrum auctions and the National Resident Matching Program, often include detailed advice for participants. The Israeli Psychology Master's Match even provided participants with a general-audience lecture demonstrating that its mechanism was strategy-proof, but some participants were clearly not convinced (Hassidim, Romm, and Shorrer 2021). Few laboratory experiments have studied such advice, possibly due to concerns about experimenter demand effects (De Quidt, Vesterlund, and Wilson 2019). The experiment conducted by Masuda et al. (2022) offers a rare exception.

■ *I thank Tilman Borgers, Eric Chen, Yunseo Choi, Jiangtao Li, Marek Pycia, and Peter Troyan for valuable comments.*

References

Abdulkadiroğlu, Atila, Parag A. Pathak, Alvin E. Roth, and Tayfun Sönmez. 2005. "The Boston Public School Match." *American Economic Review* 95 (2): 368–71.

Anderson, Ashton, Jon Kleinberg, and Sendhil Mullainathan. 2017. "Assessing Human Error against a Benchmark of Perfection." *ACM Transactions on Knowledge Discovery from Data (TKDD)* 11 (4): 1–25.

Arribillaga, R. Pablo, Jordi Massó, and Alejandro Neme. 2020. "On Obvious Strategy-Proofness and Single-Peakedness." *Journal of Economic Theory* 186: 104992.

Arribillaga, R. Pablo, Jordi Massó, and Alejandro Neme. 2023. "All Sequential Allotment Rules Are Obviously Strategy-Proof." *Theoretical Economics* 18 (3): 1023–61.

Ashlagi, Itai, and Yannai A. Gonczarowski. 2018. "Stable Matching Mechanisms Are Not Obviously Strategy-Proof." *Journal of Economic Theory* 177: 405–25.

Athey, Susan, and Jonathan Levin. 2001. "Information and Competition in US Forest Service Timber Auctions." *Journal of Political Economy* 109 (2): 375–417.

Bade, Sophie. 2019. "Matching with Single-Peaked Preferences." *Journal of Economic Theory* 180: 81–99.

Bade, Sophie, and Yannai A. Gonczarowski. 2017. "Gibbard-Satterthwaite Success Stories and Obvious

Strategyproofness." In *Proceedings of the 2017 ACM Conference on Economics and Computation*, 565. New York: ACM.

Bogomolnaia, Anna, and Hervé Moulin. 2001. "A New Solution to the Random Assignment Problem." *Journal of Economic Theory* 100 (2): 295–328.

Breitmoser, Yves, and Sebastian Schweighofer-Kodritsch. 2022. "Obviousness around the Clock." *Experimental Economics* 25 (2): 483–513.

Bulow, Jeremy, Jonathan Levin, and Paul Milgrom. 2009. "Winning Play in Spectrum Auctions." NBER Working Paper 14765.

Börgers, Tilman, and Jingtao Li. 2019. "Strategically Simple Mechanisms." *Econometrica* 87 (6): 2003–35.

Camerer, Colin, George Loewenstein, and Martin Weber. 1989. "The Curse of Knowledge in Economic Settings: An Experimental Analysis." *Journal of Political Economy* 97 (5): 1232–54.

Chatterjee, Kalyan, and William Samuelson. 1983. "Bargaining under Incomplete Information." *Operations Research* 31 (5): 835–51.

Cohen, Shani, and Shengwu Li. 2022. "Sequential Cursed Equilibrium." https://doi.org/10.48550/arXiv.2212.06025.

Dasgupta, Partha, Peter Hammond, and Eric Maskin. 1979. "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility." *Review of Economic Studies* 46 (2): 185–216.

de Quidt, Jonathan, Lise Vesterlund, and Alistair J. Wilson. 2019. "Experimenter Demand Effects." In *Handbook of Research Methods and Applications in Experimental Economics*, edited by Arthur Schram and Aljaž Ule, 384–400. Cheltenham, UK: Edward Elgar Publishing.

Dreyfuss, Bnaya, Ori Heffetz, and Matthew Rabin. 2022. "Expectations-Based Loss Aversion May Help Explain Seemingly Dominated Choices in Strategy-Proof Mechanisms." *American Economic Journal: Microeconomics* 14 (4): 515–55.

Dubins, Lester E., and David A. Freedman. 1981. "Machiavelli and the Gale-Shapley Algorithm." *American Mathematical Monthly* 88 (7): 485–94.

Edelman, Benjamin, Michael Ostrovsky, and Michael Schwarz. 2007. "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords." *American Economic Review* 97 (1): 242–59.

Esponda, Ignacio, and Emanuel Vespa. 2014. "Hypothetical Thinking and Information Extraction in the Laboratory." *American Economic Journal: Microeconomics* 6 (4): 180–202.

Esponda, Ignacio, and Emanuel Vespa. 2023. "Contingent Thinking and the Sure-Thing Principle: Revisiting Classic Anomalies in the Laboratory." *Review of Economic Studies*. https://doi.org/10.1093/restud/rdad102.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (4): 587–601.

Glazer, Jacob, and Ariel Rubinstein. 1996. "An Extensive Game as a Guide for Solving a Normal Game." *Journal of Economic Theory* 70 (1): 32–42.

Golowich, Louis, and Shengwu Li. 2021. "On the Computational Properties of Obviously Strategy-Proof Mechanisms." https://doi.org/10.48550/arXiv.2101.05149.

Gonczarowski, Yannai A., Ori Heffetz, and Clayton Thomas. 2023. "Strategyproofness-Exposing Mechanism Descriptions." NBER Working Paper 31506.

Hakimov, Rustamdjan, and Dorothea Kübler. 2021. "Experiments on Centralized School Choice and College Admissions: A Survey." *Experimental Economics* 24 (2): 434–88.

Harstad, Ronald M. 2000. "Dominant Strategy Adoption and Bidders' Experience with Pricing Rules," *Experimental Economics* 3: 261–80.

Hassidim, Avinatan, Assaf Romm, and Ran I. Shorrer. 2021. "The Limits of Incentives in Economic Matching Procedures." *Management Science* 67 (2) 951–63.

Hurwicz, Leonid. 1973. "The Design of Mechanisms for Resource Allocation." *American Economic Review* 63 (2) 1–30.

Jéhiel, Philippe. 1995. "Limited Horizon Forecast in Repeated Alternate Games." *Journal of Economic Theory* 67 (2): 497–519.

Jéhiel, Philippe, and Dov Samet. 2007. "Valuation Equilibrium." *Theoretical Economics* 2 (2): 163–85.

Kagel, John H. 2016. "Auctions: A Survey of Experimental Research." In *The Handbook of Experimental Economics*, edited by John H. Kagel and Alvin E. Roth, 501–86. Princeton: Princeton University Press.

Kagel, John H., Ronald M. Harstad, and Dan Levin. 1987. "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study." *Econometrica* 55 (6): 1275–1304.

Kagel, John H., and Dan Levin. 1993. "Independent Private Value Auctions: Bidder Behaviour in First-, Second-, and Third-Price Auctions with Varying Numbers of Bidders." *Economic Journal* 103 (419): 868–79.

Karp, Richard M. 1972. "Reducibility among Combinatorial Problems." In *Complexity of Computer Computations*, edited by Raymond E. Miller and James W. Thatcher, 85–103. New York: Plenum Press.

Klemperer, Paul. 2010. "The Product-Mix Auction: A New Auction Design for Differentiated Goods." *Journal of the European Economic Association* 8 (2-3): 526–36.

Leyton-Brown, Kevin, Paul Milgrom, and Ilya Segal. 2017. "Economics and Computer Science of a Radio Spectrum Reallocation." *Proceedings of the National Academy of Sciences* 114 (28): 7202–09.

Li, Jiangtao, and Piotr Dworczak. 2024. "Are Simple Mechanisms Optimal When Agents Are Unsophisticated?" Unpublished.

Li, Shengwu. 2017. "Obviously Strategy-Proof Mechanisms." *American Economic Review* 107 (11): 3257–87.

Li, Shengwu, and Irene Lo. 2023. "New Solution Concepts," In *Online and Matching-Based Market Design*, edited by Federico Echenique, Nicole Immorlica, and Vijay V. Vazirani, 575–90. Cambridge, UK: Cambridge University Press.

Lucier, Brendan. 2017. "An Economic View of Prophet Inequalities." *ACM SIGecom Exchanges* 16 (1): 24–47.

Lucking-Reiley, David. 2000. "Vickrey Auctions in Practice: From Nineteenth-Century Philately to Twenty-First-Century E-commerce." *Journal of Economic Perspectives* 14 (3): 183–92.

Mackenzie, Andrew. 2020. "A Revelation Principle for Obviously Strategy-Proof Implementation." *Games and Economic Behavior* 124: 512–33.

Mandal, Pinaki, and Souvik Roy. 2022. "Obviously Strategy-Proof Implementation of Assignment Rules: A New Characterization." *International Economic Review* 63 (1): 261–90.

Marszalec, Daniel, Alexander Teytelboym, and Sanna Laksá. 2020. "Epic Fail: How Below-Bid Pricing Backfires in Multiunit Auctions." CIRJE Working Paper CIRJE-F-1096.

Martínez-Marquina, Alejandro, Muriel Niederle, and Emanuel Vespa. 2019. "Failures in Contingent Reasoning: The Role of Uncertainty." *American Economic Review* 109 (10): 3437–74.

Masuda, Takehito, Ryo Mikami, Toyotaka Sakai, Shigehiro Serizawa, and Takuma Wakayama. 2022. "The Net Effect of Advice on Strategy-Proof Mechanisms: An Experiment for the Vickrey Auction." *Experimental Economics* 25 (3): 902–41.

McCabe, Kevin A., Stephen J. Rassenti, and Vernon L. Smith. 1990. "Auction Institutional Design: Theory and Behavior of Simultaneous Multiple-Unit Generalizations of the Dutch and English Auctions." *American Economic Review* 80 (5): 1276–83.

Meisner, Vincent, and Jonas von Wangenheim. 2023. "Loss Aversion in Strategy-Proof School- Choice Mechanisms." *Journal of Economic Theory* 207: 105588.

Milgrom, Paul. 2000. "Putting Auction Theory to Work: The Simultaneous Ascending Auction." *Journal of Political Economy* 108 (2): 245–72.

Milgrom, Paul, and Ilya Segal. 2020. "Clock Auctions and Radio Spectrum Reallocation." *Journal of Political Economy* 128 (1): 1–31.

Myerson, Roger B. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47 (1): 61–73.

Nagel, Lea, and Roberto Saitto. 2023. "A Measure of Complexity for Strategy-Proof Mechanisms." In *Proceedings of the 24th ACM Conference on Economics and Computation*, 1017. New York: ACM.

Niederle, Muriel, and Emanuel Vespa. 2023. "Cognitive Limitations: Failures of Contingent Thinking." *Annual Review of Economics* 15: 307–28.

Pathak, Parag A., and Tayfun Sönmez. 2008. "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism." *American Economic Review* 98 (4): 1636–52.

Prendergast, Canice. 2017. "How Food Banks Use Markets to Feed the Poor." *Journal of Economic Perspectives* 31 (4): 145–62.

Prendergast, Canice. 2022. "The Allocation of Food to Food Banks." *Journal of Political Economy* 130 (8): 1993–2017.

Pycia, Marek, and Peter Troyan. 2023. "A Theory of Simplicity in Games and Mechanism Design." *Econometrica* 91 (4): 1495–526.

Rees-Jones, Alex. 2018. "Suboptimal Behavior in Strategy-Proof Mechanisms: Evidence from the Residency Match." *Games and Economic Behavior* 108: 317–30.

Rees-Jones, Alex, and Samuel Skowronek. 2018. "An Experimental Investigation of Preference Misrepresentation in the Residency Match." *Proceedings of the National Academy of Sciences* 115 (45): 11471–76.

Roth, Alvin E. 2002. "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics." *Econometrica* 70 (4): 1341–78.

Roth, Alvin E., and Elliott Peranson. 1999. "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design." *American Economic Review* 89 (4): 748–80.

Shorrer, Ran I., and Sándor Sóvágó. 2023. "Dominated Choices in a Strategically Simple College Admissions Environment." *Journal of Political Economy Microeconomics* 1 (4): 781–807.

Thomas, Clayton. 2020. "Classification of Priorities Such That Deferred Acceptance is Obviously Strategyproof." https://doi.org/10.48550/arXiv.2011.12367.

Troyan, Peter. 2019. "Obviously Strategy-Proof Implementation of Top Trading Cycles." *International Economic Review* 60 (3): 1249–61.

Vickrey, William. 1961. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance* 16 (1): 8–37.

Williams, Kevin Jon. 1995. "A Reexamination of the NRMP Matching Algorithm. National Resident Matching Program." *Academic Medicine: Journal of the Association of American Medical Colleges* 70 (6): 470–76.

Wilson, Robert. 1987. "Game-Theoretic Analyses of Trading Processes." In *Advances in Economic Theory*, edited by Truman Fassett Bewley, 33–70. Cambridge, UK: Cambridge University Press.

Wilson, Robert. 2002. "Architecture of Power Markets." *Econometrica* 70 (4): 1299–340.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 193–210

The Problem of Good Conduct among Financial Advisers

Mark Egan, Gregor Matvos, and Amit Seru

The share of US households utilizing financial advisers has increased from 20 percent in 1995 to 30 percent over the past three decades, according to data from the Survey of Consumer Finances. Among households with nonretirement investment accounts, nearly 60 percent reported using a financial adviser in 2019. Despite this growth, a pervasive perception persists that financial advisers—and the broader financial services industry—lack integrity. For the past decade, the financial services sector has consistently ranked among the least trusted industries in the economy, as highlighted by the Edelman Trust Institute (2024). This distrust is not without cause: recent evidence indicates that misconduct within the industry is widespread, not confined to a few high-profile cases. Historically, about one in ten financial advisers in the United States has been involved in misconduct, including criminal or regulatory violations, terminations following allegations, or customer disputes resolved in the client's favor. In this essay, we explore the economic foundations of the financial advisory industry.

Advisers play a crucial role in shaping households' financial decisions, from portfolio allocation to planning for major life events, while also providing emotional

■ *Mark Egan is Associate Professor of Business Administration, Harvard Business School, Boston. Massachusetts. Gregor Matvos is Professor of Finance, Kellogg School of Management, Northwestern University, Evanston, Illinois. Amit Seru is Professor of Finance at the Graduate School of Business, and Senior Fellow at the Hoover Institution and the Stanford Institute for Economic Policy Research, all at Stanford University, Stanford, California. Egan is a Faculty Research Fellow and Matvos and Seru are Research Associates, National Bureau of Economic Research, Cambridge, Massachusetts. Their email addresses are megan@hbs.edu, Gregor.matvos@kellogg.northwestern.edu, and aseru@stanford.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.193.

support and building trust in financial markets. Intuitively, households hire financial advisors because they struggle to distinguish good investment decisions from bad ones. We argue that the same lack of sophistication that drives the demand for financial advice also makes it difficult for households to differentiate between competent and incompetent advisers.

We document that there are approximately 700,000 financial advisers in the United States. This is a high-paying industry (median wage is $95,000), but there is also considerable variation in wages: financial advisers at the top decile of the wage distribution for this industry earn almost four times as much as financial advisers at the bottom decile. In other words, there is substantial demand for financial advisory services, but also substantial heterogeneity among advisers. While this industry is highly regulated, it is far from homogenous. Even the term "financial adviser" lacks a formal legal definition, with many professionals also acting as brokers. This dual role, combined with varied regulatory oversight and certification standards, can create confusion for households and contribute to differing levels of service quality and potential conflicts of interest.

The conflicts of interest are particularly salient in cases of financial adviser misconduct. Previous research has documented significant misconduct in the financial advisory industry. We update those estimates with data until 2024 and find that approximately 6.6 percent of advisers in the United States have a misconduct record as of 2024. Misconduct varies widely across regions and firms, with certain areas and firms showing significantly higher rates. This concentration suggests that misconduct is often a result of firm-adviser sorting, where advisers with misconduct histories tend to cluster in particular firms.

The concern over quality and ethical standards in the financial advisory industry has sparked significant debate in regulatory and policy circles. These proposals range from increased information disclosure to changing the legal obligations of financial advisers. Our findings suggest that a lack of consumer sophistication is the primary friction, making increased disclosure an effective policy response. Using a difference-in-differences approach, we found that the "naming and shaming" of firms with high misconduct rates reduced misconduct by 10 percent, indicating that transparency can improve market efficiency.

The Role of Financial Advisers

Many households lack the financial sophistication needed to make basic decisions (Lusardi and Mitchell 2014). For instance, they often purchase expensive mutual funds despite the availability of cheaper alternatives (Hortaçsu and Syverson 2004; Choi, Laibson, and Madrian 2010; Brown et al. 2023), underparticipate in equity markets (Campbell 2006), and hold undiversified portfolios (Polkovnichenko 2005). As a result, households may seek the assistance of financial advisers to help navigate financial products and identify those with the most suitable characteristics or the lowest costs. In an ideal scenario, financial advisers would offer superior

knowledge—such as a deeper understanding of data, a wider awareness of available products, and insights into optimizing household financial decisions—while households would recognize and use this expertise to make optimal choices. There are two sets of forces that make it difficult to achieve this idealized outcome.

First, because the incentives of advisers are not fully aligned with the households they serve, information asymmetries can lead to inferior financial advice and cause markets to break down. It is well-documented that financial advisers face conflicts of interest, and these incentives significantly affect outcomes. Advisers often earn higher commissions for selling more expensive products that deliver lower risk-adjusted returns for their clients (Egan 2019; Egan, Ge, Tang 2022), and their incentives frequently outweigh those of households in determining which investment products clients purchase. These conflicts extend beyond product choice; for example, a financial adviser might encourage excessive trading to generate additional commissions or engage in other forms of misconduct such as fraud (Egan, Matvos, and Seru 2019). Extensive research highlights the critical role of intermediary incentives and how conflicts of interest can distort households' investment decisions (for example, Bergstresser, Chalmers, and Tufano 2009; Hackethal, Haliassos, and Jappelli 2012; Christoffersen, Evans, and Musto 2013; Anagol, Cole, and Sarkarf 2016; Chalmers and Reuter 2020). In short, misconduct can be profitable for both the adviser and the firm.

One might expect that competition, information, and reputation would tend to drive poor and conflicted financial advice out of the market (Fama 1980; Fama and Jensen 1983). However, empirical evidence suggests that firms and advisers with a history of misconduct can still thrive. For instance, Egan, Matvos, and Seru (2019) illustrate that nearly one-third of advisers with publicly available misconduct records are repeat offenders (as we will discuss in more detail later). Why are these repeat offenders not driven out of the market? A significant part of the explanation lies in the variation in household financial sophistication. Many households struggle to assess and compare financial products effectively ("search and information frictions"), which allows poor advice to persist in equilibrium. For example, survey data show that two-thirds of households mistakenly believe financial advisers have a fiduciary duty to act in their best interest (CFA 2010), and 70 percent incorrectly assume advisers are required to disclose conflicts of interest (Huang et al. 2008).

Here, we highlight the second force that prevents a well-functioning market for financial advice. The same frictions that make it difficult for households to make financial decisions on their own also hinder the seamless provision of financial advice. Just as households struggle to distinguish good investment decisions from bad ones, they also find it hard to differentiate between competent and incompetent advisers. Even after hiring, poor financial outcomes can often be rationalized, allowing subpar advice to persist in the market. If households were better equipped to assess the quality of advice, they could help eliminate low-quality advisers. However, if households could effectively evaluate the quality of financial advice, their need for such advice might diminish significantly.

In other words, financial advice is a "credence good"—a type of good that is difficult for households to evaluate both before and after a transaction (Darby and Karni 1973). Classic examples of credence goods include auto repairs, legal services, and healthcare (Balafoutas and Kerschbamer 2020). The defining characteristic of these goods is that the provider—who holds expertise—has more knowledge about the necessity and quality of the service than the consumer. This information asymmetry makes it challenging for consumers to assess both their need for the service and the quality of what is provided. For instance, a mechanic may recommend replacing an air filter, but the consumer may not know whether (1) the replacement is truly necessary or (2) the mechanic actually performed the task. The same is true in the financial advisory industry: households often do not know what advice they need or how to judge the quality of the advice they receive.

This understanding of the financial advice market suggests that efforts to improve outcomes should focus on addressing either the incentives of financial intermediaries or the information asymmetries between advisers and clients. Before exploring potential solutions, however, we first outline the role of financial advisers, who typically hires them, and key facts about the financial advisory industry.

What Do Financial Advisers Do?

Financial advisers play a vital role in assisting households with investing and financial planning. While much of their work focuses on portfolio allocation—a well-researched area in academic literature—they also guide a range of other financial decisions, including saving, debt management, and planning for major life events such as buying a home, saving for children's education, and retirement. These problems can be complex and dynamic and filled with uncertainty, and thus benefit from specialized expertise. Beyond financial guidance, advisers provide emotional support, helping households navigate the volatility and uncertainty of financial markets (Gennaioli, Shleifer, and Vishny 2015).

Much of the existing literature highlights the pivotal role financial advisers play in portfolio allocation decisions. For instance, Foerster et al. (2017) show that financial advisers are instrumental in shaping investors' portfolio choices, with the identity of a household's adviser explaining more variation in portfolio allocation than the household's own characteristics. However, while advisers are influential in this area, evidence regarding their skill is mixed. As mentioned earlier, advisers often face conflicts of interest, and research indicates that clients' portfolios are frequently suboptimal compared to rational, efficient market benchmarks (Chalmers and Reuter 2020). Additionally, studies by Linnainmaa, Melzer, and Previtero (2021) and Andries, Bonelli, and Sraer (2024) suggest that advisers often display the same behavioral biases as their clients. For example, they tend to engage in frequent trading, chase past returns, and underdiversify—even in their own portfolios (Linnainmaa, Melzer, and Previtero 2021).

The role of financial advisers extends beyond portfolio allocation. Evidence from robo-advising shows that even when algorithms manage portfolio decisions, human advisers continue to add significant value for households (Greig et al. 2024).

This added value largely comes from addressing the emotional aspects of money, investing, and retirement planning. Trust, confidence, and emotions are integral to their services. For instance, financial advisers help clients build trust in financial markets (Gennaioli, Shleifer, and Vishney 2015; Gurun, Stoffman, and Yonker 2018) and encourage them to take appropriate risks (Chalmers and Reuter 2020). The critical role of trust also highlights the potential for, and the harmful effects of, fraud and misconduct within the financial advisory industry.

Who Hires Financial Advisers

We use data from the Survey of Consumer Finances to analyze which households hire financial advisers. Panel A of Figure 1 shows the percentage of households reporting the use of a financial adviser from 1995 to 2019. The share of households using a financial adviser increased from 20 percent in 1995 to 30 percent during this time. However, many households do not participate in financial markets. Among those with nonretirement investment accounts, the percentage of households using a financial adviser rose from approximately 45 percent in 1995 to 60 percent in 2019.

Panels B–D of Figure 1 illustrate how the use of financial advisers varies by household characteristics, such as education, financial sophistication, and income levels. Panels B and C of Figure 1 show a positive correlation between education and income levels and the use of financial advice. However, this relationship is largely driven by market participation, as higher-educated and higher-income individuals are more likely to invest. Among those with investment accounts, the majority rely on financial advisers, regardless of education or income. Interestingly, panel C of Figure 1 shows no clear correlation between self-reported financial sophistication and the use of a financial adviser. This suggests that financial advice is not exclusive to wealthy or highly sophisticated investors, nor is it used solely by less experienced investors.

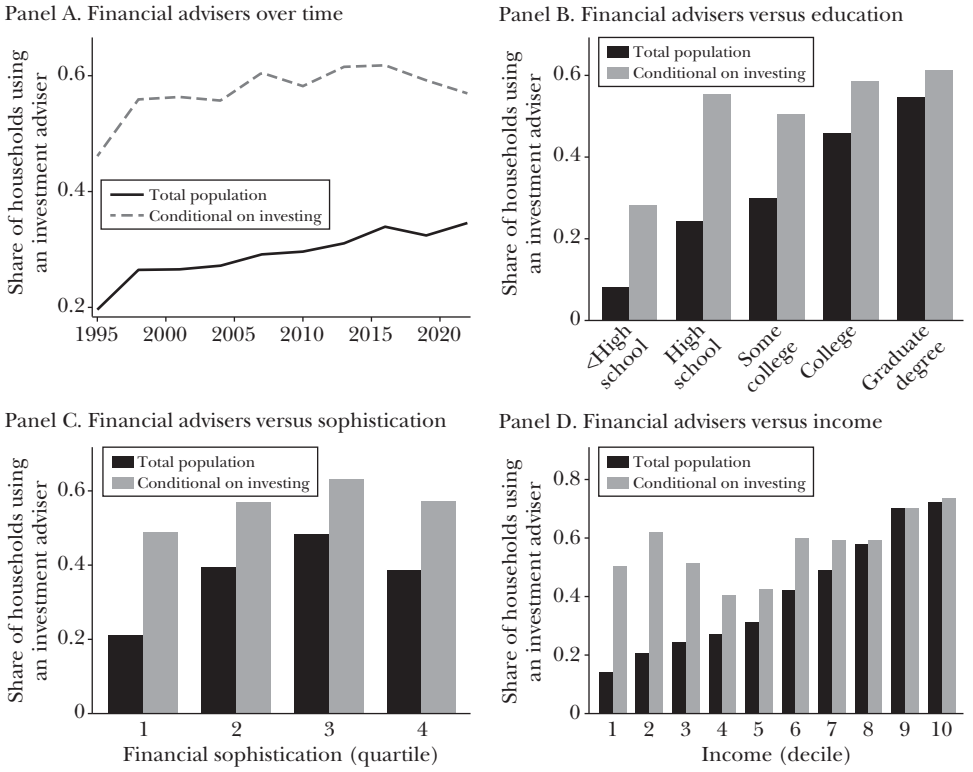
The varying degrees of sophistication among users have important implications for financial regulations, particularly given that financial advice functions primarily as a credence good. This pattern suggests either that individuals with different levels of sophistication receive varying quality of services, or that financial advice is not a substitute for using a financial adviser, at least in terms of broad adoption.

Facts about Financial Advisers

There are approximately 700,000 financial advisers in the United States, representing around 10 percent of total employment in the financial sector (NAICS 52). This share rises to 20 percent when excluding insurers and a significant 75 percent when further excluding commercial banks, according to data from "Industries by Supersector and NAICS Code" produced by the Bureau of Labor Statistics.

However, the title "financial adviser" lacks a formal legal definition, and many professionals use the label. This fact has practical implications, because it means

Figure 1
Share of Households Using a Financial Adviser



Source: Figure 1 shows the share of households reporting the use of a financial adviser in the Survey of Consumer Finances.
Note: Panel A presents this share over time, while panels B–D display the share as of 2019, categorized by education, reported financial sophistication, and income. Results are shown for the total population and for those who invest (that is, have at least one nonretirement investment account).

not all financial advisers are subject to the same regulatory oversight. From a regulatory perspective, most individuals identifying as financial advisers are actually brokers registered with the Financial Industry Regulatory Authority (FINRA), a self-regulatory organization overseen by the Securities and Exchange Commission and, ultimately, Congress.

Brokers primarily offer transaction services to clients and can only provide investment advice as an ancillary part of their business. Importantly, they cannot charge a fee for independent investment advice. While brokers are governed by FINRA rules, they are not held to a "fiduciary duty," meaning they are not legally required to act in their clients' best financial interests. Instead, brokers have historically adhered to a lower "suitability standard," which requires them to recommend products that are "suitable" for clients based on factors like age, financial goals,

and risk tolerance.[1] In response to concerns over conflicts of interest, the Securities and Exchange Commission (2019) introduced Regulation Best Interest (Reg BI) in 2020. Reg BI seeks to align the obligations of broker-dealers more closely with fiduciary responsibilities by requiring them to establish and maintain "policies and procedures reasonably designed to identify and fully and fairly disclose material facts about conflicts of interest, and in instances where we have determined that disclosure is insufficient to reasonably address the conflict, to mitigate or, in certain instances, eliminate the conflict."

Another group of professionals, known as Investment Adviser Representatives, also identify as financial advisers. Employed by investment advisory firms, these Investment Adviser Representatives provide investment advice to households and are regulated by the Securities and Exchange Commission at the federal level, as well as by state authorities. Unlike brokers, Investment Adviser Representatives are held to a fiduciary standard, meaning they are legally required to act in their clients' best financial interests.

What complicates the industry further is that many Investment Adviser Representatives are dual-registered as brokers with the Financial Industry Regulatory Authority, meaning that with the same client, a dual-registered adviser may act as a broker in some instances and as an Investment Adviser Representative in others. This dual role can create confusion for households, as the adviser's legal obligations change depending on whether they are operating as a broker or an Investment Adviser Representative. Research suggests that most households do not fully understand the distinction between these two roles (Scholl and Hung 2018).

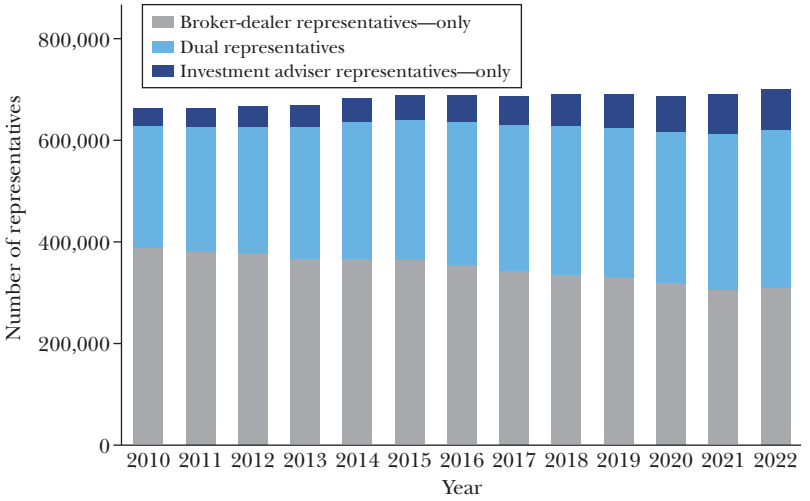
Figure 2 shows the number of brokers, Investment Adviser Representatives, and dual-registered advisers over the past 13 years. The total number of financial advisers, including both brokers and Investment Adviser Representatives, has remained steady at around 700,000. A large majority (89 percent) of financial advisers are registered as brokers, with roughly half of them also dual-registered as Investment Adviser Representatives. The remaining 11 percent are exclusively registered as Investment Adviser Representatives.

To work as a broker or an Investment Adviser Representative, individuals must pass specific licensing exams, which are legally required for operating in these roles. Brokers are required to hold a Series 7 License (General Securities Representative License), regulated by the Financial Industry Regulatory Authority, which permits them to buy and sell securities. They also typically need a Series 63 License (Uniform Securities Agent State Law Exam), a state-level exam governed by the North American Securities Administrators Association (NASAA). Investment Adviser Representatives must usually hold either a Series 65 License (Uniform Investment Adviser Law Exam) or a Series 66 License (Uniform Combined State Law Exam), both regulated by NASAA.[2]

[1] For the full FINRA rules on "Suitability," see https://www.finra.org/rules-guidance/key-topics/suitability.

[2] The Series 66 Licenses essentially combines the Series 65 and Series 63 licenses in one exam.

Figure 2
Financial Advisers in the United States over Time by Registration Type



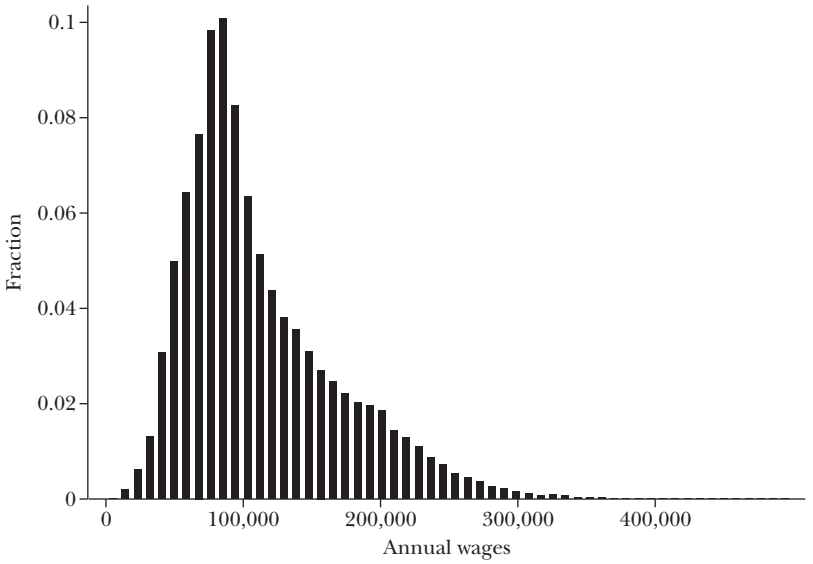
Source: FINRA (2021).
Note: Figure 2 displays the number of registered representatives/brokers, investment adviser representatives, and dual registered advisers over time.

In addition to licensing requirements, financial advisers may choose to pursue certifications. While not mandatory, these certifications are intended to demonstrate a higher level of competence and adherence to ethical standards, though they do not replace licensing requirements. Advisers can hold over 250 different certifications, with the most common being the Certified Financial Planner (CFP) and Chartered Financial Analyst (CFA). As of 2023, there were 98,875 CFPs in the United States. Although these designations are meant to signal adviser quality, evidence from Egan, Matvos, and Seru (2024) suggests that many certifications are actually associated with higher rates of misconduct and may contribute to confusion in the marketplace.

Financial advisers as a group command high wages—the average financial planner earns an annual income of $137,000 as of 2023, according to the Bureau of Labor Statistics' Occupational Employment and Wage Statistics. Figure 3 shows the wage distribution for US financial advisers as of 2024. To construct the figure, we compiled salary data for 247,961 financial advisers by merging adviser-level regulatory employment data from FINRA's BrokerCheck with individual-level salary data from Revelio Labs.[3] Consistent with Bureau of Labor Statistics data, the median wage for financial advisers is $95,000. However, there is significant variation, with

[3] The BrokerCheck data are collected as of January 2024. See Egan, Matvos, and Seru (2019) for further details on the data. We match the adviser-level BrokerCheck data with the Revelio labs data based on the adviser's first and last name and employer.

Figure 3
Distribution of Annual Wages for Financial Advisers (2024)



Source: Employment data are from FINRA's BrokerCheck website and wage data are from Revelio Labs. We truncate the distribution at $500,000.
Note: Figure 3 displays the distribution of annual wages for financial advisers as of 2024. We truncate the distribution at $500,000.

a standard deviation of $60,000. Advisers in the top decile of the wage distribution earn nearly four times as much as those in the bottom decile ($196,000 vs. $54,000).

Financial Adviser Misconduct

Previous research has documented significant misconduct in the financial advisory industry (Dimmock, Gerken, and Graham 2018; Egan, Matvos, and Seru 2019). In this section, we provide new and updated evidence on the prevalence and distribution of misconduct across firms and regions. Building on the work of Egan, Matvos, and Seru (2019), we gathered data from the Financial Industry Regulatory Authority's BrokerCheck as of January 2024. BrokerCheck includes information on all representatives for financial advisers registered with FINRA over the past ten years.

We track each adviser's complete employment, qualification, and disclosure history. According to FINRA, all individuals registered to sell securities or provide investment advice must disclose customer complaints, arbitrations, regulatory actions, employment terminations, bankruptcy filings, and criminal or judicial proceedings. FINRA classifies these disclosures into 23 categories. Following Egan, Matvos,

Table 1
Share of Advisers with One or More Misconduct Disclosure as of 2024

Disclosure	Share
Customer Dispute—Settled	3.41%
Criminal—Final Disposition	1.86%
Employment Separation After Allegations	0.97%
Regulatory—Final	0.93%
Customer Dispute—Award/Judgment	0.37%
Civil—Final	0.02%
Any Misconduct Related Disclosure	**6.56%**

Source: FINRA (2024).
Note: Table 1 displays the share of advisers with one or more misconduct disclosure as of 2024. The total number of observations is 632,271.

and Seru (2019) and subsequent studies, we define misconduct using six categories: Customer Dispute—Settled, Regulatory—Final, Employment Separation After Allegations, Customer Dispute—Award/Judgment, Criminal—Final Disposition, and Civil—Final. In essence, we define misconduct as any criminal or regulatory offenses, customer disputes resulting in a settlement, or employment separations following allegations. Our measure of misconduct likely underestimates the true extent of misconduct, as (1) some misconduct may go undetected or unreported, and (2) cases resolved in favor of the adviser, though not classified as misconduct, may still reflect underlying misconduct.

Table 1 shows the percentage of financial advisers with one or more misconduct-related disclosures on their records as of 2024. Approximately 6.6 percent of currently registered advisers in the United States have a history of misconduct. The most common type of misconduct disclosure involves customer disputes that resulted in settlements. The frequency of new misconduct disclosures has fluctuated over time. In the years leading up to the 2007–2009 Great Recession, about 0.5 percent of advisers received a misconduct disclosure annually. During the recession, this rate spiked to around 0.9 percent for a couple of years. Over the past 15 years, the annual rate has steadily declined, reaching approximately 0.3 percent in recent years.

Previous research shows that misconduct varies significantly across regions and even within counties (Gurun, Stoffman, and Yonker 2018; Parsons, Sulaeman, and Titman 2018; Egan, Matvos, and Seru 2019; Dimmock, Gerken, and Van Alfen 2021; Clifford, Ellis, and Gerken 2023). Table 2 provides an illustration of this variation. In Washington, Vermont, one in 33 financial advisers has a misconduct record, whereas in Guaynabo, Puerto Rico, the rate is much higher, with one in three advisers having a record of misconduct. Egan, Matvos, and Seru (2019) find that misconduct tends to be more prevalent in areas with wealthier, less educated, and older populations. This trend aligns with the examples in Table 2, where roughly one in six advisers in Lee, Florida, and Palm Beach, Florida, has a history of misconduct.

Table 2
Counties with the Lowest and Highest Rates of Misconduct as of 2024

Rank	County	Misc. Rate	# Advisers	Rank	County	Misc. Rate	# Advisers
1	Desoto, MS	2.00%	100	1	Guaynabo Municipio, PR	29.12%	182
2	Rock Island, IL	2.19%	137	2	San Juan Municipio, PR	28.11%	491
3	Kenton, KY	2.21%	3,625	3	Pope, AR	20.54%	112
4	Clinton, MI	2.51%	239	4	Richmond, NY	17.81%	393
5	New York, NY	2.65%	86,903	5	Lee, FL	15.89%	925
6	Elkhart, IN	2.67%	150	6	Napa, CA	15.82%	177
7	Hudson, NJ	2.80%	4,813	7	Suffolk, NY	15.78%	3,928
8	Providence, RI	2.89%	3,186	8	Palm Beach, FL	15.77%	5,820
9	Washington, VT	3.05%	131	9	Martin, FL	15.45%	369
10	St. Louis City, MO	3.06%	2,094	10	Summit, UT	15.33%	137

Source: FINRA (2024).
Note: Table 2 displays the ten counties with the highest and lowest shares of advisers with misconduct records as of 2024 among those counties with at least 100 financial advisers.

The prior literature also shows that misconduct is concentrated within certain firms (Dimmock et al. 2018; Egan, Matvos, and Seru 2019). In the early 2024 data, there are approximately 450 firms with at least 100 financial advisers. Ranking these firms by the percentage of advisers with misconduct records reveals significant variation. At 25 percent of firms, less than 2 percent of advisers have a misconduct record. In contrast, among the top 10 percent of firms, 18 percent of advisers have misconduct records, and at the top 1 percent, this figure rises to 36 percent. Table 3 lists the ten firms with the highest misconduct rates among those employing at least 1,000 advisers. For example, at Oppenheimer and Co., 17 percent of advisers have a misconduct record. This concentration of misconduct appears to result from firm-adviser sorting or firms and advisers "matching on misconduct" (Egan, Matvos, and Seru 2019).

Regulatory Responses

The challenge in addressing financial adviser misconduct stems from the inherent information asymmetry between households and advisers, as well as the misalignment of their incentives. As a result, effective policy responses would likely focus on either improving the incentives of financial intermediaries or reducing these information asymmetries.

A common proposal to improve financial advisers' incentives is to impose a higher fiduciary standard. While evidence suggests this could help mitigate conflicts of interest, it is unlikely to significantly reduce adviser misconduct more broadly. The fact that many advisers already violate existing standards of care indicates that merely raising the bar may not fully address the underlying problems.

Table 3
Firms with the Highest Share of Advisers with Misconduct Records

Rank	Firm	Misc. Rate	# Advisers
1	Oppenheimer & Co. Inc.	17.02%	1,851
2	Cetera Advisors	12.96%	2,145
3	Wells Fargo Advisors Financial Network	12.92%	3,111
4	Stifel, Nicolaus & Company	12.41%	5,054
5	UBS Financial Services	11.94%	11,211
6	Securities America	11.86%	3,161
7	Purshe Kaplan Sterling Investments	11.73%	1,620
8	Janney Montgomery Scott	11.52%	1,728
9	Cetera Advisor Networks	11.44%	4,824
10	Osaic Wealth	11.35%	5,376

Note: Table 3 displays the ten firms with the highest share of adviser with a past record of misconduct as of 2024 among those firms employing at least 1,000 financial advisers.

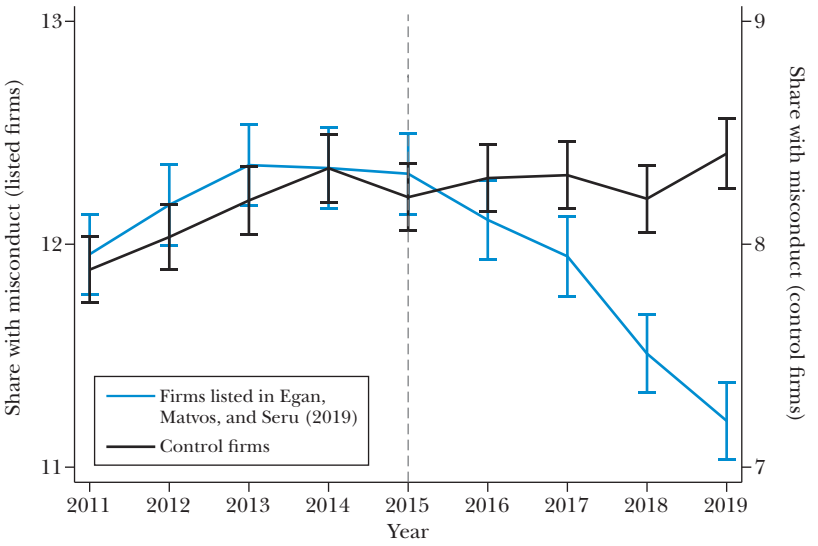
Our findings suggest that the primary friction in this market is the lack of consumer sophistication, making increased disclosure a natural policy response. As mentioned earlier, we were involved in a prior disclosure effort when Egan, Matvos, and Seru published a list of the 20 firms with the highest misconduct rates in early 2016 (the academic paper followed in 2019). The list garnered substantial attention, being featured in major outlets such as the *Wall Street Journal*, *New York Times*, CNBC, Bloomberg, and *Financial Times*. In the months following its release, several firms on the list publicly committed to addressing misconduct. For example, Oppenheimer announced that they had "made significant investments to proactively tackle risk and compliance issues in our private client division. We've made changes in senior leadership, branch managers, and significant changes in our adviser ranks" (InvestmentNews 2016).

Using our financial adviser data, we can now assess whether the firms identified on the misconduct list altered their behavior. To analyze the effects, we employ a difference-in-differences research design. The treatment group consists of the 20 firms with the highest misconduct rates that were publicly named in 2016.[4] The control group includes firms ranked 21–40, which also had elevated misconduct rates in 2016, but were not included on the published list.

For each firm in our treatment and control groups, we calculate the share of advisers with misconduct records at the year-by-firm level using updated BrokerCheck data. We then analyze how the misconduct rate changed for both treatment and control firms from 2011 to 2019, covering four years before and after the intervention.

[4] Egan, Matvos, and Seru posted a paper listing the ten firms with the highest rates of misconduct in February 2016. The paper was subsequently updated in March 2017 to include the 20 firms with the highest rates of misconduct. Consequently, we consider the firms in the top 20 as the treated group and define the firms ranked 21–40, whose names were never published, as the control group.

Figure 4
Share of Advisers with Misconduct—Treatment versus Control



Notes: Figure 4 displays the firm misconduct rate (share of advisers with a record of misconduct) for firms that were listed by Egan, Matvos, and Seru (2019) (that is, treatment) versus firms with similar rates of misconduct that were not listed (that is, control). The bars correspond to 95 percent confidence intervals.

Figure 4 below shows the share of advisers with misconduct records at treated versus control firms around the time we released the list of firms with the highest misconduct rates (that is, treatment). The dashed line marks the start of the treatment in early 2016. The black line represents the control group, where the share of advisers with misconduct records remained relatively steady at 8 to 8.5 percent throughout the sample period. The blue line represents the treatment group, where the share ranged from 11 to 12 percent. Prior to treatment, misconduct rates at both groups followed similar trends. However, after the treatment, the share of advisers with misconduct records at treated firms declines significantly, while the share at control firms remains constant.

Comparing the evolution of misconduct rates at treated firms relative to control firms after the intervention reveals the effect of the treatment. The results show that the share of advisers with misconduct records fell by 1.30 percentage points at treated firms. Given that the average misconduct rate at these firms was 12 percent, this represents a 10 percent reduction. Overall, the "naming and shaming" disclosure policy appears to have made the product and labor markets function more efficiently.

Aligned with these findings, both state and federal regulators have recently implemented policies targeting financial advisory firms with persistently high misconduct rates. For instance, the Massachusetts Securities Division launched the

"Sweep of Select Broker-Dealers that Hire Bad Agents" several years ago, focusing on firms with above-average misconduct rates in the state (Galvin 2016). The investigation examined the hiring and vetting practices of 241 broker-dealer firms in Massachusetts, where more than 15 percent of advisers had a disclosure on record. Similarly, in 2021, the Financial Industry Regulatory Authority introduced Rule 4111 to address firms with a significant history of misconduct. This rule allows FINRA to impose restrictions on firms with elevated disclosure levels and designate them as "Restricted Firms," effectively enabling FINRA to impose capital requirements on these firms.

There have been additional efforts to enhance transparency in the financial advisory industry. The Financial Industry Regulatory Authority's BrokerCheck database is a valuable tool for households and has gained popularity over the past 20 years, as FINRA has continuously improved the accessibility of its website and data. Moreover, FINRA now requires financial advisory firms to include a link to BrokerCheck on their websites, making it easier for investors to research their financial advisers and firms.

Professional certifications and designations, such as the Certified Financial Planner and Chartered Financial Analyst, are also intended to improve transparency in the financial industry and serve as private market alternatives to occupational licensing. However, evidence from Egan et al. (2024) suggests that the proliferation of certifications may actually be counterproductive. With over 250 different certifications available, many households find it difficult to distinguish between them, which in turn leads to a decline in certification standards. As a result, some certifications are linked to higher rates of misconduct and lower-quality advice. Rather than reducing informational asymmetries as intended, certifications may, in practice, worsen them. This issue is not unique to the financial industry—similar problems emerged in the organic food sector, where the abundance of certifications caused consumer confusion. In response, the US Department of Agriculture (USDA) established the National Organic Program to set uniform standards under the "USDA Organic" label.

Finally, another approach to addressing the information challenges households face is through financial education. While raising the overall financial sophistication of households is inherently difficult, recent research suggests that educating advisers could be more effective. Kowaleski, Sutherland, and Vetter (2020) found that a regulatory change reducing the emphasis on ethics in exams for financial advisers led to higher rates of misconduct. The ethics component not only heightened advisers' awareness of the rules but also influenced their perceptions of misconduct. This is significant, as misconduct appears to be at least partially a learned behavior, concentrated in certain firms, and contagious among coworkers (Dimmock, Gerken, and Graham 2018).

Financial advisers play a pivotal and privileged role in the economy, guiding the savings and investment decisions of households. Given their unique position, research and policy efforts should continue to focus on improving transparency, accountability, and the overall effectiveness of the financial advisory industry.

■ *We thank Sam Hanson, Erik Hurst, Adi Sunderam, Luigi Zingales, and editors Nina Pavcnik, Jonathan Parker, and Timothy Taylor for helpful comments and suggestions. Roshan Mahant, Yiyang Han and Winston Xu provided excellent research assistance.*

References

Anagol, Santosh, Shawn Cole, and Shayak Sarkar. 2017. "Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance market." *Review of Economics and Statistics* 99 (1): 1–15.

Andries, Marianne, Maxime Bonelli, and David Sraer. 2024. "Financial Advisors and Investors' Bias." Unpublished.

Balafoutas, Loukas, and Rudolf Kerschbamer. 2020. "Credence Goods in the Literature: What the Past Fifteen Years Have Taught Us about Fraud, Incentives, and the Role of Institutions." *Journal of Behavioral and Experimental Finance* 26: 100285.

Barber, Brad M., and Terrance Odean. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *Journal of Finance* 55 (2): 773–806.

Barber, Brad M., and Terrance Odean. 2001. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics* 116 (1): 261–92.

Bergstresser, Daniel, John M. R. Chalmers, and Peter Tufano. 2009. "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry." *Review of Financial Studies* 22 (10): 4129–56.

Brown, Zach Y., Mark L. Egan, Jihye Jeon, Chuqing Jin, and Alex Wu. 2023. "Why Do Index Funds Have Market Power? Quantifying Frictions in the Index Fund Market." NBER Working Paper 31778.

Campbell, John Y. 2006. "Household Finance." *Journal of Finance* 61(4): 1553–1604.

Célérier, Claire and Boris Vallée. 2017. "Catering to Investors through Security Design: Headline Rate and Complexity." *Quarterly Journal of Economics* 132 (3): 1469–1508.

Chalmers, John, and Jonathan Reuter. 2020. "Is Conflicted Investment Advice Better than No Advice?" *Journal of Financial Economics* 138 (2): 366–87.

Christoffersen, Susan E. K., Richard Evans, and David K. Musto. 2013. "What Do Consumers' Fund Flows Maximize? Evidence from their Brokers' Incentives." *Journal of Finance* 68 (1): 201–35.

Choi, James J., David Laibson, and Brigitte C. Madrian. 2010. "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies* 23 (4): 1405–32.

Clifford, Christopher P., Jesse A. Ellis, and William Christopher Gerken. 2023. "Childhood Exposure to Misbehavior and the Culture of Financial Misconduct." Unpublished.

Consumer Federation of America (CFA). 2010. "Survey: Vast Majority of US Investors Support Clear 'Fiduciary Standard' for Financial Professionals, Widespread Confusion Seen Linked to Current SEC Rules." Press release, September 15. https://consumerfed.org/press_release/survey-vast-majority-of-u-s-investors-support-clearfiduciary-standard-for-financial-professionals-widespread-confusion-seen-linked-to-currentsec-rules/.

Darby, Michael R., and Edi Karni. 1973. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics* 16 (1): 67–88.

Dimmock, Stephen G., William C. Gerken, and Nathaniel P. Graham. 2018. "Is Fraud Contagious? Coworker Influence on Misconduct by Financial Advisors." *Journal of Finance* 73 (3): 1417–50.

Dimmock, Stephen G., William C. Gerken, and Tyson Van Alfen. 2021. "Real Estate Shocks and Financial Advisor Misconduct." *Journal of Finance* 76 (6): 3309–46.

Edelman Trust Institute. 2024. *2024 Edelman Trust Barometer.* New York: Edelman Trust Institute.

Egan, Mark. 2019. "Brokers versus Retail Investors: Conflicting Interests and Dominated Products." *Journal of Finance* 74 93): 1217–60.

Egan, Mark, Shan Ge, and Johnny Tang. 2022. "Conflicting Interests and the Effect of Fiduciary

Duty—Evidence from Variable Annuities." *Review of Financial Studies* 35 (12): 5334–86.

Egan, Mark, Gregor Matvos, and Amit Seru. 2019. "The Market For Financial Adviser Misconduct." *Journal of Political Economy* 127 (1): 233–95.

Egan, Mark, Gregor Matvos, and Amit Seru. 2021. "Arbitration with Uninformed Consumers." NBER Working Paper 25150.

Egan, Mark, Gregor Matvos, and Amit Seru. 2022a. "When Harry Fired Sally." *Journal of Political Economy* 130 (5): 1184–1248.

Egan, Mark, Gregor Matvos, Amit Seru, and Hanbin Yang. 2024. "What Do Professional Certifications Signal?" Working Paper.

Egan, Mark, Gregor Matvos, and Amit Seru. 2024. *Data and Code for: "The Problem of Good Conduct among Financial Advisers."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E209181V1.

Fama, Eugene F. 1980. "Agency Problems and the Theory of the Firm." *Journal of Political Economy* 88 (2): 288–307.

Fama, Eugene F., and Michael C. Jensen. 1983. "Agency Problems and Residual Claims." *Journal of Law and Economics* 26 (2): 327–49.

FINRA. 2021. "Regulatory Notice 21-09: FINRA Adopts Rules to Address Brokers with a Significant History of Misconduct." March 10. https://www.finra.org/rules-guidance/notices/21-09.

FINRA. 2020a. "BrokerCheck Website." https://brokercheck.finra.org/ (accessed December 2020).

FINRA. 2020b. "FINRA Industry Snapshot: Figure 1.5 Securities Industry Registered Individuals by Type of Registration, 2010–2019." https://www.finra.org/sites/default/files/2020-07/2020-snapshot-data-fig-1-5.xlsx (accessed September 2024).

FINRA. 2023. "FINRA Industry Snapshot: Figure 1.1.8 Securities Industry Registered Individuals by Type of Registration, 2013–2022." https://www.finra.org/sites/default/files/2023-04/2023-snapshot-data-fig-1-1-8.xlsx (accessed September 2024).

FINRA. 2024. "BrokerCheck Website." Financial Industry Regulatory Authority. https://brokercheck.finra.org/ (accessed January 2024).

Foerster, Stephen, Juhani T. Linnainmaa, Brian T. Melzer, and Alessandro Previtero. 2017. "Retail Financial Advice: Does One Size Fit All?" *Journal of Finance* 72 (4): 1441–82.

Galvin, William Francis. 2016. *Massachusetts Securities Division's Sweep of Select Broker-Dealers that Hire Bad Agents*. Boston: Massachusetts Secretary of the Commonwealth Securities Division.

Gârleanu, Nicolae, and Lasse Heje Pedersen. 2018. "Efficiently Inefficient Markets for Assets and Asset Management." *Journal of Finance* 73 (4): 1663–1712.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2015. "Money Doctors." *Journal of Finance* 70 (1): 91–114.

Greig, Fiona, Tarun Ramadorai, Alberto G. Rossi, Stephen Utkus, and Ansgar Walther. 2024. "Algorithm Aversion: Theory and Evidence from Robo-Advice." Unpublished.

Gurun, Umit G., Noah Stoffman, and Scott E. Yonker. 2018. "Trust Busting: The Effect of Fraud on Investor Behavior." *Review of Financial Studies* 31(4): 1341–76.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli. 2012. "Financial Advisors: A Case of Babysitters?" *Journal of Banking and Finance* 36 (2): 509–24.

Hortaçsu, Ali, and Chad Syverson. 2004. "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds." *Quarterly Journal of Economics* 119 (2): 403–56.

HUD. 2024. "HUD-USPS ZIP Crosswalk Files." Washington, DC: United States Department of Housing and Urban Development. https://www.huduser.gov/apps/public/uspscrosswalk/home (accessed September 2024).

Hung, Angela A., Noreen Clancy, Jeff Dominitz, Eric Talley, Claude Berrebi, Farrukh Suvankulov, and the United States Securities and Exchange Commission. 2008. "Investors' Level of Understanding." In *Investor and Industry Perspectives on Investment Advisers and Broker-Dealers*, 87–114. Santa Monica, CA: RAND Corporation. http://www.jstor.org/stable/10.7249/tr556ussec.14.

Hung, Angela A., and Joanne K. Yoong. 2013. "Asking for Help: Survey and Experimental Evidence on Financial Advice and Behavior Change." In *The Market for Retirement Financial Advice*, edited by Olivia S. Mitchell and Kent Smetters, 182–212. Oxford: Oxford University Press.

InvestmentNews. 2016. "Why it's more difficult than ever to trust financial advisers." February 25. https://www.investmentnews.com/retirement-planning/why-its-more-difficult-than-ever-to-trust-financial-advisers/66588.

Kowaleski, Zachary T., Andrew G. Sutherland, and Felix W. Vetter. 2020. "Can Ethics Be Taught? Evidence from Securities Exams and Investment Adviser Misconduct." *Journal of Financial Economics* 138 (1): 159–75.

Linnainmaa, Juhani T., Brian T. Melzer, and Alessandro Previtero. 2021. "The Misguided Beliefs of Financial Advisors." *Journal of Finance* 76 (2): 587–621.

Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52 (1): 5–44.

Parsons, Christopher A., Johan Sulaeman, and Sheridan Titman. 2018. "The Geography of Financial Misconduct." *Journal of Finance* 73 (5): 2087–2137.

Piskorski, Tomasz, Amit Seru, and James Witkin. 2015. "Asset Quality Misrepresentation by Financial Intermediaries: Evidence from RMBS Market." *Journal of Finance* 70 (6): 2635–78.

Polkovnichenko, Valery. 2005. "Household Portfolio Diversification: A Case for Rank-Dependent Preferences." *Review of Financial Studies* 18 (4): 1467–1502.

Revelio Labs. 2024. "Individual Level Data." New York, NY: Revelio Labs Inc.

Robles-Garcia, Claudia. 2019. "Competition and Incentives in Mortgage Markets: The Role of Brokers." Stanford Graduate School of Business Working Paper 3796.

SCF. 1995. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 1998. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2001. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2004. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2007. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2010. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2013. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2016. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2019. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

SCF. 2022. "Survey of Consumer Finances." Washington, DC: Board of Governors of the Federal Reserve System

Scholl, Brian, and Angela A. Hung. 2018. *The Retail Market for Investment Advice*. Washington, DC and Santa Monica, CA: SEC Office of the Investor Advocate and RAND Corporation.

Securities and Exchange Commission. 2019. "Regulation Best Interest: The Broker-Dealer Standard of Conduct." *Federal Register* 84 (134): 33318–33492. https://www.govinfo.gov/content/pkg/FR-2019-07-12/pdf/2019-12164.pdf.

Zingales, Luigi. 2015. "Presidential Address: Does Finance Benefit Society?" *Journal of Finance* 70 (4): 1327–63.

US Census Bureau. 2020. "Census County Codes." Washington, DC: US Department of Commerce. https://www2.census.gov/geo/docs/reference/codes2020/national_county2020.txt (accessed September 2024).

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 211–228

Retrospectives: Friedman and Schwartz, Disaggregated

Jennifer Burns

This feature addresses the history of economic terms and ideas. The hope is to deepen the workaday dialogue of economists, while perhaps also casting new light on ongoing questions. If you have suggestions for future topics or authors, please contact either Beatrice Cherrier, CNRS & CREST, ENSAE-Ecole Polytechnique (beatrice.cherrier@gmail.com) or Joseph Persky, University of Illinois at Chicago (jpersky@uic.edu).

Most economists know the name of Anna Jacobson Schwartz. After all, "Friedman and Schwartz" remains a basic reference point in the study of inflation and economic crisis, and the book Schwartz authored with Milton Friedman, A Monetary History of the United States, 1867–1960, stands as a landmark publication in twentieth-century economics.

But how much of that book was Friedman, and how much was Schwartz? Assessments vary widely. Friedman's student Michael Bordo noted: "[H]e couldn't have done it without her" (as quoted in Burns 2023, p. 218). In a recent account of monetarism, Lothian and Tavlas (2018, pp. 783, 762) credit Schwartz's research as having a "*decisive*" influence on Friedman's views, but also insist "he was the lead investigator" on their twelve-year project. Nelson (2020, especially pp. 28–36) gives generous credit to Schwartz in his two-volume account of Friedman's ideas, while largely referring to the two as a unified analyst.

The lack of consensus about what Schwartz contributed to *A Monetary History of the United States* perhaps explains, without excusing, the egregious omission of

■ *Jennifer Burns is Associate Professor of History, Stanford University, Stanford, California, and Research Associate at the Hoover Institution. Her email address is jenniferburns@stanford.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.211.

her contribution from the original press release of Friedman's 1976 Nobel Prize, which lauded *A Monetary History of the United States* as "his major work." By the time of the prize ceremony, the error had been corrected, with the book now called "the large work" and an additional sentence clarifying "Here Friedman has collaborated with an economic historian" (compare Royal Swedish Academy of Sciences 1976; Lundberg 1976).

Indeed, the pervasive discrimination against women that marked twentieth-century economics makes contemporaneous assessments of Schwartz's contribution hard to credit. In her lifetime, for example, Schwartz faced extraordinary difficulties securing a PhD from Columbia University. She received a master's degree in economics from Columbia in 1935, but it was only when publication of *A Monetary History* was imminent that the Columbia economics department agreed, under heavy pressure from Friedman, to grant her a doctorate in 1964. These long-ago struggles have continued to obscure our current understanding of what Schwartz accomplished.[1]

Teasing out the fullness of Schwartz's contribution to the 860-page *A Monetary History of the United States*, and to Friedman's work more generally, would require an equivalent amount of pages. However, their partnership was epistolary, with in-person meetings only about once a year. As a result, the ample archives of Friedman's correspondence held at the Hoover Institution provide an opportunity to disaggregate at least partially the contributions of each economist to this seminal work. Archival evidence, paired with a contextual understanding of the distortions caused by pervasive sexism in the economics profession, suggest three primary contributions Schwartz made to the work, and to Friedman's career more generally.

The first contribution was meeting the classic challenge of quantitative economic history: going into the field to locate and collect archival data that had been assembled for purposes unrelated to economic research, and deciding how best to use those data. Schwartz's prodigious research gave Friedman's establishment of monetarism a sound empirical bedrock, both in *A Monetary History* and other works. Second, Schwartz had a decades-long role as technical sounding board and shaper of the statistical approach taken in the book.

Schwartz's third and arguably greatest contribution was to transform *A Monetary History of the United States* into a compelling narrative argument that made an impact far beyond the economics profession. Covering nearly a century of American history, the book is an extended brief for one central idea: that money matters, that it is the hidden force behind the ups and downs, the breadlines and the bubbles. Combining novel empirical research into the quantity of money with a deeply human story of institutional folly, Friedman and Schwartz also offered a striking reinterpretation of the Great Depression as a preventable liquidity crisis. With great vigor and detail, they argued that the Federal Reserve System's actions "sharply intensified the banking crisis." They also asserted that "the System had

[1] There is no existing biography of Schwartz. The best sources on her life and career are interviews, including Nelson (2004) and Schwartz (1993, 1995). See also the remarks of eight economists at her memorial service in McElroy (2013). Her papers are housed at Duke University.

ample powers to cut short the tragic process of monetary deflation and banking collapse" (Friedman and Schwartz 1963, pp. 10–11).

This stark depiction of Federal Reserve inaction during the Great Depression became a lodestar of public policy and an instruction manual for central bankers. More than sixty years after publication, Friedman and Schwartz still remains the basic playbook on how to avert economic collapse: provide liquidity, and lots of it, fast. Subsequent reactions to economic crisis—from Alan Greenspan's 1987 pledge of Federal Reserve support after the Black Monday stock market crash to the extraordinary Fed-sponsored coronavirus relief programs of 2020—are all driven by the book's analysis of what not to do. Ben Bernanke (2002) famously paid tribute to the pair: "Regarding the Great Depression. You're right, we did it. We're very sorry. But thanks to you, we won't do it again."

The book's economic analysis of the Depression has by now been augmented with new emphases on the banking system and the international gold standard, although monetary factors are still considered primary causes of the slump and its persistence (Bernanke 1995; Eichengreen 1992). This advancement in economic thinking is unremarkable. What is remarkable, however, is that the ideas of Friedman and Schwartz still percolate through best-selling accounts of financial meltdown, haunt officials of the Federal Reserve, and guide policy on the most basic level. The combination of deep empirical research with a historical narrative resulted in the book's astonishing longevity, and owes nearly everything to Schwartz.

Quantitative Historical Research

Friedman and Schwartz became collaborators through the National Bureau of Economic Research, then headquartered in New York City, where Schwartz had taken a job as a researcher in 1941. When Arthur Burns took the helm of NBER in 1948, he convinced his protégée Friedman to join Schwartz on a project studying the role of money in business cycles. While this turned out to be an inspired pairing, it was not an obvious one. There was a vast gulf in status between the two—Schwartz was a researcher with a master's degree, and Friedman was a tenured professor at the University of Chicago. At the time, Schwartz thought of Friedman as "a statistician, and not particularly an up-and-coming economist." Furthermore, she had the impression "there really was no role for money in his analysis" (Nelson 2004, p. 401). Nonetheless, Burns well knew Friedman's brilliance—he had taught him as an undergraduate at Rutgers. And after fighting off Simon Kuznets for the top spot at NBER, Burns knew he could count on Friedman's loyalty, too. Still, Friedman knew almost nothing of economic history—meaning it was Schwartz, already a coauthor of a three-volume history of the British economy (Gayer, Rostow, and Schwartz 1953)—who drew up a reading list to get him oriented.[2]

[2]This reading list has been lost, but is referenced in Schwartz's first letter to Friedman. Schwartz to Friedman, April 5, 1948, Box 90, Folder 1, Milton Friedman Papers, Hoover Institution, Stanford, CA.

Although it may be hard to imagine from today's vantage point, when Friedman and Schwartz began working together, most economists considered monetary policy unimportant. Indeed, the National Bureau of Economic Research's interest in money was a holdover from the preoccupations of its now-deceased founder, the institutionalist Wesley Mitchell, whose career spanned the years in which few doubted that money mattered. At the height of the gold standard regime in 1911, the Yale economist Irving Fisher illuminated the workings of the system with a simple quantity theory equation that linked the money supply, velocity (how frequently money changed hands), price level, and total volume of trade: MV = PT. The quantity theory had flourished alongside the emergence of the classical gold standard, which became the dominant currency regime of the world's major trading nations in the 1870s (Laidler 1991). Another version of the quantity theory was developed in Britain by Alfred Marshall and his followers, among them the young John Maynard Keynes. At the University of Chicago in the early 1930s, Friedman had been steeped in these ideas (Burns 2023, chap. 2, 5; Tavlas 2023).

By 1948, however, the quantity theory had been pushed aside. Holding gold was illegal in the United States, and the US Department of the Treasury controlled the ratio of gold to paper money in accordance with the Bretton Woods agreement hammered out after World War II. At the same time, most economists had accepted the analysis of the Great Depression proposed by Keynes in his 1936 *General Theory of Employment, Interest, and Money*, which explained the economic crisis as a failure of aggregate demand. In this scenario, the solution was government spending, and money or monetary policy was largely irrelevant. To look to the central bank for help would only be "pushing on a string," as the saying went. The real levers of economic policy, everyone knew, were the taxing and spending policies of the federal government. But even as many US economists embraced Keynesianism and left the quantity theory behind as a relic of another era, Friedman saw a revitalized quantity theory as a potential counterpoint to the rising tide of Keynesian economics.

But at first, Friedman had little idea how to go about supporting his hunch, and here Schwartz's emphasis on systematic empirical methodology, knowledge of banking institutions, and feel for novel data proved invaluable. In her first letter to Friedman, she pushed him to specify how concrete data from quantitative historical research applied to a general concept like "money." Practically, did money mean vault cash, government bonds, and reserve balances? How about postal savings deposits and the surrender value of life insurance policies? "What role do the banks play in your concept?" she wondered. Schwartz was forcing Friedman and his theories down to earth, compelling him to recognize the variegated and diverse instruments that could be considered money, along with their many institutional channels. "You are raising exactly the kind of questions that should be raised and that I would have overlooked because of not recognizing the actual difficulties of getting the figures together," returned Friedman, energized by the barrage.[3]

Henceforth cited as MFHI.

[3] Milton Friedman to Anna Schwartz, April 22, 1948, Box 90, Folder 1, MFHI.

Schwartz was educating Friedman in the practice of economic history, where even seemingly rote tasks required nuanced judgement. An "ingenious" comment from Schwartz about insured deposits "troubles me a good deal," he reported, sharing he would "think this point through some more since I am not entirely sure of my ground." In another letter, while expressing pleasant surprise that much of the data he wanted was available, he wondered if they should measure the total stock of money, or just currency in the treasury.[4] In the end, the duo decided the only way to get at money supply as a meaningful aggregate concept was to measure its many potential components. Schwartz took the lead in putting together a time series of both bank deposits and currency in public circulation. Eventually, these would become the M1 and M2 aggregates that monetarism made famous.

Today, similar data can be found with a few keystrokes. In an era before computers were widely used, assembling these figures required painstaking research and a considerable appetite for ambiguity. In the summer of 1949, Schwartz travelled to Washington, DC, to scrounge for data. The two were attempting to create a time series of only bank deposits, but this measure combined many different types of data. Schwartz described parsing through figures including "National Bank float, call dates, 1865–1914 . . . Government deposits, call dates, 1874–1914, National Bank individual deposits corrected for float, and including government deposits, 1865–1868; 1872–1914."[5] Sometimes these data could be found in bank publications, but often they were in ledgers held by individual banks. Schwartz hit the road to find them.

To pull these figures into any meaningful form required creative quantitative problem-solving and dogged persistence. Each state gathered banking data differently. Some calculated month to month; others midmonth to midmonth. It was hard to know where to begin, where to end, and how to compare across categories. Making progress required intense intellectual labor, both to stay with the details and to make big-picture decisions about what mattered and why. For example, a 1951 report from Schwartz to Friedman provides a glimpse of her process in calculating of "vault cash"—cash that banks held on hand to serve customers. Integrally connected to the velocity of money, vault cash was an essential building block in Friedman's revitalized quantity theory and related policy proposals. Schwartz wrote:

> I have been in a perfect fury of work with the vault cash figures. You have no idea what I've been through. Getting the non-national Bank figures by themselves, even when original sources are consulted, wouldn't have been too time-consuming, but converting them into nonmember bank figures, from 1929 on, at any rate, means using both the abstracts of reports of national banks (which until recently don't show an aggregate state figure, but County banks, reserve cities, separately, so that the copying job is mean) and member

[4]Friedman to Schwartz, April 22, 1948, and Friedman to Schwartz, March 4, 1949, Box 90, Folder 2, MFHI.
[5]Schwartz to Friedman, July 24, 1949, Box 90, Folder 2, MFHI.

bank call reports, getting the residual state member figure, and finally the nonmember figure.[6]

Schwartz's reference to her "copying job" highlights that much of her output was handwritten—Friedman's archive is stuffed with legal papers filled with columns of numbers scratched out by Schwartz. She often concealed her exceptional efforts with brief summaries: "Enclosed is a series of individual deposits at all national banks less cash items and clearinghouse exchanges, call dates, 1865–1883," she wrote in a typical phrase to Friedman.[7] By the end of 1951, after more than three years of work, only two decades of historical data had been assembled—and only deposits at that. To bring the project to fruition would require, in the end, twelve years of intensive research and writing.

Even at the early stages, however, this ingathering of data opened new avenues of analysis—which Friedman and his students quickly transformed into publications. In 1952 Friedman published an *American Economic Review* article that compared episodes of wartime inflation, although in a glaring omission, he thanked only his graduate students Phillip Cagan and David Fand, while noting the article was a summary of "some of the main findings of a larger and still unfinished study" (Friedman 1952, p. 612). In Schwartz's hands, these "main findings" were not mere columns of figures. For example, she wrote up a short paper on "resumption"—the post-Civil War shift away from fiat greenbacks to a specie-backed currency. Friedman immediately shared it with graduate students in his Chicago money and banking workshop, writing to Schwartz, "I hope you don't mind my using your product for this kind of educational purpose. In addition to the interest of the period, I think your piece particularly valuable for the students on another score; namely, its careful and intensive use of detailed empirical evidence." Schwartz's account proved particularly useful to workshop students like James Kindhal (1961) and Eugene Lerner (1956), who would publish articles on the history of money in the Civil War era. But many workshop students knew little of Schwartz, who appeared as some distant figure with little connection to the important work happening at Chicago.[8]

Technical Sounding Board

Schwartz helped Friedman solve many of the most vexing statistical issues that intrinsically arose from the diverse, nonstandard nature of historical sources.

Schwartz excelled in comparative analysis and drawing coherence out of messy empirical findings. Once the duo had data in hand from multiple states, it became clear that in a country as large as the United States, creating an integrated

[6] Schwartz to Friedman, November 16, 1951, Box 90, Folder 3, MFHI.
[7] Schwartz to Friedman, November 16, 1951, Box 90, Folder 3, MFHI.
[8] Friedman to Schwartz, April 5, 1956, Box 90, Folder 1, MFHI.

national picture would be nearly impossible. The behavior of vault cash, in particular, seemed to vary wildly between different states. Friedman was looking for an empirical pattern to money's ebbs and flows that would support a broader theory about the role of money in business cycles. However, "I must confess that despite several days intense work on the stuff you sent me, I am nearly as baffled as ever," Friedman wrote in 1952. The only idea he could imagine was that "that geography plays a big part—East versus West or Northeast versus rest of country."[9] When Schwarz's additional analysis disproved that theory, and invalidated another one that mutual savings banks made the difference, the mood turned to "unmitigated gloom," as Schwartz wrote.[10] Recovering, Schwartz then apparently suggested they isolate the Southern states. This would create three groups: states with mutual savings banks, states without mutual savings banks, and the southern states. Friedman resisted: "my reason revolts somewhat—it says we are spending altogether too much time on this step, and that the improvement likely to be attained not be worth the cost."[11] Nonetheless, he eventually agreed that she should proceed.

Schwartz had a feeling for institutional details and historical particularity that Friedman lacked. Her fascination with Confederate money led her to make the case for Southern distinctiveness. After all, the Civil War had unleashed not only wartime inflation but a brief proliferation of new currencies existing side by side. In addition, the similarities between the states of the former Confederacy seemed to warrant separate treatment. Unlike the commercial centers of the Northeast, southern states were rural, minimally banked, agricultural societies. It made sense that money would circulate differently in these types of places. Schwartz quickly found correlations in velocity between these states, asking Friedman modestly as she reported her findings: "do you think it's a third group?" Friedman at first demurred, but in subsequent letters the two discussed the southern states as a distinct group, and by December Friedman had agreed to the division.[12]

In general, there were few interpersonal tussles or power plays between the two—in part because Friedman so clearly had all the power. Still, relative to the other men that surrounded Schwartz, Friedman's willingness to consider a suggestion from his research partner stood out. At the National Bureau of Economic Research, Schwartz routinely found her authority challenged on matters large and small relating to the project. When it came time to publish an early summary of the research, for example, her NBER superiors fought Schwartz's suggestion to publish a table instead of a chart—until learning that Friedman supported the idea, too.

[9] Friedman to Schwartz, July 3, 1952, Box 90, Folder 5, MFHI.
[10] Schwartz to Friedman, August 8, 1952, Box 90, Folder 5, MFHI.
[11] Friedman to Schwartz, October 8, 1952, Box 90, Folder 5, MFHI.
[12] Schwartz to Friedman, October 13, 1952, and Friedman to Schwartz, December 2, 1952, in Box 90, Folder 5, MFHI. This group appears to have survived into *A Monetary History* as nine states (Alabama, Arizona, Arkansas, Colorado, Delaware, Florida, Georgia, Idaho, Indiana) separated from the larger category of 38 rural states. Friedman and Schwartz, 726.

Nor would they take her word about "the nature of our volume" or accept her seasonal adjustment calculations until Friedman had spoken.[13]

Interpolation was a major focus of Friedman and Schwartz's early correspondence. For all Schwartz's efforts, the data were inevitably piecemeal. In countless letters, the two discussed ways to clean and scrub the incomplete data and patch over gaps. Schwartz understood the value of their work to other researchers. After Friedman wrote up some notes about their procedures, Schwartz encouraged him to publish the result: "I think your notes point the way to a higher standard of statistical integrity then at present exists . . . I should therefore like to see you recast the notes, placing the chief emphasis on this question of how one goes about choosing an interpolator."[14] Friedman pushed back, asserting "I would much rather spend my time on economics and in particular in learning about money . . . we have learned what we need to know."[15] Yet eventually Schwarz prevailed, with the notes appearing as an article in the *Journal of the American Statistical Association* (Friedman 1962).

While the grouping of states and interpolation are two areas in which Schwartz's interchanges with Friedman are easy to see, they are not the only ones. Their correspondence is studded with moments where Friedman ceded ground to her ideas, accepted her suggestions, or asked her for help. There was the velocity calculation that yielded "crazy results," which he was hoping Schwartz would check for him. In another letter, typical of their collaboration, he wrote: "The more I work over the cycle analyses, the more I am persuaded that the rate of change dating to which you forced me by your chart is likely to be preferable to what I did."[16]

As a whole, the letters show two scholars working closely together in an atmosphere of mutual respect. Undoubtedly, the long-distance set up of their working relationship, which mirrored gender norms of the 1950s more broadly, minimized opportunities for dispute and disagreement. It is thus all the more striking that Friedman came to rely on Schwartz as a true intellectual partner, never doubting her basic capacities or ability. After a few years working together, Schwartz had become fundamental to the project.

Some decades later, Friedman affirmed her contribution with a self-deprecating joke: "From my point of view, it was an almost perfect example of collaboration. Anna did all the work and I got a lot of the credit. How much more can you ask than that?" More seriously, Friedman went on to underscore his respect for Schwartz: "I always knew that everything she did was going to be done right. It was going to be precise, it was going to be accurate, it was going to be thoughtful" (Brunner and Friedman 1989, pp. 247, 249) This description matches the recollection of other scholars, including David Laidler, who assisted Schwartz for a summer at a late stage of the manuscript. "Working for Anna, even for ten weeks or so, was transformative,"

[13] Schwartz to Friedman, January 14, 1955, and Friedman to Schwartz, January 20, 1955, Box 90, Folder 8; and Schwartz to Friedman, March 7, 1955, Box 90, Folder 8, MFHI.
[14] Schwartz to Friedman, September 26, 1951, Box 90, Folder 3, MFHI.
[15] Friedman to Schwartz, October 5, 1951, Folder 90, Box 3, MFHI.
[16] Friedman to Schwartz, August 19, 1956, and Friedman to Schwartz, July 24, 1956, Box 91, Folder 1, MFHI.

he recalled.[17] Schwartz's billing on the cover of *A Monetary History* was not generosity on Friedman's part, but testament to her profound impact on the book, from statistical details to its scope and scale.

A Narrative History

The best way to understand Schwartz's contribution to *A Monetary History of the United States* is with a counterfactual: What would the book have looked like without her input? Assuming—a big assumption—that Friedman could have gotten the data from another source or person, what would he have done with it? Friedman put the matter succinctly: "Anna is an historian and I am not" (Brunner and Friedman 1989, p. 249).

Indeed, *A Monetary History* opens with a curious confession: the authors had not actually intended to write a book (Friedman and Schwartz 1963, p. xxi). Using one authorial voice, Friedman and Schwartz introduce a deceased colleague, Walter W. Stewart, who encouraged them to write "an 'analytic narrative' of post-Civil War monetary developments in the United States as a background for the statistical work." They recount that following Stewart's suggestion, they decided to "include a chapter on the historical background of the money stock in our planned monograph." But then this chapter "took on a life of its own. The one chapter became two, then a separate part, and has now become a separate book."[18] This was definitely not what Friedman originally had in mind.

Back in February 1954, Schwartz had for the first time asked Friedman what the final product might look like, even as she drew up a tentative outline entitled "Money Supply in the United States, 1907–1953." She wrote: "I shall be highly disappointed if you will decide in favor of a restricted data report with no attempt to tackle the basic economic issues, the importance of which is the only justification for our time expenditure on the estimates." With a three-volume comprehensive history of the British economy under her belt, Schwartz hankered to do more.[19]

But Friedman's response showed him inclining to just what she feared. "This will not be a long drawn out project for my feeling is that we should skim the cream without being either, on the one hand, superficial, or on the other, exhaustive," he wrote Schwartz, after the two had worked together for six years. [20] A year later, Friedman was still visualizing "a book in which part one would be relatively brief, perhaps 50 or 60 printed pages and would present the estimates and analyze them to a rather limited extent," followed by a much longer section that "would

[17] David Laidler, personal communication to author, April 21, 2024.
[18] The original planned monograph, the authors revealed, was still forthcoming (ultimately it would appear as *Monetary Statistics of the United States: Estimates, Sources, Methods* (Friedman and Schwartz 1970), with another comparative volume to follow (Friedman and Schwartz 1982).
[19] Schwartz to Friedman, February 5, 1954, Box 90, Folder 7, MFHI.
[20] Friedman to Schwartz, February 10, 1954, Box 90, Folder 4, MFHI.

contain the description of the preparation of the estimates."[21] He was essentially proposing a fact sheet for economists, the sort of output that the National Bureau of Economic Research regularly published at that time.

In other letters, Friedman expanded on a vision of a book composed of interchangeable parts that could be read in any order. Part 1 would analyze the economic significance of their estimates, followed by "a self-contained Part 2, that would describe in considerable detail their derivation." This structure would benefit readers, Friedman argued, for they could skip over whatever was of less interest. It would also benefit the writers, he suggested, "because each part can be written without having to worry about how it integrates with the other in any section-by-section way, because it will mean that the two parts can be written independently."[22] Friedman was describing not a narrative, but rather a compendium of charts and graphs.

Schwartz had to tread delicately. For all his deference and encouragement, Friedman was in charge. She pushed her own vision of a broader book indirectly; in fact, she ran a sort of shadow campaign for a more comprehensive history from the start. In 1951 she confessed to having "stole some time from the vault cash series" to explore a few works on Confederate finance, calling it "fascinating material." When she sent more "Civil War material," Friedman took Schwartz's writing and "put a student to work on it."[23] A few years later, Schwartz's fascination with that period led her to produce the comprehensive paper on resumption, tracing the post-Civil War shift away from fiat greenbacks to a specie backed currency. Although this essay does not survive in the archive, in a letter to Friedman's student Phillip Cagan she described it as "a story of resumption, from 1865 on, as well I understand it."[24] Ultimately, while *A Monetary History* did not cover the Civil War, the resumption essay anchored it firmly in the immediate postwar years, stretching the book's chronology to nearly a century.

The appearance of Friedman's 1956 book *Studies in the Quantity Theory of Money* must have given Schwartz encouragement, for it showed Friedman and his students engaging deeply with monetary history. Introducing this collection of papers written in his Chicago workshop, Friedman (1956, p. 1) used his students' historical forays to buttress what he called "a restatement" of the quantity theory which he believed could be "a flexible and sensitive tool for interpreting movements in aggregate economic activity and for developing relevant policy prescriptions." The same year the book appeared, Schwartz sent a draft of material covering 1897–1914, noting, "I have worked on the section on and off for the past half year and I have revised it a

[21] Friedman to Schwartz, March 1, 1955, Box 90, Folder 8, MFHI.
[22] Friedman to Schwartz, February 10, 1954, Box 90, Folder 4, MFHI.
[23] Schwartz to Friedman, October 31, 1951, and Friedman to Schwartz, October 5, 1951, Box 90, folder 3, MFHI. The student could have been Eugene Lerner or James Kindhal, both of whom published on the period a few years later (Lerner 1956; Kindahl 1961). Or it may have been Phillip Cagan: Cagan (1965), in which he acknowledged benefitting from the "work and suggestions" of Friedman and Schwartz, began in 1875, the date of the Resumption Act.
[24] Anna Schwartz to Phillip Cagan, February 23, 1956, Box 91, Folder 1, MFHI.

number of times when I have come back to it after interruptions." She did not point out that compared to her outline of two years earlier, the book had lengthened, from a starting point in 1907 to one in 1897. True, Friedman had earlier suggested they bring in Federal Reserve estimates from as early as 1892, to complement "our stuff." However, Schwartz was layering together these materials in a more comprehensive way.[25]

Surviving letters do not capture the eureka moment when the duo shifted to the capacious, integrated analytic narrative that makes *A Monetary History* so distinctive. We can see, however, that Friedman did not intend to write a book of history, but Schwartz seems to have wanted to do so all along. As Friedman later admitted, "Anna, with her economic history background, reinforced [Stewart's] suggestion, that it was important to have a historical background before we got started on a primarily statistical study" (Brunner and Friedman 1989, p. 249).

When the scholars began seriously drafting chapters for *A Monetary History*, Friedman's earlier idea of the two authors working separately was quickly abandoned. From their correspondence, it appears that Schwartz frequently—but not always—wrote the first draft, which Friedman then commented on and mailed back. Summer was prime time for writing, with Friedman sequestered in his New Hampshire vacation home and Schwartz relatively close in New York City. While she did not visit him in New Hampshire, Friedman's trip from Chicago at the start of summer created a regular in-person touch point. Friedman encouraged Schwartz to be forthright in her criticism, and she was. After drawing up a particularly long and critical response to one of Friedman's draft chapters, she worried, "I hope you will not think these comments come for me with poor grace." Friedman was undeterred: "I very much want your frank criticism and the more destructive the better . . . After all, this is going to come out under your name and you will have to share the responsibility for it."[26] By 1956, the scholars had shifted to a deeply collaborative method of writing, with the idea of separate lanes set aside. Friedman had communicated his intent to treat Schwartz as a full coauthor, and the scale of the book was expanding significantly.

With the assurance in hand that her contribution would be recognized, Schwartz began drafting more boldly, and her writing began to influence Friedman. He chided her when one section came in "much too long in its present form for the present purpose: it is too long to provide a convenient summary and background for the estimates and the analysis that follows; yet too short to permit of a really thorough analysis of the individual incidents." Yet less than a week later, he confessed that "my face is red. Having started out with the objective of greatly reducing the size of your chapter 2, I herewith submit the first section, which is, unless I am

[25] Schwartz to Friedman, July 5, 1956, Box 90, Folder 1, MFHI. Friedman to Schwartz, February 10, 1954, Box 90, Folder 4, MFHI.

[26] Anna Schwartz to Friedman, August 17, 1956, and Friedman to Schwartz, August 21, 1956, Box 90, Folder 1, MFHI.

mistaken, rather longer than your draft!" While Friedman still hoped to cut the material down, the template set by Schwartz was having an effect.[27]

If existing correspondence does not capture the moment when the two settled into a more expansive narrative history dating to the Civil War era, it does show Schwartz enjoying the aftermath. Sending back material that both authors had revised once before, Schwartz added, "I must tell you how much I have enjoyed working through the resumption section. Except for the question about the gold reserve ratio, your piece makes the whole so much more comprehensible than anything now in print. I hope you will not cut any of it. I shall now get on to the section on silver."[28] The Civil War era, in particular, was a place Schwartz felt confident about her own take in the material, bombarding Friedman with questions and challenges regarding greenbacks and gold. Her letters are thick with references to scholars to whom she had spoken, primary sources and data, and other historical accounts.

Back and forth, the book emerged. It is organized around institutional formations and political events, linked by interwoven histories and recurring dynamics. While attentive to the shifts and changes in each time-period, terms and definitions introduced at the start of the book recur. A reader could skip around, but the analysis that emerges draws fully on the complete historical record. For example, a central finding introduced in the book's early pages is that adjusted for population growth, the US money supply grew at an annual rate of 3.7 percent. This figure informed Friedman's later policy proposal, central to his monetarism of that time, that the money supply grow at a fixed rate. Although Friedman did not at first specify a growth rate, when pressed for details he suggested around 4 percent as ideal (Friedman and Schwartz 1963, p. 5).

As an economic historian, Schwartz had a fundamental comfort thinking about different monetary regimes, which would pay rich dividends in Friedman's career. Economists of their generation had lived through profound shifts in global monetary regimes, including the destruction of the classical gold standard amid the Great Depression and World War II, and the reestablishment of a modified standard in the Bretton Woods agreement. Despite this experience, Friedman's peers greeted his forays into history with disbelief. At one conference, MIT economist Robert Solow derided Friedman for focusing on "Emperor Diocletian . . . the German and the Bolshevik and the Brazilian inflation . . . It really is not clear to me why extreme instances like this, or in what way extreme instances like this, are at all relevant to the problem that economic policy faces right now" (as quoted in Shultz and Aliber 1966, p. 63).

With Schwartz, however, Friedman had a peer who understood that the monetary arrangements of today were likely to be gone tomorrow, just like Confederate money and greenbacks. This consciousness fed into Friedman's 1951 paper on floating exchange rates, which envisioned a world beyond Bretton Woods; his insistence on the difference between nominal and real interest rates (a distinction the

[27] Friedman to Schwartz July 6, 1957, and Friedman to Schwartz July 14, 1957, Box 91, Folder 1, MFHI.
[28] Schwartz to Friedman, September 19, 1957, Box 91, Folder 2, MFHI.

Fed ignored for most of his career); and his prediction and theoretical explanation of stagflation. Thus, when Schwartz blasted Friedman's take on gold reserve ratios from the late nineteenth century, she was also blasting away the short-term thinking that marked much of the profession.

At this high peak of productivity and satisfaction, Schwartz received a troubling phone call. Nearly 30 years earlier—in 1932—she had enrolled in Columbia's graduate program in economics, and received a master's degree in 1934. Since then, she had been one of three authors on a three-volume study of the British economy, and begun work as a professional economist, including more than 16 years as a researcher at the National Bureau of Economic Research. All this time, she had laid claim to a cubicle in Columbia's Butler Library. Now the chairman of the economics department notified her that unless she registered a dissertation with a supervisor, the cubicle would be reassigned.

The issue was not the cubicle. The issue was the doctorate, which Schwartz had hoped to gain long ago. "Many years ago when I worked on the British business cycle study, the expectation was that I would offer as a dissertation the part of the manuscript for which I was primarily responsible," she told Friedman in February 1958, but "for many reasons this plan fell through." Now she wondered if she could offer some of their manuscript instead. Friedman thought this was a fine idea, encouraging her to submit the whole thing. "The preface can indicate that you have primary responsibility, both for the work and authorship of Part 2 and joint responsibility for the rest," he wrote. At this point, Part 2 referred to the statistical tables that would appear in 1970. Friedman was certain Columbia would accept the work, based in part on his own experience there. He wondered why she had not thought to do it years ago.[29]

A Kafka-esque ordeal then unfolded, where Schwartz was rebuffed for ever-shifting reasons. A proposal turned down by one professor was resurrected by another. Then came a demand that she pass an examination on the work, which Schwartz feared would be a hostile inquisition. Everything was put on hold while mimeographs were prepared, and then the process fizzled out. Friedman had put in a good word for her, and assumed the situation was resolved.

But five years later, as publication of *A Monetary History* approached, Schwartz was still trying to get her degree. Now she was told that a coauthored work was unacceptable as a dissertation—even though Friedman had a Columbia doctorate based on coauthored work. Finally came the excuse that because her "dissertation" was now in bound galleys at Princeton University Press, it was ineligible. This last pretext was flimsy enough to infuriate Friedman, who finally seemed to grasp it was not mere bureaucracy that Schwartz faced. Now it was his reputation on the line, too—was Columbia saying Friedman's magnum opus was not doctoral material? "As of this date, a degree conferred on you will honor Columbia more than it will honor you," Friedman wrote testily to Schwartz, resolving to call the department chair

[29] Schwartz to Friedman, February 19, 1958, and Friedman to Schwartz, February 26, 1958, Box 91, Folder 2, MFHI.

himself. As it turned out, the Columbia department was at that very moment trying to hire Friedman, giving him critical leverage as he demanded—successfully—that Schwartz receive her degree.[30]

Within a few months, Friedman's verdict had been borne out, as reviews of *A Monetary History* were widespread and rapturous. "A massive piece of erudition, exhibiting the highest scholarly qualities . . . a truly great book," judged Harrod (1963) in the *University of Chicago Law Review*. Even reviewers like Goodhart (1964, p. 314) who disagreed with the author's conclusions nonetheless were lavish in their praise; while scorning the quantity theory, his review in *Economica* dubbed the book "magnificent." Perhaps the most important verdict came from longtime Friedman critic and rival James Tobin (1965, p. 485) in a lengthy review published in the *American Economic Review*. "I have not done justice to the scope of this book," Tobin admitted, praising its "brilliance and finesse." He advised readers "in no event to omit the footnotes, which contain many gems of monetary theory," not realizing Schwartz had authored these. Tobin concluded with the ultimate academic complement: "This is one of those rare books that leave their mark on all future research on the subject." The book earned reviews across the disciplinary spectrum, with Kemmerer (1964, p. 195) in *American Historical Review* judging it "one of the most important books of our time." This reception was based not just on the book's ideas, but on their presentation in the analytic narrative that Schwartz had done so much to inspire and bring to reality. None of these positive reviews mentioned Schwartz. And none understood that without her, Friedman would have created a statistical report of interest only to other researchers.

Indeed, the book's appearance as a narrative—a story—meant it reverberated beyond academic audiences. Friedman and Schwartz's arresting argument that the Fed could—and should—have prevented the Great Depression focused new attention on the then secretive and relatively anonymous central bank. *Business Week* published a series of articles on *A Monetary History* and the reactions to it among economists, policymakers, and Fed officials, while Congress took a new interest in monetary policy, conducting hearings in 1964. As a *Washington Post* writer summarized, "[The book] is unsparing in its criticism of both the men and ideas that have been identified with the monetary establishment . . . Officials of the Federal Reserve System will hardly welcome this stout volume" (Segal 1963). Indeed, the book's publication forced change at the Federal Reserve, starting its transformation into an institution known for communication and transparency. Responding to Friedman's critique, the Fed began publishing more of its own data, including the monetary aggregates Friedman and Schwartz had constructed on their own. And in a sad irony, during the inflation of the 1970s Friedman became the most famous critic of the Fed, then chaired by his old friend Arthur Burns.

A Monetary History gave a great boost to Friedman's career. Among the book's more important readers was future Republican presidential candidate

[30] For more on this episode, see Burns (2023) and Friedman to Schwartz, December 10, 1963, Box 91, Folder 3, MFHI.

Barry Goldwater, who wrote Friedman an enthusiastic personal letter. "I think it is superb . . . Professors sometimes have the habit of writing only for other professors," Goldwater noted, "but your book is written in a way that the man on the street will understand and get your message."[31] Before long Friedman was serving as an adviser to Goldwater's campaign, a move that would kick off his career as a public intellectual with a regular column in *Newsweek*. Within economics, even Friedman's strongest detractors began to allocate money more importance in their thinking. Admirers calling themselves "monetarists" formed a distinct school that challenged the Keynesian consensus and pushed forward Friedman's proposal for a monetary growth rule. Eventually, this idea would evolve into two major frameworks that continue to shape monetary policy today: inflation targeting and Taylor rules (Taylor 2001). And a separate rational expectations school grew out of Friedman's emphasis on inflationary expectations (in this journal, Hall and Sargent 2018). In one form or another, these were all the children of *A Monetary History* (for more discussion, see Burns 2023, chap. 12, 15).

Over time, Schwartz would be widely recognized for her coauthorship of *A Monetary History*, and her star rose along with Friedman's. New collaborators sought her out, and admirers came to the New York offices of the National Bureau of Economic Research to meet her. She was one of three economists who founded the Shadow Open Market Committee, becoming an influential Fed watcher in the 1970s. During the Reagan administration, she chaired the US Gold Commission, bringing her historical knowledge to bear on renewed debates on the place of gold in the US monetary system. By the time of her death in 2012, nine universities had bestowed honorary doctorates on Schwartz.

It was fitting recognition for the indispensable role Schwartz played in the book's construction and conclusions. Its first and most basic argument—that money was fundamental to business cycles—could not have been reached without Schwartz. This was an empirical finding based on the monetary data she gathered and the analysis of them she and Friedman hammered out. Similarly, the reinterpretation of the Great Depression as a monetary phenomenon, and the indictment of the Federal Reserve's management at that time, could not possibly have been presented in the book's original conception. Only the embrace of the analytic narrative, made possible by Schwartz's partnership, made this argument viable. These chapters of the book, in particular, relied upon archival findings Schwartz dug out of the Columbia University library—near her cubicle (Nelson 2020, p. 31).

Conclusion

As for Schwartz's impact on Friedman's larger career, it is useful to circle back to citation on Friedman's 1976 Nobel prize, which identified three areas of work

[31] Barry Goldwater to Milton Friedman, July 13, 1962, box 27, folder 24, MFHI.

that had earned him the prize: "for his contribution to consumption analysis and to monetary history and theory, including his observations of the complexity of stabilization policy." Friedman's monetary history is impossible to imagine without Schwartz. As to monetary theory and stabilization policy, it is true she wasn't directly involved in the production of his major works in this area, such as the 1967 address to the American Economic Association, his *American Economic Review* article "The Lag in Effect of Monetary Policy," or his 1960 *Program for Monetary Stability*. Yet how plausible is it that Friedman would have ventured these bold theories without the grounding in monetary data and history Schwartz afforded him? Would he have come to see the lags in policy without a long-term perspective, or come to question the long run trade-off between inflation and unemployment? Would he had advocated a monetary growth rule for stabilization without nearly a hundred years of monetary data at his fingertips? Friedman championed an economic methodology that blended theory and empirical work, stressing that all theories must be validated against experience. And he used history as a sort of natural experiment, a testing ground that gave him clarity and confidence. His lengthy partnership with Schwartz enabled Friedman to formulate the new approaches, perspectives and ideas that the Nobel Prize celebrated.

The first area identified in the citation, "consumption analysis," refers to *A Theory of the Consumption Function*, a 1957 book interestingly similar to *A Monetary History* in its origins. Schwartz was not a collaborator on this work—rather Friedman developed the ideas in tandem with three other women, Rose Friedman, Dorothy Brady, and Margaret Reid, even calling it "a joint work" in the introduction (Burns 2022). By the time he started this project, Friedman had been working with Schwartz for years, and well knew the enormous benefit to collaborating with economists who his peers ignored because they happened to be women.

In the end, the effort to systemically disaggregate the contributions of Friedman and Schwartz to *A Monetary History* is futile. They were intertwined minds, working in tandem, feeding off one another, strengthening each other's weaknesses and buttressing each other's strengths.

■ *The author would like to thank Michael Bordo and David Laidler for helpful comments.*

References

Ahamed, Liaquat. 2009. *Lords of Finance: The Bankers Who Broke the World*. New York: Penguin Books.

Bernanke, Ben S. 1995. "The Macroeconomics of the Great Depression: A Comparative Approach." *Journal of Money, Credit, and Banking* 27 (1): 1–28.

Bernanke, Ben. 2002. "On Milton Friedman's Ninetieth Birthday." Speech, Conference to Honor Milton Friedman, University of Chicago, Chicago, IL, November 8, 2002. https://www.federalreserve.gov/boarddocs/speeches/2002/20021108/.

Brunner, Karl, and Milton Friedman. 1989. "Anna J. Schwartz: An Appreciation: A life of Scholarship." In *Money, History, and International Finance: Essays in Honor of Anna J. Schwartz*, edited by Michael D. Bordo, 245–50. Chicago: University of Chicago Press.

Burns, Jennifer. 2022. "Hidden Figures: A New History of the Permanent Income Hypothesis." *History of Political Economy* 54 (S1): 43–68.

Burns, Jennifer. 2023. *Milton Friedman: The Last Conservative* New York: Farrar, Straus, and Giroux.

***Business Week*.** 1963. "Theorizing for Goldwater?" November 23, 108.

Cagan, Philip. 1965. *Determinants and Effects of Changes in the Stock of Money, 1875–1960* New York: Columbia University Press.

Eichengreen, Barry. 1992. *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939*. New York: Oxford University Press.

Friedman, Milton. 1952. "Price, Income, and Monetary Changes in Three Wartime Periods." *American Economic Review* 42 (2): 612–25.

Friedman, Milton. 1956. "The Quantity Theory of Money—A Restatement." In *Studies in the Quantity Theory of Money*, edited by Milton Friedman. Chicago: University of Chicago Press.

Friedman, Milton. 1962. "The Interpolation of Time Series by Related Series." *Journal of the American Statistical Association* 57 (300): 729–57.

Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.

Friedman, Milton, and Anna J. Schwartz. 1970. *Monetary Statistics of the United States: Estimates, Sources, Methods*. New York: Columbia University Press.

Friedman Milton, and Anna J. Schwartz. 1982. *Monetary Trends in the United States and United Kingdom: Their Relations to Income, Prices, and Interest Rates 1867–1975*. Chicago: University of Chicago Press.

Gayer, Arthur, W. W. Rostow, and Anna J. Schwartz. 1953. *The Growth and Fluctuation of the British Economy, 1790–1850: An Historical, Statistical, and Theoretical Study of Britain's Economic Development*. Oxford: Clarendon Press.

Goodhart, C. A. E. 1964. Review of *A Monetary History of the United States, 1867–1960*, by Milton Friedman and Anna J. Schwartz. *Economica* 31 (123): 314–18.

Hall, Robert E., and Thomas J. Sargent. 2018. "Short-Run and Long-Run Effects of Milton Friedman's Presidential Address." *Journal of Economic Perspectives* 32 (1): 121–34.

Harrod, Roy F. 1963. Review of *Friedman and Schwartz: A Monetary History of the United States*, by Milton Friedman and Anna J. Schwartz. *University of Chicago Law Review* 31: 188–96.

Kemmerer, Donald L. 1964. Review of *A Monetary History of the United States, 1867–1960*, by Milton Friedman and Anna J. Schwartz. *American Historical Review* 70 (1): 195–97.

Kindahl, James K. 1961. "Economic Factors in Specie Resumption in the United States, 1865–79." *Journal of Political Economy* 69 (1): 30–48.

Laidler, David E. W. 1991. *The Golden Age of the Quantity Theory*. Princeton: Princeton University Press.

Lerner, Eugene. 1956. "Inflation in the Confederacy, 1861–65." In *Studies in the Quantity Theory of Money*, edited by Milton Friedman. Chicago: University of Chicago Press.

Lothian, James R., and George S. Tavlas. 2018. "How Friedman and Schwartz Became Monetarists," *Journal of Money, Credit, and Banking* 50 (4): 757–87.

Lundberg, Eric. 1976. "Award Ceremony Speech." Speech, Nobel Prize. https://www.nobelprize.org/prizes/economic-sciences/1976/ceremony-speech/.

McElroy, Marjorie B. 2013. "A Celebration of the Life of Anna Jacobson Schwartz." *American Economic Association Committee on the Status of Women in the Economics Profession* (*CSWEP*) *News*: 1.

Nelson, Edward. 2004. "An Interview with Anna J. Schwartz." *Macroeconomic Dynamics* 8 (3): 395–417.

Nelson, Edward. 2020. *Milton Friedman and Economic Debate in the United States 1932–1972*. Vol. 2. Chicago: University of Chicago Press.

Royal Swedish Academy of Sciences. 1976. "Press Release: This Year's Economics Prize to an American." October 14. https://www.nobelprize.org/prizes/economic-sciences/1976/press-release/.

Schultz, George P., and Robert Z. Aliber. 1966. *Guidelines, Informal Controls, and the Market Place: Policy Choices in a Full Employment Economy*. Chicago: University of Chicago Press.

Schwartz, Anna J. 1993. "Interview with Anna J. Schwartz." By David Fettig. *The Region*. https://minneapolisfed.org/publications/the-region.

Schwartz, Anna J. 1995. "An Interview with Anna J. Schwartz." By Eugene N. White. *Newsletter of the Cliometric Society* 10 (2): 3–7.

Segal, Harvey H. 1963. "Economic Scene: Some US Monetary Blunders." *Washington Post,* November 25, D11.

Tavlas, George S. 2023. *The Monetarists: The Making of the Chicago Monetary Tradition, 1927–1960*. Chicago: University of Chicago Press.

Taylor, John B. 2001. "An Interview with Milton Friedman." *Macroeconomic Dynamics* 5 (1): 101–31.

Tobin, James. 1965. "The Monetary Interpretation of History." *American Economic Review* 55 (3): 464–85.

Journal of Economic Perspectives—Volume 38, Number 4—Fall 2024—Pages 229–236

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

Five years ago, the Institute for Fiscal Studies launched a mammoth project called "Inequality: IFS Deaton Review" (https://ifs.org.uk/inequality/). An "Evidence Volume" has now been published, with a mix of 81 (!) chapters, commentaries, articles, and reports divided into 18 (!) categories: (1) what's wrong with inequality; (2) attitudes toward inequality; (3) trends in economic inequality; (4) history of inequality; (5) political inequality; (6) gender; (7) race; (8) immigration; (9) health; (10) geography; (11) families; (12) early childhood; (13) immigration; (14) the labour market; (15) firms; (16) trade and globalization; (17) top income inequality and tax policy; and (18) benefits and public services. In the very first essay, Debra Satz and Stuart White ask: "What is wrong with inequality?" "Because we point to a plurality of reasons to be concerned with inequality, our account is complex. This

■ *Timothy Taylor is Managing Editor, Journal of Economic Perspectives, based at Macalester College, Saint Paul, Minnesota.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.38.4.229.

arguably makes our approach less tractable than more simple one-dimensional alternatives. For example, it is surely simpler to adopt a single focus on income inequality without looking at its effects on specific institutions or particular aspects of life such as health, social relations, and political influence. But we believe that it is possible and better to develop richer models for measuring those aspects of inequality that matter. This can lead, in some cases, to more tailored policies than income transfers, such as benefits in kind (although income transfers are often better) or different ways of accomplishing income transfers. It might suggest the need to focus on wealth as well as income. It might suggest the need to focus on policies that shape the associational context of economic and political life, such as levels and patterns of unionization. And it is worth bearing in mind that one person's simplification for the sake of tractability is another person's life."

As a complementary effort, the June 2024 issue of *Fiscal Studies* offers a ten-paper special issue on "Changing labour market and income inequalities in Europe and North America: a parallel project to the IFS Deaton Review of Inequalities in the 21st century" (https://onlinelibrary.wiley.com/toc/14755890/2024/45/2). Bradley L. Hardy, Elizabeth Krause, and James P. Ziliak contribute "Income inequality in the United States, 1975–2022." "Overall after-tax and transfer income inequality increased more than 25 per cent since the mid-1970s, and by as much as 50 per cent when comparing the 90th and 10th percentiles. While there has been substantial upgrading in formal education credentials among both men and women—an inequality-reducing development—those with fewer credentials have increasingly been less likely to work and marry, each of which could result in higher inequality. The latter effects are exacerbated by those selecting into marriage and cohabitation being more likely to partner with those holding similar educational credentials and earning power. Moreover, the decline in work among the less skilled coincided with the transformation of the safety net to rewarding work. These demographic and policy changes have resulted in a pulling apart of the US income distribution."

The International Comparison Project at the World Bank has released the 2021 version of its "purchasing power parity" (PPP) measures for the global economy (May 2024, https://www.worldbank.org/en/programs/icp/data) . From the "Highlights of Main Findings": "PPPs convert different currencies to a common currency and, in the process of conversion, equalize their purchasing power by controlling for differences in the price levels of goods and services between economies. PPPs allow international comparisons of GDP and its components that avoid the over- or under-estimation of economic output that is inherent in market exchange rate-based comparisons, as the latter does not adjust for price levels. PPP-based estimates are also not prone to fluctuations in market exchange rates. PPPs are calculated by the ICP based on the prices of items within a common basket of goods and services and expenditure shares, used as expenditure weights, on groups of items in each of the participating economies." Using the PPP measures, "The largest economy in the world in 2021 was China, recording a PPP-based GDP of $28.8 trillion, reflecting 18.9 percent of the global GDP. The United States was the second largest, with

nearly $23.6 trillion or 15.5 percent of the global GDP. India's economy was the third largest at $11. trillion, accounting for 7.2 percent."

Viral V. Acharya, Nicola Cetorelli, and Bruce Tuckman have written "Where Do Banks End and NBFIs Begin?" (from the Sverges Riksbank at https://www.riksbank.se/globalassets/media/konferenser/2023/session-5-acharya_cetorelli_tuckman_where_do_banks_end_and_nbfis_begin_draft.pdf). "Non-bank financial intermediaries (NBFIs) have surpassed banks as the largest global financial intermediaries. And yet, most NBFIs continue to be lightly regulated relative to banks for safety and soundness, whether in terms of capital and liquidity requirements, supervisory oversight, or resolution planning. . . . [T]he global financial assets of NBFIs have grown faster than those of banks since 2012, to about $239 trillion and $183 trillion in 2021, respectively. In percentage terms, the share of the NBFI sector has grown from about 44% in 2012 to about 49% as of 2021, while banks' share has shrunk from about 45% to about 38% over the same time period. . . . As in the global data, NBFIs in the United States have accumulated substantially more assets than banks over the period shown. However, the NBFI sector in the United States accounts for a much higher share of financial assets, which was over 60% in 2021."

David McKenzie writes of "Fears and Tears: Should More People Be Moving within and from Developing Countries, and What Stops this Movement?" (*World Bank Research Observer*, February 2024, 39:1, 75–96, https://academic.oup.com/wbro/article-abstract/39/1/75/6982896). McKenzie estimates that only "one in seven of the world's population have ever migrated," either internationally or between regions within a country, and offers a "fears and tears" explanation to address the puzzle of why moving is not more widespread. "I think economists have devoted far less attention to what I call fears, which is the enormous uncertainty associated with migration that is difficult to quantify. This type of unquantifiable uncertainty, also known as Knightian uncertainty, may include fears about the safety conditions at destination . . ., the ability to make friends and fit in, about whether one will like living in the new location, and soon. With Knightian uncertainty, there is no unique probability distribution of possible outcomes of employment, wages, and amenities . . . Bewley (2002) argues that in such cases, there can be a bias towards inertia. . . . I think a bigger part of the reason for tears upon moving and not having these same tears upon deciding to stay is the inability to picture what you are giving up when you do not move . . . Gabaix and Laibson (2017) argue that people have only noisy information about the future and that it is harder to forecast the further into the future one looks, with these features causing individuals to behave as if they have very myopic preferences, preferring the present."

Those who would like more on industrial policy, in addition to the symposium in this issue, can turn to three articles in the 2024 *Annual Review of Economics* (https://www.annualreviews.org/content/journals/economics/browse). Réka Juhász, Nathan Lane, and Dani Rodrik explore "The New Economics of Industrial Policy" (pp. 213–42): "The salience of industrial policy has risen greatly in recent years, as governments have increasingly engaged in self-conscious industrial policies as they address a variety of problems—the green transition, the resilience of supply chains, the

challenge of good jobs, and geopolitical competition with China. . . . We focus on three types of cases: episodes of infant industry promotion (e.g. in textiles, shipbuilding, and heavy industries), large-scale public R&D efforts (as in the space race between the United States and Soviet Union), and selective place-based policies targeting specific industries (as in the US manufacturing drive during World War II and contemporary regional European subsidies). Interestingly, the most recent vintage of papers, paying serious attention to identification and observability difficulties, produces results that are much more favorable to industrial policy." Chad P. Bown discusses "Modern Industrial Policy and the World Trade Organization" (pp. 243–70): "To remain relevant in the international trading system, the World Trade Organization (WTO) may need its members to engage directly over the issue of industrial policy. The staff at the major international organizations—the International Monetary Fund (IMF), the Organisation for Economic Co-operation and Development (OECD), the World Bank, and the WTO—have put out an explicit plea for a renewed work program and policymaker engagement on the issue. . . ." Réka Juhász and Claudia Steinwender examine "Industrial Policy and the Great Divergence" (pp. 27–54): "We revisit the historical track record of industrial policy in the context of the nineteenth and early twentieth century—a critical juncture in economic history. . . . In particular, while much attention in economics has been paid to the developmental effects (or lack thereof) of protective import tariffs in the nineteenth century, our review of recent work suggests that tariffs were neither the only nor perhaps the most important policy lever in countries' industrial policy tool kit. Rather, many independent countries deployed a multitude of complementary policies that foreshadow modern industrial policy, such as state-led technology acquisition, human capital development, intellectual property rights (IPR) protection, low industrial input tariffs, and subsidies for prioritized activities. We also highlight an aspect of the nineteenth century which cannot and should not be ignored, namely, that colonial powers used colonies in the service of their own industrial development goals."

Economic History

Maristella Botticini delivered the 2023 Presidential Address to the European Economic Association, drawing on a research paper written with Pietro Buri and Massimo Marinacci, titled "The Beauty of Uncertainty: The Rise of Insurance Contracts and Markets in Medieval Europe" (*Journal of the European Economic Association*, 21:6, December 2023, 2287–326, https://academic.oup.com/jeea/article/21/6/2287/7319353; video at https://www.youtube.com/watch?v=IZzTHeIrGGw). From the abstract: "Maritime insurance developed in medieval Europe is the ancestor of all forms of insurance that appeared subsequently. . . . [W]e show that medieval merchants had to bear more frequently natural risks (they traveled longer distances) and new human risks with unknown probabilities (they faced unpredictable attacks by corsairs due to increased political fragmentation and commercial competition in Europe). The increased demand for protection in medieval seaborne

trade met the supply of protection by a small group of wealthy merchants with a broad information network who could pool risks and profit from selling protection through a novel business device: the insurance contract. A new market—the market for insurance—was then born. Next, analyzing more than 7,000 insurance contracts redacted by notaries and about 100 court proceedings housed in the archives of Barcelona, Florence, Genoa, Palermo, Prato, and Venice, we study the main features of medieval trade, the type of risks faced by merchants, and the characteristics of insurance contracts and markets from 1340 to 1500."

James Feigenbaum and Daniel P. Gross have been studying a classic example of US jobs lost to automation: telephone switchboard operators. In "Answering the Call of Automation: How the Labor Market Adjusted to Mechanizing Telephone Operation" (*Quarterly Journal of Economics*, August 2024, 139:3, pp. 1879–939, https://academic.oup.com/qje/article/139/3/1879/7614605), they explore the period from 1920 and 1940: "[W]e show that after a city was cut over to mechanical [switchboard] operation, the number of 16- to 25-year-old women in subsequent cohorts employed as telephone operators immediately fell by 50% to 80%. These jobs made up around 2% of employment for this group, and even more for those under age 20—and given turnover rates, this shock may have foreclosed entry-level job opportunities for as much as 10% to 15% of peak cohorts. The effect of this shock on incumbent operators was to dispossess many of their jobs and careers . . . In contrast, however, automation did not reduce employment rates in subsequent cohorts of young women, who found work in other sectors—including jobs with similar demographics and wages (such as typists and secretaries), and some with lower wages (such as food service workers). Feigenbaum and Gross also consider the perspective of AT&T in "Organizational and Economic Obstacles to Automation: A Cautionary Tale from AT&T in the Twentieth Century" (*Management Science*, published online, February 27, 2024, https://pubsonline.informs.org/doi/abs/10.1287/mnsc.2022.01760). "Although manual switching served early telephone networks well, expansion revealed its limits, as its complexity rose quickly in large markets with billions of possible connections, and switchboards became system bottlenecks. As AT&T grew, its service quality thus fell, and operator requirements exploded: by the 1920s AT&T was the largest U.S. employer, with operators over half its workforce. Company records show the limits of manual switching were known as early as the 1900s, when automatic technology was already being tested—yet it took AT&T several more decades to adopt it widely. We show in this paper that automation was hindered by interdependencies between call switching and the rest of AT&T's business: the mechanization of call switching required complementary innovation and adaptation across the firm, which were only resolved over time."

William Deringer provides an early practical example of present discounted value calculations in "Mr. Aecroid's Tables: Economic Calculations and Social Customs in the Early Modern Countryside," *Journal of Modern History*, March 2024, 96:1, https://www.journals.uchicago.edu/doi/10.1086/728594). From the abstract: "In the 1610s and 1620s, a new computational technology took hold in England: printed mathematical tables for compound interest and discounting ('present

value') problems. Historians of finance and accounting have long recognized these paper tools as predecessors of essential modern techniques like 'discounted cash flow.' Yet the history of these tables remains hazy. . . . [A]mong the leading 'early adopters' were institutions of the Church of England. Amidst the inflation of the early modern 'price revolution,' bishops, cathedrals, and colleges confronted a complex of economic, political, and social pressures. Mathematical tables like Acroyd's emerged out of long-running conflicts between church landlords and tenants over how to determine just and reasonable fines on church lands."

Construction Costs

Zachary Liscow, Will Nober, and Cailin Slattery discuss "Procurement and Infrastructure Costs," in a paper presented at the 13th annual Municipal Finance Conference at the Brookings Institution (July 11, 2024, https://www.brookings.edu/events/13th-annual-municipal-finance-conference/). "The United States spends a large amount on infrastructure costs: state and local governments spent $266 billion on highways alone in 2022. The spending, on a per-project basis, is very high by international standards—over three times as high as other upper- and middle-income countries." The authors conduct a survey of state-level Departments of Transportation (DOTs) as well as the companies that bid to build roads. "States that flag concerns about consultant costs have higher costs—a one standard deviation increase in reported consultant costs is associated with an almost 20% increase ($70,000) in cost per lane-mile. States where contractors and procurement officials expect more change orders have significantly higher costs: one additional change order correlates with $25,000 in additional cost per lane-mile at the mean. . . . More directly, we find that states with (perceived) higher quality DOT employees have lower costs. A state with 'neither low nor high quality' employees has almost 30% higher costs per mile than one that rates the DOT employees as 'moderately high quality', all else equal. . . . A one standard deviation increase in DOT employment per capita is correlated with 16% lower costs. . . . [W]e find that an additional bidder on a project is associated with 8.3% lower costs, or a savings of approximately $30,000 per lane-mile ($460,000 for the average project)."

Stephen Smith reports on costs of installing "Elevators" (Center for Building in North America, May 2024, https://static1.squarespace.com/static/634dfe3176afcc36f569d83d/t/6689cb0e8ac6370940a122ff/1720306458871/Elevators.pdf). "Single-family houses aside, the United States has over 32 million apartments, while Spain has fewer than 13 million apartments but about the same number of elevators. The U.S. has 40 percent fewer elevators per capita than the Netherlands, despite 30 percent of the American housing stock being in multifamily dwellings (and 19 percent in buildings with at least 10 units), compared to a total multifamily housing share of just 21 percent in the Netherlands. New York City has roughly the same population as Switzerland and even more New Yorkers live in apartment buildings than Swiss residents do, but New York only has half the

number of passenger elevators. No matter how you slice the numbers, America has fallen behind on elevators." The report investigates how labor costs and building codes across countries drive these differences.

Interviews

Tyler Cowen interviews Joseph Stiglitz on his "Conversations with Tyler" podcast: "Joseph Stiglitz on Pioneering Economic Theories, Policy Challenges, and His Intellectual Legacy" (June 26, 2024, https://conversationswithtyler.com/episodes/joseph-stiglitz/). As Cowen mentions, Joe's CV now runs to 153 pages, which "is neither complete nor really has any chaff." For a sample, here's Stiglitz on a 1980 paper written with Sandy Grossman: "The title of that paper was 'On the Impossibility of Informationally Efficient Markets.' It was an argument against the view that was held by people like Eugene Fama that markets were informationally efficient, that they transmitted efficiently all the information from the informed to the uninformed. We made the obvious observation that if that were the case, there would be no incentive for anybody to gather information. So the market might be transmitting information, but it would be all free information. It would be information that nobody had done any work to collect. That idea, actually, in another context worries me very much today, that with Google and AI scraping so much information off of our newspapers, off of our podcasts, off of everything they can get a hold of, they're trying to appropriate the value of the knowledge that's been created by other people without paying for it. If they succeed in doing that, of course, that will decrease the incentives for others to produce information of high quality and of value. It's that kind of interaction that was at the heart of our 1980 paper, and the themes that we talked about there are still the critical themes that we're talking about today."

Janet Bush of the McKinsey Global Institute interviews Chad Syverson in "Unpacking the Mysteries of the Global Economy" (July 2, 2024, audio and text at https://www.mckinsey.com/mgi/forward-thinking/unpacking-the-mysteries-of-productivity). Bush: "There's a perception that productivity means efficiency and lost jobs. I remember somebody said to me, 'Oh, productivity—you're fired.' Unpack that for us." Syverson: "That is an example of the fallacy of reasoning causality from an accounting identity. There are many specific ways to measure productivity, but they're all basically ratios of output to input, how much comes out of a production process divided by how many inputs go into it. And the notion that you're describing with that person's comment comes from looking at that definition and thinking, 'Oh, that's how you causally affect productivity. So OK, I want productivity to be higher. It's outputs over inputs, so if I make inputs smaller, productivity will go up.' Well, the problem with that is—and this is true whenever you reason from an accounting identity—it's not just inputs that are changing when you decide to cut inputs. You know, those inputs are doing something, presumably, and you're going to affect what they're doing if you try to cut those inputs, like, say, workers

or worker hours. And that might be useful stuff that makes output. . . . You know, I totally understand sort of the sentiment behind what that person said. I hear it a lot, but it's usually in that direction, the messing up the identity for causality. Because if someone said, 'Oh, I need productivity to go up, I know what I'll do, I'll just make more output'—if you said that to someone, they'd say, 'OK, what magic wand do you have that lets you wave and get more output for nothing?'"

Discussion Starters

Alexander Budzier and Bent Flyvbjerg present "The Oxford Olympics Study 2024: Are Cost and Cost Overrun at the Games Coming Down?" (May 2024, University of Oxford Said Business School, Working paper | 2023–24, https://www.politico.eu/wp-content/uploads/2024/07/15/The_Oxford_Olympics_Study_2024_Are_Cost_and_Cost_O_240715_145740_cleaned.pdf). "Given that the last three Summer Games cost USD 51 billion (in 2022 prices) and overran budgets by 185% in real terms—not including road, rail, airport, hotel, and other infrastructure, which often cost more than the Games themselves—the financial size and risks of the Games warrant study. . . . The Paris 2024 games, for instance, have seen costs surge from EUR 3.6 billion to 8.8 billion. Similarly, Los Angeles 2028 has revised its forecast from USD 5.3 billion to 6.8 billion. . . . For instance, cost overrun and associated debt from the Athens 2004 Games weakened the Greek economy and contributed to the country's deep financial and economic crises, beginning in 2007 and still playing out almost a decade later. For Rio 2016, the Brazilian economy was doing well when the city bid for the Olympics. Fast forward a decade to two months before the opening ceremony and this was no longer the case. Rio was now in such dire straits that the governor declared a state of emergency to secure additional funding for the Games from money reserved for dealing with natural and other disasters."

Stephanie M. Blalock, Kevin McMullen, Stefan Schöberlein, and Jason Stacy tell the story of "'One of the Grand Works of the World': Walt Whitman's Advocacy for the Brooklyn Waterworks, 1856–59" (*Technology and Culture*, January 2024, 65:1, 237–63, https://muse.jhu.edu/pub/1/article/920522). From the abstract: "When the Brooklyn Waterworks opened in 1859, it was one of America's most advanced water and sewer systems. Yet after Brooklyn was annexed by New York City, the waterworks' history slipped into obscurity, despite having a now-famous champion: the 'poet of America,' Walt Whitman, whose brother worked on the project. This article shows the Brooklyn poet's fierce, multiyear lobbying effort for the waterworks in various newspapers and introduces a wealth of newly recovered Whitman writings on the issue. As a journalist, Whitman exemplifies the nineteenth-century press as an intermediary between expert engineers and popular readers. The poet brought precise expertise, translated engineers' technical arguments into everyday language for his readers, and fought the resulting day-to-day political battles over construction in print. Whitman, then, is an underappreciated case study of the confluence of technology, public health, and local journalism."

MWI

AMERICAN ECONOMIC ASSOCIATION

Meet the AEA Ombuds Team

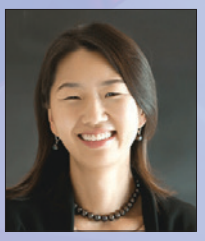
The American Economic Association Ombuds Team is an **independent, impartial, confidential, and informal** third-party resource to help AEA members obtain information about AEA policies and procedures, discuss issues, and consider options for resolving conflicts.

The Ombuds Team features three ombuds professionals with a range of experiences and specialties who can provide informal guidance on the evaluation of any interpersonal challenges, including sexual harassment or discrimination issues, and determining how, when, or if a formal complaint to leadership may be an appropriate path.

Affected AEA members who are not quite ready to discuss or file a formal complaint also have the option to use a separate and secure reporting lockbox to log allegations of harassment, discrimination, and retaliation involving other members via an encrypted, time-stamped online tool.

Contact your Ombuds Team to learn more:

Brent Epperson
brent-ombuds
@mwi.org



Charmhee Kim
charmhee-ombuds
@mwi.org

Nnena Odim
nnena-ombuds
@mwi.org

The Ombuds Team provides regular educational sessions throughout the year on topics of interest to AEA members.

For more details about upcoming workshops, please visit **www.aeaweb.org/ombuds**.

Learn more about the role of the Ombuds Team resource and how it works to meet your needs.

JOE NETWORK

The career network designed by economists—for economists.

The AEA's JOE Network is the preferred hiring tool for the economics job market, matching qualified candidates with relevant economics positions in leading institutions.

Job Seekers

Showcase your profile to hundreds of hiring managers and quickly apply to your preferred positions. The JOE Network allows you to securely request reference letters and share job market materials.

Employers

Target search requirements to locate your preferred candidates, set up hiring committee access, and easily manage job postings and applications.

Faculty Letter-Writers

Reply to reference requests with default or customized letters, assign proxies, and easily monitor task completion. Over 150,000 requests are fulfilled on the JOE Network each year.

Use the JOE Network as your comprehensive tool for the economics job market!

AMERICAN ECONOMIC ASSOCIATION

aeaweb.org/JOE


ASSA
ASSA
ALLIED SOCIAL SCIENCE ASSOCIATIONS
Annual Meeting
January 3–5, 2025
San Francisco, CA
Mark Your Calendars!
• Over 500 sessions covering important topics in economics research
• New “lightning round” sessions, mentoring and networking luncheon, and lectures on recent developments in economics
• Distinguished speakers
• Unique opportunities to connect and collaborate
Registration Now Open
www.aeaweb.org/conference
2025

The American Economic Association

Correspondence relating to advertising, business matters, permission to quote, or change of address should be sent to the AEA business office: aeainfo@vanderbilt.edu. Street address: American Economic Association, 2014 Broadway, Suite 305, Nashville, TN 37203. For membership, subscriptions, or complimentary access to JEP articles, go to the AEA website: http://www.aeaweb.org. Change of address notice must be received at least six weeks prior to the publication month.

The following Statement of Ownership, Management, and Circulation is provided in accordance with the requirements, as contained in 39 U.S.C. 3658. Journal of Economic Perspectives is owned, managed, and published by the American Economic Association, a nonprofit educational organization, located at 2014 Broadway, Suite 305, Nashville, Davidson County, TN 37203-2425. The Editor is Heidi Williams. The Managing Editor is Timothy Taylor, AEA Publications, 2403 Sidney Street, Suite 260, Pittsburgh, PA 15203. The tax status of the American Economic Association has not changed during the preceding twelve months. During the preceding twelve months, the average number of copies printed for each issue was 2,620; the average total paid and/or requested circulation was 2,225; the average total non-requested distribution was 0; the average number of copies not distributed was 396. Corresponding figures for May 2024, the issue reported for the filing date: total number of copies printed, 2,310; total paid and/or requested circulation, 1,955; total non-requested distribution, 0; number of copies not distributed 355. During the preceding twelve months, the average number of requested and paid electronic copies for each issue was 905; the total average requested, and paid print and electronic copies was 3,130. Corresponding figures for May 2024, the issue reported for the filing date: number of requested and paid electronic copies, 910; the total requested and paid print and electronic copies, 2,865. Certified by Barbara Fiser, Director of Finance and Administration.

EXECUTIVE COMMITTEE

Elected Officers and Members

President
JANET CURRIE, Princeton University

President-elect
LAWRENCE F. KATZ, Harvard University

Vice Presidents
NICHOLAS BLOOM, Stanford University
EBONYA L. WASHINGTON, Columbia University

Members
AMANDA BAYER, Swarthmore College
MELVIN STEPHENS, JR., University of Michigan
KEVIN LANG, Boston University
LISA M. LYNCH, Brandeis University
AMANDA E. KOWALSKI, University of Michigan
LARA D. SHORE-SHEPPARD, Williams College

Ex Officio Members
CHRISTINA ROMER, University of California, Berkeley
SUSAN C. ATHEY, Stanford University

Appointed Members

Editor, The American Economic Review
ERZO F.P. LUTTMER, Dartmouth College

Editor, The American Economic Review: Insights
MATTHEW GENTZKOW, Stanford University

Editor, The Journal of Economic Literature
DAVID H. ROMER, University of California, Berkeley

Editor, The Journal of Economic Perspectives
HEIDI WILLIAMS, Dartmouth College

Editor, American Economic Journal: Applied Economics
BENJAMIN OLKEN, Massachusetts Institute of Technology

Editor, American Economic Journal: Economic Policy
LUCAS W. DAVIS, University of California, Berkeley

Editor, American Economic Journal: Macroeconomics
AYŞEGÜL ŞAHIN, Princeton University

Editor, American Economic Journal: Microeconomics
NAVIN KARTIK, Columbia University

Secretary-Treasurer
PETER L. ROUSSEAU, Vanderbilt University

OTHER OFFICERS

Director of AEA Publication Services
ELIZABETH R. BRAUNSTEIN

Counsel
LAUREN M. GAFFNEY, Bass, Berry & Sims PLC, Nashville, TN

ADMINISTRATORS

Director of Finance
ALLISON BRIDGES

Director of Administration
BARBARA H. FISER

Convention Manager
REBEKAH LOFTIS

The Journal of
Economic Perspectives

Fall 2024, Volume 38, Number 4

Symposia

Industrial Policy

Behavioral Incentive Compatibility

Articles

Features

AMERICAN ECONOMIC ASSOCIATION