NBER WORKING PAPER SERIES

LIFE AND DEATH AT THE MARGINS OF SOCIETY: THE MORTALITY OF THE U.S. HOMELESS POPULATION

Bruce D. Meyer
Angela Wyse
Ilina Logani

Working Paper 31843
http://www.nber.org/papers/w31843

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2023

We are indebted to Charles Hokayem, Josh Mitchell, Nikolas Pharris-Ciurej, Mandana Vakil, and John Voorheis for their assistance, and Margot Kushel and Derek Wu for their comments. We are grateful to the Alfred P. Sloan Foundation, the Russell Sage Foundation, the Charles Koch Foundation, the Menard Family Foundation, the National Institute on Aging, and the American Enterprise Institute for their support of the Comprehensive Income Dataset project. Disclaimer: This paper is released to inform interested parties of research and to encourage discussion. The Census Bureau has reviewed this data product for unauthorized disclosure of confidential information and has approved the disclosure avoidance practices applied to this release, authorization number: CBDRB-FY2023-CES005-012. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Life and Death at the Margins of Society: The Mortality of the U.S. Homeless Population
Bruce D. Meyer, Angela Wyse, and Ilina Logani
NBER Working Paper No. 31843
November 2023
JEL No. I0,J0,R0

ABSTRACT

This paper examines the relationship between extreme socioeconomic disadvantage and poor health by providing the first detailed and accurate picture of mortality patterns among people experiencing homelessness in the U.S. Our analyses center on 140,000 people who were sheltered or unsheltered homeless during the 2010 Census, by far the largest sample ever used to study this population and the only sample designed to be nationally representative. These individuals, along with housed comparison groups, are linked to Social Security Administration data on all-cause mortality from 2010-2022 to estimate the magnitude of health disparities associated with homelessness. We find that non-elderly people experiencing homelessness have 3.5 times the mortality risk of those who are housed, accounting for differences in demographic characteristics and geography, and that a 40-year-old homeless person faces a similar mortality risk to a housed person nearly twenty years older. Our results reveal notable patterns in relative mortality risk by age, race, gender, and Hispanic ethnicity and suggest that within the homeless population, employment, higher incomes, and more extensive observed family connections are associated with lower mortality. The mortality hazard of homeless individuals rose by 33 percent during the COVID-19 pandemic, an increase that, while similar in proportional terms to the increase for the housed population, affected a much larger share of the homeless population due to their substantially elevated baseline mortality rate. These findings elucidate the persistent hardships associated with homelessness and show that the well-documented gradient between health and poverty persists into the extreme lower tail of socioeconomic disadvantage.

Bruce D. Meyer
Harris School of Public Policy
University of Chicago
1307 E 60th Street
Chicago, IL 60637
and NBER
bdmeyer@uchicago.edu

Ilina Logani
University of Oxford
ilina.logani@history.ox.ac.uk

Angela Wyse
University of Chicago
1307 E 60th Street
Chicago, IL 60637
awyse@uchicago.edu

A data appendix is available at http://www.nber.org/data-appendix/w31843

1. Introduction

Of the many hardships associated with poverty, heightened mortality risk is arguably the most alarming. More than a grim outcome, mortality is a fundamental indicator of quality of life, reflecting numerous dimensions of physical and mental health and one's sense of safety and well-being. Extensive research spanning academic disciplines, countries, and time periods has established a robust correlation between heightened mortality risk and socioeconomic disadvantage (Kitigawa and Hauser 1973, Deaton and Paxson 1999, Cutler et al. 2006). This correlation holds true whether privilege is defined by income and wealth (Chetty et al. 2016, Boen et al. 2010), education (Cutler and Lleras-Muney 2006, Cutler et al. 2011), social and occupational class (Cutler et al. 2012), or geography (Currie and Schwandt 2016). Yet despite this broad literature, little is known about the mortality risk faced by people in the extreme lower tail of socioeconomic disadvantage, due in part to the difficulty of accurately identifying the most deprived individuals in existing data sources like household surveys (Meyer et al. 2021).

This paper advances our understanding of the relationship between extreme poverty and health by examining the mortality of one of the most deprived segments of the U.S. population, people experiencing homelessness. Homelessness is both a stark indicator of material deprivation and an adverse life event, one that can have detrimental effects on health and personal safety. Recent developments like the rise in unsheltered homelessness, the COVID-19 pandemic, and surging deaths from opioids and other substances have drawn renewed attention to the humanitarian toll of homelessness, while also highlighting important gaps in our present understanding of the health vulnerabilities faced by this population. Although anecdotal evidence and numerous localized studies suggest that homelessness is associated with substantially elevated mortality risk, the extent of mortality disparities between homeless and housed individuals has not been examined nationally or with representative data, and little is known about heterogeneity in mortality risk within this population.

This paper provides the first detailed and accurate picture of mortality in the U.S. homeless population. Our approach centers on 140,000 people who were experiencing sheltered or unsheltered homelessness during the 2010 Census, by far the largest sample ever used to study this population's mortality and the only sample designed to be nationally representative. We follow these individuals using linked administrative data on all-cause mortality for twelve years,

including the first two years of the COVID-19 pandemic, and compare their mortality risk to representative samples of the overall housed and housed poor populations drawn from the Census and American Community Survey (ACS). Unlike prior work, we use the same, or closely comparable, datasets to calculate homeless and housed mortality risk and apply methods uniformly to both groups, an approach that facilitates direct and reliable comparisons and allows us to examine heterogeneity in mortality disparities with much greater detail than in prior work. This approach also allows us to examine the evolution of mortality disparities over time and to account for the full time-varying distribution of characteristics when comparing groups.

Our main finding is that non-elderly people who have experienced homelessness face 3.5 times the mortality risk of people who are housed, accounting for differences in demographic characteristics and geography. This disparity far exceeds the mortality gap between Black and white housed individuals, which we estimate to be 1.4, and is only slightly smaller than the mortality gap between disabled and non-disabled housed individuals, which we estimate to be 4.6. Comparing the mortality risk of people who are homeless and those who are poor but housed, we find that homelessness is associated with about a sixty percent greater mortality risk than poverty alone. Our estimates suggest that a 40-year-old homeless person has a mortality risk similar to a housed person who is nearly twenty years older and a poor housed person who is nearly ten years older.

Our analyses reveal notable patterns in mortality risk by age, race, income, family status, and type of homelessness. Homeless individuals' mortality risk relative to housed individuals differs over the life cycle and is greatest when they are in their 30s and 40s. Beginning in their 50s, however, homeless individuals' mortality hazard begins to converge with people who are housed, a pattern that may reflect both excess mortality of exceptionally vulnerable homeless individuals at younger ages and shared health vulnerabilities for elderly homeless and housed individuals. We also find that Black homeless individuals have lower mortality risk than those who are white, a pattern that may in part reflect a lower prevalence of substance abuse and behavioral health conditions among Black homeless individuals and may suggest important heterogeneity in the predominant pathways to homelessness by race. Within the homeless population, people who do not have a recent employment history, those with lower formal incomes, and those without observed family connections are especially vulnerable relative to their more advantaged and connected counterparts. Surprisingly, we find similar mortality risk

for people who were initially observed in shelters and those who were unsheltered once we control for gender, a finding that illustrates the substantial health risks faced by people experiencing homelessness even when they are not sleeping on the streets.

We estimate that average annual mortality risk rose by about half a percentage point for homeless individuals during the first two years of the COVID-19 pandemic, translating to about a 33 percent increase over their average during the two years preceding the pandemic after accounting for the expected increase in mortality due to aging. While the proportional rise in mortality risk was similar for people who were housed (30 percent) and poor and housed (34 percent), the pandemic affected a much larger share of the homeless population because of their substantially elevated baseline mortality risk. Homeless men experienced a larger rise in both absolute and proportional mortality risk during the pandemic (about 0.7 percentage points and 38 percent, respectively) than homeless women (about 0.3 percentage points and 24 percent).

Our findings illustrate, for the first time, the substantial health disparities associated with homelessness using data that are designed to be representative of the U.S. homeless population, while also calling attention to subsets of this population that are especially vulnerable and helping to establish the most broadly true patterns from among the many, often conflicting findings in previous work. In doing so, this paper adds to a growing body of research establishing fundamental facts about the size, characteristics, material circumstances, and housing transition dynamics of the U.S. homeless population. More broadly, this paper contributes to an expansive literature in economics on the association between socioeconomic disadvantage and poor health, suggesting that this gradient persists even into the extreme lower tail of socioeconomic disadvantage.

This paper proceeds as follows. Section 2 reviews available literature on homeless individuals' mortality and health and the broader literature on the relationship between socioeconomic status and health. Section 3 describes the decennial Census and American Community Survey (ACS) datasets from which we draw our homeless and comparison samples, as well as the administrative data on mortality, disability status, income, and family connections to which we link these samples. Section 4 describes our methods for linking datasets and estimating mortality hazards and relative risks. Section 5 presents our findings, including results from comparisons of homeless and housed individuals, comparisons of subsets of the homeless

population, and changes in mortality risk during the COVID-19 pandemic. Section 6 discusses and analyzes key findings and Section 7 concludes.

2. Background and related literature

2.1 Prior work on homelessness and mortality

Challenges in studying homeless individuals' mortality

Researchers often turn to mortality as an indicator of health and wellbeing for vulnerable populations because it is straightforward to measure and reflects unambiguous hardship. To this end, a small body of research examines the mortality patterns of people experiencing homelessness and estimates mortality disparities between housed and homeless individuals. These efforts are complicated, however, by the lack of representative data and the difficulty of obtaining longitudinal information. As a result, many prior studies are based on small, non-random samples of homeless individuals in major cities, primarily Boston, New York, or Philadelphia. Other studies focus on narrow subsets of the homeless population for whom data are more readily available, such as male veterans (Schinka et al. 2018), youth (Auerswald, Lin, and Parriott 2016), or people with post-traumatic stress disorder (Kasprow and Rosenheck 2000).

Table 1 summarizes the data, methods, and findings from prior studies with relatively large samples and sound methodologies. Even these analyses, however, face limitations that make it difficult to assess the generality of their findings. Barrow et al. (1999) and Metraux et al. (2011) draw large samples from New York City's administrative shelter databases, an approach that offers reliable mortality estimates for the city's sheltered homeless population but does not illuminate mortality patterns nationally or among people who are unsheltered. The remaining studies draw their samples from users of homeless health services, an approach that could bias findings to people who are either unwell or health-conscious enough to use these services (Baggett et al. 2013, Hibbs et al. 1994, Hwang 1997, Roncarati et al. 2018, Roncarati et al. 2020). Another limitation of these studies lies in their ability to obtain comparable mortality estimates for the housed population. For instance, Baggett et al. (2013) and Barrow et al. (1999) measure homeless individuals' mortality using linked microdata from the Massachusetts Department of Health and the National Death Index, but they obtain housed individuals' mortality rates using aggregated data from the Center for Disease Control. These and other data

and methodological discrepancies between housed and homeless mortality estimates complicate the interpretation of comparisons and limit the authors' ability to account for demographic differences between groups.

Key findings from prior work

Despite these data challenges, prior studies agree on several qualitative observations about homeless individuals' mortality risk relative to people who are housed. While point estimates of relative mortality risks differ widely across studies, most find that non-elderly people experiencing homelessness face a substantially elevated mortality risk relative to housed individuals of the same age group and gender. Relative mortality risks tends to be higher in early adulthood (Baggett et al. 2013, Barrow et al. 1999, Hwang et al. 1997, Hibbs et al. 1994), and compared to housed people of their gender, homeless women appear to have a greater risk than men (Baggett et al. 2013, Barrow et al. 1999, Hwang et al. 1997, Henwood et al. 2015, Hibbs et. Al. 1994).

Many of these studies also examine heterogeneity in mortality risks within the homeless population. Prior work suggests that white homeless adults face a heightened mortality risk relative to those who are Black or other races, a pattern that contrasts with mortality disparities by race in the housed population (Baggett et al. 2013, Hibbs et al. 1994, Roncarati 2018, Metraux et. al. 2011, Hibbs et. al. 1994, Roncarati et al. 2022). Prior work also broadly agrees that homeless men face higher mortality risk than homeless women, especially homeless women in families (Roncarati 2018, Hwang et al. 1997, Barrow et al. 1999, and Metraux et al. 2011). Studies that center on the unsheltered find that they face higher mortality risk relative to sheltered homeless populations (Roncarati 2018, Roncarati et al. 2020), and, unsurprisingly, that substance abusers face a particularly high mortality risk (Hibbs at al. 1994, Barrow et. al. 1999). These studies find no apparent seasonal pattern in mortality risks (Hibbs et al. 1994, Hwang et al. 1997), and there is little consensus on relationship between length of time spent homeless and Bmortality risks (Barrow et al. 1999, Metraux et al. 2011, Kasprow and Rosenheck 2000).

Causes of death and health risks

Prior work has suggested that the leading causes of death among people experiencing homelessness have changed over time and differed by age. Drug overdose appears to be the leading cause of death for homeless individuals under 45 in recent years, having superseded

HIV/AIDS in the mid-2000s. While substance abuse disorders have long been prevalent in this population, the most frequently abused substances have changed over time from alcohol, to cocaine, to methamphetamine, fentanyl, and other opioids in more recent years (Fischer and Breakey 1991, North et al. 2004, Cawley et al. 2022, Baggett et al. 2013, Roncarati et al. 2018, Roncarati et al. 2020). Traumatic injuries, including traffic accidents and homicides, appear to constitute the second leading cause of death for younger homeless individuals (Hwang et al. 1997, Roncarati et al. 2018, Roncarati et al. 2020, Schinka et al. 2018, Cawley et al. 2022, Hwang et al. 1997). For homeless individuals aged 45-64, heart disease and cancer appear to be the leading causes of death, followed by drug overdose and traumatic injury. Few studies have examined mortality among elderly homeless adults because this group is a small share of the homeless population.

Aside from cause of death, many studies also examine health conditions and health risks that disproportionately affect people experiencing homelessness. In the most recent survey designed to be nationally representative, Burt et al. (1999) found high rates of victimization and assault, difficulties in accessing medical attention, and alcohol, drug, and mental health conditions among people experiencing homelessness. More recent evidence suggests that homeless individuals experience accelerated aging, as evidenced by the early onset of chronic medical conditions and functional and cognitive impairments typically seen in housed adults aged 75 and older (Brown et al. 2022, Baggett et al. 2013, Hwang et al. 1997, Schinka et al. 2016, Garibaldi et al. 2005, Gelberg et al. 1990). For instance, older homeless adults are more likely than older housed individuals to have functional and mobility impairment, frailty, visual impairment, and urinary incontinence, and the prevalence of these and other "geriatric" conditions is equal to or higher than that seen in housed and housed poor adults twenty years older (Brown et al. 2012, Brown et al. 2017). Trick et al. (2021) also reported that the most frequently cited reasons for homeless individuals' emergency room (ER) visits are schizophrenia or auditory hallucinations, foot pain, and suicidal ideation. In a recent study designed to be representative of California's homeless population, two-thirds of respondents indicated symptoms of mental health conditions, but much smaller shares had received counseling or medication (Kushel and Moore 2023). These findings speak to a broad array of physical and behavioral health conditions and vulnerabilities that are likely linked to elevated mortality risk.

2.2 Relationship to literature on the health-socioeconomic status gradient

The study of homeless individuals' mortality relates to an expansive body of economic literature on the association between socioeconomic disadvantage and poor health more broadly (e.g. Kitigawa and Hauser 1973, Deaton and Paxson 1999). This association, often called the health-socioeconomic status gradient, arises whether disadvantage is defined or proxied by income and wealth (e.g. Chetty et al. 2016, Boen et al. 2010), education (e.g. Cutler and Lleras-Muney 2006, Cutler et al. 2011), social and occupational class (Cutler et al. 2012), or geography (Currie and Schwandt 2016). Unlike prior studies, the present work examines health disparities using one of the most extreme indicators of economic hardship available, homelessness.

A key question in this literature is whether a causal relationship exists between socioeconomic advantage and health, and if so in which direction and through which channels. On the one hand, human capital models suggest that poor health impedes the ability to work productively, limits the ability to invest in human capital, and reduces returns to such investments (Schultz 1962, Becker 1962, Grossman 1972, Becker 2007). Conversely, because health is a normal good, economic theory predicts that higher-income individuals should spend more money on health, which could produce a causal relationship in the other direction. At the same time, this theoretical literature emphasizes the dynamic nature of human capital processes and likely interactions between causal mechanisms over the life course, meaning that it may be difficult to identify a predominant causal direction or channel (e.g. Currie and Moretti 2003, Case et al. 2005, Almond and Currie 2011).

Empirical work offers support for numerous causal channels. For example, studies have found that poor health in childhood limits educational attainment and reduces earnings and labor force participation in adulthood (Brown et al. 2020, Case et al. 2005, Case and Paxson 2011) and that health shocks depress wages and reduce labor force participation (Smith 1999). Mental health conditions, which are prevalent among people experiencing homelessness, may be particularly important, with Currie and Madrian (1999) finding that this facet of health is one of the most important determinants of adult working days lost. Other studies have found that socioeconomic status, typically as proxied by income, wealth, and education, causally affects health by reducing health expenditures and investments or by affecting health behaviors (Boen and Yang 2016, De Walque 2007, Gramard and Parent 2007, Lleras-Muney 2005).

While little causal evidence exists on the relationship between poor health and homelessness specifically, this broader literature suggests channels through which the two may be related. Behavioral health conditions, substance use, and physical or mental health shocks could be important drivers of homelessness, while at the same time the experience of homelessness likely causes health to deteriorate through direct effects on physical and mental health and indirect effects on access to and continuity of medical care. Early life disadvantage in the form of parental resources and behaviors, health endowments (including behavioral health and vulnerability to addiction), and adverse childhood experiences may elevate the risk of both homelessness and mortality later in life. As with health and socioeconomic status more broadly, homelessness and poor health are likely causally related in both directions, with the importance of various channels differing substantially within this population.

3. Data

3.1 Census data on the U.S. homeless population

Our homeless sample is comprised of individuals counted during the 2010 Census's Service-Based Enumeration (SBE), an operation that took place March 29-31, 2010. The SBE included in the Census people sleeping in homeless shelters, people using soup kitchens or food vans who said they lacked a residence, and people sleeping outdoors at sites called Targeted Non-Sheltered Outdoor Locations (TNSOLs). We include all individuals with sufficient personal information to be linked to death records in our analysis. The linked subset consists of 140,000 individuals who are weighted to account for the probability of linkage.[1]

The SBE's enumeration frame was based on the address list of homeless service locations from the 2000 Census and augmented using internet research, queries to local officials and service providers, and a series of validation and advance visit operations. Prior work has shown that the coverage of the sheltered homeless population in the Census was surprisingly good, with about 90-95 percent of shelter users being included in its count, although it is worth noting that the Census's shelter definition excludes some facilities classified by HUD as homeless shelters

[1] While our sample is drawn from the tail-end of the Great Recession, prior work suggests that our findings likely generalize to people who experienced homelessness in the surrounding years. For example, Meyer et al. (2023) compare income and safety net participation for homeless individuals from the 2010 Census to people who were surveyed in homeless shelters by the American Community Survey (ACS) in surrounding years and find similar levels and longitudinal patterns of these outcomes.

(Meyer, Wyse, and Corinth 2023). The SBE also arrived at an unsheltered homeless population estimate similar in magnitude to the Department of Housing and Urban Development (HUD)'s point-in-time (PIT) estimate of the unsheltered homeless population.

Our homeless sample therefore consists of people who were literally homeless at a point in time in late March 2010. Because the study period continues through 2022, and because people frequently transition between homeless and housed statuses, it is likely that many or most of those in our sample were housed for some of the study period. HUD estimates that about one-quarter of people experiencing homelessness at a point in time are chronically homeless, i.e. experiencing frequent or extended homeless spells, while the rest are experiencing shorter or less frequent homeless spells (HUD 2022). Even when housed, however, prior work has shown that this population faces markedly worse material deprivation than the average housed poor individual, with extremely low incomes and high reliance on the safety net persisting for at least the decade surrounding the 2010 Census enumeration date (Meyer et al. 2023). Moreover, we find no evidence of heightened mortality risk for this population in 2010 and 2011, the years closest to when we observe them as homeless, relative to later years, a finding that suggests our results are applicable to people contemporaneously experiencing homelessness.

3.2 Administrative data on mortality, income, and family and disability status

We obtain death dates from the Census Bureau's Numerical Identification File (Numident), which is derived from Social Security Administration (SSA) records and frequently updated. The Numident has been shown to be a "high-quality and timely source of data to study all-cause mortality" (Finlay and Genadek 2021). A limitation of our study is that the Numident does not indicate cause of death.

We draw on several additional sources of administrative data to examine heterogeneity in mortality risk by income and employment, family status, and disability status. Specifically, we use Internal Revenue Service (IRS) 1040 extract files and W-2s (2005-2009) to determine income, employment status, and identify the presence of co-filing spouses and dependents prior to our study period. We also draw on administrative data from the Centers for Medicare and Medicaid Services (CMS) to identify 2009 recipients of Disability Insurance (DI) in Medicare records and Supplemental Security Income (SSI) in Medicaid records.

3.3 Census and ACS data on housed comparison groups

We compare homeless individuals' mortality risk to people who are housed and to a subset of the housed population that is also poor. The overall housed comparison group consists of a one percent random sample of housed adults from the 2010 Census. The housed poor comparison group is drawn from the 2009-2010 American Community Survey (ACS), which indicates household income relative to the poverty line. To obtain a large sample of poor individuals while keeping the selection date as close to the Census as possible, we keep individuals surveyed in the last three months of the 2009 ACS or the first three months of the 2010 ACS who were alive on April 1, 2010, the beginning of our study period.

4. Methods

4.1 Linking datasets

Our approach requires us to link birth and death dates from social security records to the homeless and comparison samples from the Census and ACS. We also link administrative data on transfer programs and tax records to determine disability status, connections to others, and income. We link these datasets using unique anonymized linkage keys known as Protected Identification Keys (PIKs), which are assigned by a Census Bureau program that searches for matches based on name, date of birth, gender, and address (or, in the case of homeless individuals, enumeration site address) in a reference file based on social security records.[2] PVS assigned a linkage key to 69 percent of those counted in homeless shelters, 42 percent of those counted at food vans and soup kitchens, and 17 percent of those counted at outdoor locations (TNSOLs) (Meyer et al. 2021).[3] Linkage rates are close to 90 percent for the housed comparison groups. Most homeless individuals who were not assigned a linkage key did not provide sufficient personal information to enumerators, in many cases because they were sleeping during

[2] The system, known as the Person Identification Verification System (PVS), uses addresses to narrow the number of potential matches for a Census record in the reference file, but if this approach does not yield a linkage key, PVS proceeds to search for matches using name, date of birth, and gender only (Layne and Wagner 2014). In this way, PVS can assign linkage keys to homeless individuals in the Census even if its reference file does not include the address where they were found during the SBE.

[3] Linkage rates for people experiencing homelessness in our sample meet or exceed linkage and attrition rates for marginalized populations in other studies. For example, the PIK rate Collinson et al. (2022)'s sample of court-involved individuals from Cook County is 52 percent. The attrition rate for Medicaid recipients surveyed in Finkelstein et al. (2012) was about 50 percent.

the count or were enumerated by sight at a bustling service location (Meyer et al. 2022). We adjust for non-linkage using inverse probability weights where the probability of linkage is estimated as a probit function of age, race, gender, Hispanic origin, state, and homeless location type.

4.2 Homeless individual and comparison sample restrictions

In our main results, we estimate mortality hazard rates and survival rates for three groups of non-elderly adults (individuals who are homeless, housed, housed and poor), defined as those ages 18-54 in 2010.[4] We focus on this age cohort in our main results because homelessness is rare among the elderly; in 2010, only 6.6 percent of the adult homeless population was 65 or older, compared to 17.3 percent of the overall housed adult population and 12.2 percent of the housed poor population. We do, however, produce results for some key outcomes with a sample that includes elderly people in all three groups to document differences in the mortality hazard by age.

Tables 2 and 3 display summary statistics for the non-elderly sample of homeless individuals and comparison groups. The non-elderly homeless sample consists of about 140,000 linked individuals and the housed comparison group includes about 1.3 million linked individuals. The housed poor sample consists of 110,000 linked individuals. Among the non-elderly, homeless individuals are older and are disproportionately likely to be between 45-49 and 50-54 years old. The homeless are also more likely to be male, especially compared to the housed poor, are more likely to be Black, and are more concentrated in the Northeast and West, reflecting the substantial homeless populations in New York and California.

4.3 Mortality hazard model

We specify the mortality hazard $\lambda_i(t)$ using a discrete time proportional hazard model with a non-parametric baseline hazard:

$$\lambda_i(t) = \lambda_0(t)\exp(z_i(t)'\beta)$$

In this equation, $\lambda_0(t)$ is the baseline hazard at time t (which is unknown, but estimated non-parametrically), where *t* indexes six-month periods between April 2010 and March 2022. $z_i(t)$ is a vector of time-dependent explanatory variables and covariates for individual *i*, and β is

[4] Surviving individuals' ages ranged from 30-66 at the end of the study period in 2022.

a vector of unknown parameters. The covariates included differ across specifications. We also include group interactions with the baseline hazard parameters in some specifications.

This model is a natural choice in our setting for several reasons. Proportional hazard survival models are frequently employed to study time-to-event data, particularly mortality. Moreover, while the underlying data-generating process is continuous, our data are discrete, with ties in the form of same-day deaths occurring not infrequently. The discrete model allows us to estimate the model without relying on approximations that would be required if using the Cox partial likelihood estimation method. Nevertheless, the estimates from our model are parameters of a continuous time hazard and thus retain an easy interpretation. We employ a non-parametric baseline because approaches that assume a parametric form for the baseline hazard provide inconsistent estimates when the assumed baseline hazard is incorrect, which likely leads to bias when events like the COVID-19 pandemic or other period effects give the hazard an unusual shape (Meyer 1990).

4.4 Mortality accounting for demographic differences between groups

We compare groups accounting for differences in their demographic characteristics in two ways. Our first approach is to estimate a hazard model with controls for covariates with key covariates being indicators for groups and a common baseline hazard. We estimate several different specifications with different sets of controls. We then interpret the coefficient on a group dummy as the proportional difference in the hazard between that group and a base group, accounting for covariate differences. The advantage of this approach is that it provides a simple summary measure of the relative hazard rate. The drawback is that it assumes a common baseline mortality hazard across all groups, which may not be correct.

Our second approach allows the baseline hazard to vary more flexibly between groups. Under this approach, we estimate a hazard model including the covariates but also interacting group indicators with the baseline hazard parameters. We then use these estimates and the distribution of covariates for the homeless population to simulate a hazard and survivor function for the homeless and our comparison groups. This approach provides us with predicted hazard rates and survivor functions for each group under the counterfactual scenario where they had the same covariates as the homeless group. In addition to allowing the baseline hazard function to vary between groups, this approach has the advantage of allowing us to estimate differences in

twelve-year survival accounting for demographic differences, not just semi-annual hazards, and to see the evolution of differences in mortality hazard rates over time.

Both approaches constrain the effect of a covariate to be the same for homeless and housed groups. This assumption may not be plausible in all cases. For example, comparisons of means for several outcomes suggest that among the homeless, groups that are more disadvantaged in the overall population fare better in certain respects than those who are more advantaged in the overall population. Meyer et al. (2023) find that homeless individuals who are Black have higher incomes and are more likely to be employed than those who are white. In this case, the assumption that race has the same effect on mortality for the homeless and the housed may be incorrect, suggesting that controlling for race may not make the homeless and our comparison groups more comparable. For this reason, we estimate specifications without controls and then with the progressive addition of age, gender, race and ethnicity, and geographic controls.

5. Results

5.1 Mortality disparities between the homeless and housed populations

In this section, we consider differences in mortality risk between people who are homeless and people who are housed and compare our findings to the previous literature. We estimate the magnitude of disparities between groups with and without accounting for demographic and geographic differences. We also estimate relative mortality risk for subsets of the population defined by gender, race, Hispanic ethnicity, disability status, and age.

Empirical mortality hazard and survivor functions

Figure 1 displays the empirical mortality hazard, defined as the probability of death in a six-month period conditional on being alive at the beginning of that period, for the non-elderly homeless population and for the housed and housed poor comparison groups. The shaded portion of the figure indicates periods during the COVID-19 pandemic. The mortality hazard increases over time as people in each cohort age, rising from 0.38 percent in the first period to about 1.19 percent in the final period for the homeless, 0.09 to 0.30 percent for the housed, and 0.18 to 0.47 percent for the housed poor. Homeless individuals' mortality hazard ranges from 3.9 to 4.9 times that of the housed population over the study's twelve-year period.

Unlike previous studies, we also compare the mortality of homeless individuals to people who are poor but housed. Our homeless sample's mortality hazard is 2.1 to 3.2 times that of the housed poor over the twelve years. We also find that housed poor individuals' mortality hazard is 1.4 to 2.1 times that of the housed population more broadly, but as we show in the next section, this disparity increases when we account for age and gender.

Figure 2 displays the empirical survivor function for the three groups, defined as the share of those alive at the beginning of our study who are alive at end of each six-month period. After 12 years, 96.1 percent of the housed population is still alive, compared to 93.8 percent of the housed poor and just 84.2 percent of the homeless population.

Mortality hazard accounting for differences between groups

Figure 3 summarizes the mortality hazard rate of the homeless and housed poor groups relative to the overall housed population and shows how the relative hazard changes after accounting for demographic and geographic differences between groups. Specifically, the figure displays the estimated coefficient on group indicators from the first estimation approach described in Section 3, where we regress mortality on group indicators for the homeless and housed poor samples and a common set of duration indicators for the three groups, as well as various sets of controls.

Without controls, the mortality hazard of homeless individuals is about 4.4 times that of the housed, but when we account for age and gender differences the relative hazard falls to 3.4. This estimate, which is higher than some estimates of relative mortality risk from previous literature and lower than others, is much more precise than prior work, with a 95 percent confidence interval with age and gender controls ranging from about 3.5 to 4.1 (Hwang et al. 1997, Hibbs et al. 1994, Baggett et al. 2013, Barrow et al. 2011).[5] Adding race, ethnicity, and geographic controls has little effect on the relative mortality rate, suggesting that age and gender are the key demographic differences between samples affecting relative mortality rates. Without controls, the housed poor are 1.6 times as likely to die as the broader housed population, but after accounting for age and gender their relative mortality risk rises to 2.1. Accounting for age and gender, we estimate that people who have experienced homelessness are about 60 percent more

[5] The mortality ratios calculated by Baggett et al. (2013) and Hwang et al. (1997) adjust for age, race, and gender, while Barrow et al. (2011) adjusts for both age and gender and Hibbs et al. (1994) only adjusts for age.

likely to die than those who are poor but housed, suggesting that homelessness is an important risk factor for mortality that is distinct from poverty alone.

Figure 3 also indicates the mortality risk of sheltered and unsheltered homeless individuals relative to the housed population. Without controls, the mortality hazard is slightly higher for the unsheltered than the sheltered, but after accounting for the larger male share among the unsheltered homeless, we find that these two subsets of the homeless population face similar mortality risk.

In Figure 4, we display the age- and gender-adjusted mortality hazard for the homeless and comparison groups. These results correspond to the second approach described in Section 4, where we estimate a model with group-specific baseline hazard parameters and simulate the mortality hazard for housed and housed poor groups using the distribution of characteristics of the homeless sample.[6] The main difference between the empirical and covariate-adjusted hazards is that the housed poor have a higher mortality hazard when we account for age and gender, as we saw in Figure 3, reflecting the fact that when we align their characteristics with the older, more male homeless population, their mortality hazard increases. Table 4 reports cumulative mortality over the twelve-year study period using the empirical and covariate-adjusted hazards. When considering cumulative rather than period-specific mortality, we find people in our homeless sample were 3.2 times as likely to die during the study period as the housed and about 1.6 times as likely to die as those who are housed but poor, accounting for age and gender.

Figure 4 illustrates in stark terms the considerable health disparities associated with poverty and homelessness. People who are poor but housed are about twice as likely to die as the average housed person, and people who have experienced homelessness face a mortality risk that is about 60 percent higher than those who are poor but housed.

Gender and mortality

Figure 5 displays the mortality risk of homeless and poor housed individuals relative to the housed by gender, controlling for age. The first set of points in this figure indicates the hazard by gender and housing status relative to housed men. This set allows us to see how the

[6] Figure 4 includes confidence intervals for the covariate-adjusted mortality hazard since these are predicted according to the methodology described in the text rather than observed in our data, as was the case for the empirical mortality hazard.

mortality hazard differs by gender for a given housing status and across housing statuses for a given gender. To make it easier to see how the mortality hazard differs by housing status among women, the second set of points indicates the hazard of homeless and housed poor women relative to housed women.

We find that men have higher mortality risk than women with a given housing status. For example, housed men have mortality hazard that is 35 percent higher than housed women, and homeless men have mortality hazard that is 29 percent higher than homeless women. At the same time, homeless women face mortality risk that is four times that of their housed counterparts, whereas homeless men are only about 3.3 times as likely to die as housed males, estimates that reflect the higher mortality risk of housed men compared to housed women. These findings are consistent with prior literature suggesting that homeless women face especially elevated mortality risk relative to their housed counterparts, although the magnitude of homeless-to-housed mortality disparities for women in our study is smaller than in previous work (Baggett et al. 2013, Barrow et al. 1999, Hwang et al. 1997, Henwood et al. 2015, Hibbs et al. 1994). These findings suggest that gender differences in mortality risk found in some past studies may not generalize to the U.S. homeless population more broadly.

Race and mortality

Figure 6 displays the mortality risk of homeless and poor housed individuals relative to the housed by race, controlling for age. The first set of points in this figure indicates the hazard by race and housing status relative to housed people who are white, while the second and third sets of points indicate the relative hazard of homeless and housed poor people who are Black and of other races, respectively, relative to housed people of the same race.

For housed and poor housed individuals, mortality risk is highest for people who are Black, followed by those who are white, and then then those of other races. Among the homeless population, however, white individuals have the highest mortality risk, followed by people who are neither white nor Black. In a reversal of the pattern observed in the housed population, Black individuals have the lowest relative mortality risk among the homeless population. This pattern in mortality risk by race is consistent with previous work, which also found lower mortality risk for Black individuals experiencing homelessness (Hibbs et al. 1994, Metraux et al. 2011, Baggett et al. 2013).

Comparing homeless individuals' mortality risk to housed individuals of the same race, we find that white homeless individuals and those of other races have the most elevated mortality risk relative to their housed counterparts, at 4.7 and 4.6 times, respectively. Black homeless individuals are 2.3 times as likely to die as their housed counterparts, a fact that reflects both the relatively low mortality hazard of Black individuals within the homeless population and the elevated mortality risk of Black housed individuals compared to those who are white and of other races. This finding, too, is consistent with previous work (Hibbs et al. 1994, Metraux et al. 2011, Baggett et al. 2013).

Hispanic ethnicity and mortality

Figure 7 displays relative mortality risk by housing status and Hispanic ethnicity. Hispanic individuals have lower mortality risk than non-Hispanics in each of the three housing statuses.[7] For example, a homeless Hispanic person has, on average, 23 percent lower mortality risk than a non-Hispanic person, controlling for age. Non-Hispanics who are homeless have slightly higher mortality risk relative to their housed counterparts (3.8 times) than do homeless Hispanics (3.5 times). No previous study, to our knowledge, has looked at differences in mortality risk by Hispanic ethnicity in the homeless population.

Disability status

Figure 8 displays relative mortality risk by housing status and disability. We define a person to be disabled if Medicare records indicate that they received Disability Insurance (DI) in 2009 or if Medicaid records indicate that they received Supplemental Security Income (SSI) in 2009. As Table 2 indicates, a much larger share of the homeless population was disabled before the beginning of our study (20.6 percent) than of the housed poor (10.7 percent) or of the broader housed population (3.9 percent). We note, however, that Meyer et al. (2023) find that DI and SSI receipt increase at a faster rate for the Census homeless population after 2010 than for the housed and housed poor populations, meaning that a larger share of homeless individuals indicated as non-disabled in our study became disabled or enrolled in DI or SSI during the study period.

[7] Hispanics' lower mortality risk is not a novel finding. Hispanic individuals are frequently found to experience similar or better health outcomes than non-Hispanic individuals in the United States despite socioeconomic disadvantage, a pattern that is sometimes called the "Hispanic mortality paradox" (Ruiz, Steffen, and Smith 2013).

People who are disabled face substantially higher mortality risk than non-disabled individuals with the same housing status, controlling for age. A housed disabled person is 4.5 times as likely to die in a six-month period as a non-disabled housed person, while a housed poor disabled person is 2.8 times as likely to die as a non-disabled housed poor person, accounting for age. A homeless disabled person is 1.6 times as likely to die as a non-disabled homeless person. Notably, disabled housed and housed poor individuals have even a higher mortality risk than non-disabled homeless individuals.

The mortality risk is very similar for all three groups of disabled individuals. In contrast, a non-disabled homeless person is about 4 times as likely to die as a non-disabled housed person. In other words, mortality disparities between housed and homeless individuals are much smaller among people with disabilities than people without disabilities. This latter fact may reflect in part the prevalence of disabilities in the homeless population not captured by our measure, as suggested by the steep increase in disability program receipt after 2010 found in Meyer et al. (2023). Nevertheless, it appears that mortality disparities by housing status are concentrated almost exclusively among people who were not enrolled in disability assistance programs at the beginning of our study period. Our study is the first, to our knowledge, to look at mortality hazard by disability status in the U.S. homeless population.

Age and mortality

Figure 9 displays relative mortality risk by housing status and age, where we have selected age bins to facilitate comparisons to prior literature. Homeless individuals in the youngest age category, 18-24, have the lowest mortality risk relative to their housed counterparts; they are slightly less than twice as likely to die in a six-month period. Relative mortality risk is highest for homeless individuals ages 45-54, who are about 4.2 times as likely to die as their housed counterparts.

These findings are largely consistent with prior work, which finds a peak in the homeless-to-housed mortality ratio between the ages of 25 and 49 (Hibbs et al. 1994, Barrow et al. 1999, Baggett et al. 2013). Our estimate of homeless-to-housed mortality risk for those ages 18 to 24, however, is smaller than prior studies' estimates, which range from 2.7 to 11.8 times that of the housed young population (Hibbs et al. 1994, Hwang et al. 1997, Barrow et al. 1999,

Baggett et al. 2013).[8] Hwang et al. (1997), in particular, find that homeless adults aged 18 to 24 had the highest mortality risk relative to the housed in their sample.

Figure 10 presents estimates of mortality hazard by housing status in two-year age bins relative to housed 30- to 31-year-olds. While all three groups' mortality risk increases as they age, the three groups' relative mortality risk begins to converge after the age of 50, a fact that is more readily apparent in Figure 11, which displays the ratio of estimated homeless and housed poor coefficients relative to the housed coefficient in the mortality hazard model. In their late 70s, homeless individuals face the same mortality risk as housed poor individuals and are only about 1.4 times as likely to die as their housed counterparts.

The convergence in relative mortality risk between groups may reflect the declining relative frailty of the surviving homeless population as the highest-risk individuals die. This pattern could also reflect the fact that risks such as cardiovascular disease that rise with age that affect housed and unhoused individuals similarly. We also note that safety net eligibility is changing over time and as people age, and that such shifts in eligibility may affect relative mortality risk between groups. We do not account for the safety net here, but Wyse and Meyer (2023) find that the effect on mortality of safety net programs like Medicaid and social security does not appear to be large, at least as indicated by changes around eligibility ages and policy implementation dates.

Figure 10 also illustrates the age at which each group will face a given level of mortality risk relative to the baseline group (30- to 31-year-old housed individuals). The dashed line on this figure indicates the mortality risk of a 40-year-old homeless person. Its intersection with the comparison groups' lines indicates the age at which people in those groups will face the same mortality risk as a 40-year-old homeless person. We see that a 40-year-old homeless individual faces a mortality risk that is similar to that of a 58-year-old housed person and a 48-year-old housed poor person. In other words, homelessness is associated with a health detriment equivalent to nearly twenty years of aging relative to the typical housed person.

5.2 Identifying the most vulnerable subsets of the homeless population

In this section, we consider differences in the mortality risk among subsets of the homeless population with the goal of identifying the most vulnerable groups and factors which

[8] This range includes relative mortality risks estimated separately for men and women of this age range.

may be protective against mortality risk. We also consider differences by type of homelessness during the 2010 Census (i.e. sheltered or unsheltered), state of residence, income and employment status, and the extent of observed family connections.

Mortality risk by type of homelessness

Figure 12 shows the mortality hazard in each period for sheltered and unsheltered men and sheltered and unsheltered women, controlling for age.[9] Both sheltered and unsheltered men are about half a percentage point more likely to die in a six-month period than sheltered and unsheltered women. As in Figure 3, conditional on gender and age, sheltered and unsheltered people have very similar mortality hazard rates, conflicting with prior work finding that unsheltered individuals have higher mortality risk than sheltered individuals (Roncarati et al. 2018, Roncarati et al. 2020). This result may reflect the fact that our sample is designed to be representative of the overall homeless population, not just of health services users as in Roncarati et. al (2018, 2020). We caution, however, that our study indicates sheltered or unsheltered status in the year 2010, and we are unable to ascertain people's living situations at other points in time. It is likely that many people in our sample transitioned between sheltered and unsheltered homelessness and other housing statuses throughout our study period.

This finding is surprising, too, because we know based on prior work that people who were sheltered homeless during the 2010 Census had greater incomes, employment, and connections to the safety net than unsheltered people of the same gender. For example, Meyer et al. (2023) found that about 55 percent of sheltered women had formal employment in 2010, compared to 42 percent of unsheltered women. About 50 percent of sheltered males and 40 percent of unsheltered males were formally employed that year. Yet despite important apparent differences in these populations' material well-being, mortality appears to be similar between sheltered and unsheltered homeless individuals.

Figure 13 indicates mortality risk relative to the sheltered white group by sheltered status, race, and gender. Black homeless individuals have lower mortality risk than those who are white even conditional on gender and type of homelessness. White women who are unsheltered have slightly lower mortality risk than sheltered homeless white women, while sheltered and unsheltered white men have nearly the same mortality hazard.

[9] We use the age distribution of sheltered males to simulate the covariate-adjusted hazard for the other three groups.

Figure 14 displays mortality risk by age relative to the youngest sheltered homeless cohort (ages 18-24), controlling for gender. Once again, sheltered and unsheltered individuals have similar mortality risk by age group.

Mortality risk in New York, California, and other states

Figure 15 displays the relative mortality hazard rate by state of residence (New York, California, or other states), controlling for demographic characteristics, type of homelessness, and income according to 2005-2009 tax records. People who are homeless in New York have a mortality risk that is about 13.4 percent lower than those in other states, while the mortality risk for California's homeless population is not statistically significantly different from the risk for homeless individuals in states besides New York.

New York residents' lower mortality does not reflect differences in demographic characteristics, income, or type of homelessness, because we have controlled for these variables in our estimation. Their lower mortality risk also does not appear to reflect differences in disability status or safety net engagement. Meyer et al. (2023) find that homeless individuals in New York and in the rest of the country have similar rates of disability program receipt (23 percent and 19 percent, respectively) and similar rates of receipt of other major safety net programs (93 and 89 percent).[10] One possible explanation lies in the generosity of homelessness services in New York, where a court-mandated "right to shelter" policy has increased the availability and quality of shelter beds, especially for families (O'Flaherty 2019). Better services could improve the health of people experiencing homelessness. Higher shelter quality could also affect the relative affluence of the average shelter resident by making shelters preferable to some extremely undesirable housed situations, resulting in a sheltered homeless population that is drawn from a slightly less disadvantaged population.[11]

[10] These shares reflect receipt of benefits through the Supplemental Nutrition Assistance Program (SNAP), housing benefits through the Department of Housing and Urban Development (HUD), Medicare or Medicaid, or service-connected disability through the Veterans Benefit Administration. A key exception in New York's safety net generosity concerns Medicaid, which was available to all poor adults at the beginning of our study period in the state but only became available to poor adults in most other states after 2014, under provisions of the Affordable Care Act (ACA). However, Wyse and Meyer (2023) do not find evidence that Medicaid enrollment reduces homeless individuals' mortality risk, suggesting that a causal relationship between Medicaid availability and lower mortality risk in New York is weak, if present.

[11] Families in New York must be vetted before being admitted to the shelter system and O'Flaherty (2019) notes that most families who apply are rejected, meaning they were determined to have access to other housing options and may hence be less disadvantaged than the typical homeless person in other states.

Mortality risk by employment and income status

Figure 15 also shows how mortality differs by employment status and income among people experiencing homelessness. We define someone as employed if they had formal earnings in 2009 according to IRS 1040 and W2 datasets.[12] We define someone as being in the top half of the income distribution by taking the average of their inflation-adjusted pre-tax cash income according to tax records over 2005-2009 and comparing this to the median for people with the same sheltered status.[13]

We find that people with a recent history of employment and people in the top half of the income distribution are about 34.6 and 32.7 percent less likely to die in a six-month period than those who were not employed and those in the bottom half of the income distribution. These findings show that even among people who have experienced homelessness, those who are more economically disadvantaged and more disconnected from the formal labor market have worse health outcomes.

Mortality risk by extent of observed family connections

The third set of results on Figure 15 show how mortality differs by the extent of observed family connections. We classify individuals as having or having once had a spouse if they ever had a co-filer on a 1040 in 2005-2009, and we classify them as having a child if they ever included a dependent on a 1040 in those years. We also attribute family connections to individuals who were recorded in the Census as being housed in addition to homeless. Meyer et al. (2022) document widespread double-counting of people experiencing homelessness in the Census and find that duplicate records often reflect those individuals' inclusion on the Census form of a housed family member, oftentimes their parent. In addition to demographic characteristics, we control for income to ensure that our estimates are not confounded by the fact of tax filing, which is in turn associated with higher income.

[12] For people who link to a 1040, we define earnings as the sum of 1040 wage and salary income, estimated non-negative 1040 self-employment income (when a self-employment schedule was filed), and W2 deferred compensation, minus any W2 wages and tips associated with a co-filer. For people who do not link to a 1040 but do link to a W2, earnings are equal to wages and tips across W2s. For people who link to neither, earnings are zero.
[13] For people who link to a 1040, pre-tax income is equal to the sum of total money income and VA service-connected disability compensation. For people who do not link to a 1040, pre-tax income is equal to the sum of wages and tips and deferred compensation in W2s, VA service-connected disability compensation, and IRA and employer-sponsored retirement distributions across 1099-Rs.

Having at least one observed family connection is associated with 17.3 percent lower mortality risk for people experiencing homelessness. Homeless individuals who have a current or former spouse face a mortality risk that is 21.3 percent lower than those who do not, and people who have children are 21.6 percent less lower mortality risk than those who do not. Homeless individuals who were recorded on a housed family member's Census form are 13.6 percent less likely to die than those who were not. Family connections appear to be an important protective factor against mortality for people who have experienced homelessness, albeit one that is not as potent as our measures of income and employment.

5.3 Mortality during the COVID-19 pandemic

The empirical hazard in Figure 1 suggests a steep rise in mortality hazard during the COVID-19 pandemic. In this section, we examine the pandemic-era rise in mortality and compare its magnitude across groups. Specifically, we consider the absolute and proportional change in average annual mortality risk in the two years prior to the pandemic (April 2018-March 2020) and the first two years of the pandemic (April 2020-March 2022). We calculate these changes using both the empirical mortality hazard and the covariate-adjusted mortality hazard, which uses the distribution of age and gender among the homeless to provide a comparable hazard for the three groups.

In describing the COVID-era rise in mortality hazard, we wish to account for the fact that aging would have caused our cohorts' mortality hazards to rise over these four years regardless of the COVID-19 pandemic. To do so, we regress the mortality hazard in six-month periods indexed by $j = 1, \ldots, 20$ on a constant and a time trend. We then take the estimated coefficient on this time trend and multiply it by eight to obtain the estimated effect of aging on the average annual mortality hazard, which we subtract from the observed rise in the mortality hazard to obtain an aging-adjusted estimate of the hazard increase.[14]

Table 5 displays the results from applying this methodology. We find that all three groups experienced an approximately 30-35 percent increase in their average annual mortality hazard in the two years of the pandemic relative to the two preceding years exceeding the change we would have expected due to aging. At the same time, the absolute increase in mortality was

[14] Multiplying this estimate by four gives us the effect of aging on average biannual hazard between the midpoint of our pre-pandemic period and the midpoint of our post-pandemic period. We then multiply the estimate by two to convert the estimate's effect on the biannual hazard to its effect on an annual one.

much larger for the homeless population than other housing status groups given their substantially elevated baseline mortality risk. Figure 1 displays the observed mortality hazard alongside the predicted hazard accounting for aging, which is indicated by a dashed line. The gap between predicted and observed hazard illustrates the magnitude of excess mortality.

We also calculate differences in the pandemic-era mortality rise by gender and type of homelessness and display these findings in Table 6. Sheltered and unsheltered males saw a much larger absolute and proportional rise in their mortality risk (35 and 28 percent, respectively) during the pandemic beyond what we would have expected due to aging. Sheltered and unsheltered females saw a 24 and 21 percent increase in their mortality risk. Again, the distinction between genders is much more pronounced than the distinction between sheltered and unsheltered homelessness. An important caveat is that type of homelessness reflects status in 2010 and may not reflect people's living situations during the pandemic.

Because we lack information on cause of death, however, we caution against attributing excess mortality to COVID-19 directly. Previous research has indicated that excess mortality during the pandemic could be attributed to rising fentanyl, other opioid, or methamphetamine use over the last two decades, or, possibly relatedly, to difficulties in obtaining medical services for life-threatening situations like overdoses and traumatic injuries due to hospital overcrowding (Cawley et al. 2022, Baggett et al. 2011). Cawley et al. (2022) found that the substantial rise in homeless individuals' mortality in San Francisco during the pandemic was driven by difficulties in obtaining care for emergencies from an overburdened medical system. COVID-19 itself was not a leading cause of death for homeless individuals in their sample. The pandemic-era rise in mortality should therefore be interpreted as the combined effect of the pandemic and any associated changes in all-cause mortality risk.

6. Discussion

The findings in this paper, which are based on the largest and most representative sample used to study homeless mortality to date, establish the most broadly true patterns among the mixed findings in prior work. For example, we find, as do many prior studies, that homeless individuals' mortality risk relative to the housed population is greatest when they are in their 30s and 40s and that homelessness is associated with more elevated mortality risk for women than for men when compared to housed people of the same gender. We also find that Black homeless

individuals have lower mortality than white homeless individuals, a pattern that is consistent not only with others' results on mortality but also with work on race group differences for other indicators of wellbeing for this population, such as income and connections to the formal labor market (Hibbs et al. 1994, Baggett et al. 2013, Metraux et al. 2011, Meyer et al. 2023).

Mortality differences by race in the homeless population are especially noteworthy because they diverge from those observed in the housed population and may provide insight into the relative importance of different pathways to homelessness between racial groups. For example, a key question in homelessness research is the extent to which homelessness is driven by personal vulnerabilities like addiction and poor mental health versus poverty and economic shocks (Lee, Tyler, and Wright 2010). Prior work suggests that a proximate explanation for white homeless individuals' elevated mortality risk lies in the higher prevalence of substance abuse and behavioral health conditions in this group, which could in turn suggest that personal vulnerabilities are more important drivers of homelessness among white individuals while economic circumstances are relatively more important for Black individuals (Hibbs et al. 1994, Baggett et al. 2013). Such a pattern could arise because white individuals, on average, have access to better-resourced social and family networks to protect against homelessness, meaning that only those with especially difficult personal circumstances become homeless. This pattern could also arise because Black individuals face greater housing discrimination, meaning that even relatively better-off individuals find themselves homeless because no affordable housing options are available to them. These hypotheses merit further study in future work.

Another advantage of our data lies in the richness of demographic, economic, and social information available, which in turn allows us to examine in great detail the association between these factors and mortality and to highlight especially vulnerable segments of an already extremely deprived population. For example, we find that people with lower incomes and those who are less connected to family and to the labor market have especially poor health outcomes. Moreover, while disabled homeless individuals have a higher mortality rate than those who are not disabled, even non-disabled homeless individuals have a substantially elevated mortality rate relative to their housed counterparts. Somewhat surprisingly, we find no major difference in mortality rates between unsheltered and sheltered homeless individuals after controlling for gender, and we find that both of these groups experienced substantial increases in mortality during the COVID-19 pandemic. These findings may prove useful to policymakers and service

providers looking to target resources to the neediest individuals. They also highlight the significant health risks associated with homelessness even among people sleeping in shelters, a group that figures less prominently in policy debates than their more visible unsheltered counterparts.

While our analyses indicate strong associations between homelessness, individual characteristics, and mortality, we caution that our study does not allow us to identify a causal or directional relationship between homelessness and elevated mortality risk. As with socioeconomic status and health in the housed population, homelessness and mortality are likely causally related in both directions, with the importance of various channels differing across sub-groups. Behavioral health disorders, addiction, and substance abuse, for instance, could drive both homelessness and heightened mortality risk for some. For others, particularly those who are chronically homeless, heightened mortality risk could be a consequence of the exceedingly severe long-term material deprivation the homeless population experiences. We also caution that, in relating our findings to the literature on the health-socioeconomic status gradient, it is probably incorrect to view homelessness purely as a proxy for extreme poverty. Mortality patterns within the homeless population point to a complicated selection procedure into homelessness that differs across racial groups and genders appears to be determined only in part by economic circumstances.

7. Conclusions

This study examines health disparities in the United States using one of most fundamental indicators of well-being, mortality, for one of the most disadvantaged segments of the population, people experiencing homelessness. We base our analyses on by far the largest and most representative data ever used to compare the mortality risk of homeless and housed populations, data that include a rich set of demographic characteristics that facilitate detailed comparisons by age, gender, race, Hispanic ethnicity, and disability status. Within the homeless population, we examine mortality differences by sheltered status, income, employment, and the extent of observed family connections to identify factors associated with heightened mortality risk, analyses which may serve to help researchers, service providers, and policymakers identify the most vulnerable subsets of this already exceptionally vulnerable population.

In addition to the size, national scope, and representativeness of these data, this approach benefits from important advantages. Unlike prior studies, we compare housed and homeless mortality using data from the same sources and applying a common methodology to both groups, allowing for more nuanced and reliable comparisons. We also link the individuals in our study to administrative tax and program data to access rich longitudinal information on income, employment, disability status, and safety net participation, which in turn allows us to compare mortality risk between subsets of the homeless population and to characterize the patterns of long-term material deprivation that accompany elevated mortality risk. We supplement our analyses using a nationally representative sample of poor housed individuals from a closely comparable data source to learn about homelessness as a risk factor for mortality that is distinct from poverty in general. We view this work as complementary to an extensive and growing body of clinical and public health research into the relationship between homelessness and health, adding a nationally representative perspective to prior findings and helping to establish the most broadly true patterns from the wide array of results in prior work.

Our findings reveal severe disparities in health and wellbeing between people who are homeless and those who are housed. People who have experienced homelessness face 3.5 times the mortality risk of people who are housed, accounting for differences in demographic characteristics and geography. Put differently, a 40-year-old homeless person faces a mortality risk similar to a housed person nearly twenty years older. These disparities reflect more than economic disadvantage: homeless individuals' mortality risk is about 60 percent greater than poor housed individuals of the same age and gender. Mortality disparities change over the course of the life cycle, however, with homeless individuals' relative mortality risk peaking between the ages of 30 and 50 before falling to converge with the poor housed population's mortality risk by the age of 70.

Many of the groups that face higher mortality risk in the housed population – men, non-Hispanics, people with lower incomes, people who are disabled – also face higher risk in the homeless population, but race is a notable exception. Among people who are housed, someone who is Black has 40 percent greater mortality risk than someone who is white, but among people who have experienced homelessness, someone who is Black has 27 percent lower mortality risk than someone who is white. This finding mirrors prior work showing that Black homeless individuals have higher incomes and are more connected to the formal labor market and safety

net than those who are white (Meyer et al. 2023). These patterns merit further examination, as they may suggest that the predominant pathways to homelessness differ by race, with individual conditions like addiction and behavioral health issues perhaps playing a greater role for white individuals, while structural issues like discrimination and poverty being more important drivers for Black individuals.

Our findings also speak to the exceptionally severe toll of the COVID-19 pandemic on people experiencing homelessness. The pandemic coincided with a 33 percent increase in mortality for people experiencing homelessness beyond what we would have expected due to the aging of our cohort. While housed and poor housed people saw similar proportional rises in their mortality risk, the pandemic affected a much larger share of the homeless population because of their already elevated baseline mortality rate. Homeless men seem to have been hit especially hard, with their mortality risk rising by 35-38 percent during the pandemic, compared to 22-24 percent for homeless women. The causes of excess mortality during the pandemic are not apparent from our data, however, and could include both the direct effects of virus itself, as well as indirect effects from strain on healthcare systems and reduced access to emergency services, with these latter issues potentially interacting with surging fentanyl use to exacerbate the pandemic's harm on people experiencing homelessness.

Within the homeless population, connection to the formal labor force and to family are associated with lower mortality risk; having been observed in a shelter in the 2010 Census, as opposed to an unsheltered location, is not. People who have or had spouses, have children, or who were included on a family member's housed Census form all have mortality risk that is about 20 percent lower than their counterparts. People who are homeless in New York have a lower mortality risk, but those who are homeless in California face a similar mortality risk as people in other states. Perhaps surprisingly, we find very little difference in mortality between people we initially observe in homeless shelters and those who are unsheltered, conditional on gender. This last finding highlights the substantial health risks faced even by people experiencing homelessness in shelters, a group that is less visible and receives less attention than those who sleep on the streets, but who nevertheless experience substantial health disparities. Our findings on mortality differences between sub-groups of the homeless population point to dimensions linking health and socioeconomic status that may be especially important among extremely

disadvantaged individuals more broadly, such as housing quality and stability, social connections and family resources, and disability status.

This paper joins a growing body of work through the Comprehensive Income Dataset (CID) project that aims to establish fundamental facts about homelessness in the United States by linking Census and administrative data to unlock new insights. Recent work has improved our understanding of the size of the U.S. homeless population, established the surprisingly good coverage of people experiencing homelessness in the Census, and revealed persistent extremely low incomes and high reliance on the safety net. Ongoing and planned work aims to understand the effects of safety net programs on homeless individuals' mortality and material well-being and to learn about the dynamics of transitions between housing, institutional settings, and homelessness. In providing the first national estimates of homeless mortality in the U.S., this paper not only adds to the emerging picture of the persistent hardships and stark health disparities associated with homelessness, but also sheds light on some of the most vulnerable subsets of an already exceptionally vulnerable population and contributes to efforts to more effectively mitigate the mortality risks faced by people experiencing homelessness.

8. Citations

Auerswald, Colette L., Jessica S. Lin, and Andrea Parriott. 2016. "Six-Year Mortality in a Street-Recruited Cohort of Homeless Youth in San Francisco, California." *PeerJ* 4 (April): e1909. https://doi.org/10.7717/peerj.1909.

Baggett, Travis P., Stephen W. Hwang, James J. O'Connell, Bianca C. Porneala, Erin J. Stringfellow, E. John Orav, Daniel E. Singer, and Nancy A. Rigotti. 2013. "Mortality Among Homeless Adults in Boston: Shifts in Causes of Death Over a 15-Year Period." *JAMA Internal Medicine* 173 (3): 189. https://doi.org/10.1001/jamainternmed.2013.1604.

Barrow, S M, D B Herman, P Córdova, and E L Struening. 1999. "Mortality among Homeless Shelter Residents in New York City." *American Journal of Public Health* 89 (4): 529–34. https://doi.org/10.2105/AJPH.89.4.529.

Becker, G.S., 1962. Investment in human capital: A theoretical analysis. Journal of political economy, 70(5, Part 2), pp.9-49.

Becker, G.S., 2007. Health as human capital: synthesis and extensions. Oxford economic papers, 59(3), pp.379-410.

Boen, C. and Yang, Y.C., 2016. The physiological impacts of wealth shocks in late life: Evidence from the Great Recession. Social Science & Medicine, 150, pp.221-230.

Boen, C., Keister, L. and Aronson, B., 2020. Beyond net worth: Racial differences in wealth portfolios and black–white health inequality across the life course. Journal of health and social behavior, 61(2), pp.153-169.

Brown, D.W., Kowalski, A.E. and Lurie, I.Z., 2020. Long-term impacts of childhood Medicaid expansions on outcomes in adulthood. The Review of economic studies, 87(2), pp.792-821.

Brown, Rebecca T., Jennifer L. Evans, Karen Valle, David Guzman, Yea-Hung Chen, and Margot B. Kushel. 2022. "Factors Associated With Mortality Among Homeless Older Adults in California: The HOPE HOME Study." *JAMA Internal Medicine* 182 (10): 1052. https://doi.org/10.1001/jamainternmed.2022.3697.

Brown, Rebecca T., Kaveh Hemati, Elise D. Riley, Christopher T. Lee, Claudia Ponath, Lina Tieu, David Guzman, and Margot Kushel. 2017. "Geriatric Conditions in a Population-Based Sample of Older Homeless Adults." *The Gerontologist* 57 (4).

Brown, Rebecca T., Dan K. Kiely, Monica Bharel, and Susan L. Mitchell. 2012. "Geriatric Syndromes in Older Homeless Adults." *Journal of General Internal Medicine* 27 (1).

Burt, Martha R, United States, Interagency Council on the Homeless, and Urban Institute. 1999. *Homelessness: Programs and the People They Serve.* Collingdale, Penn.: DIANE Pub. Co.

Case, A., Fertig, A. and Paxson, C., 2005. The lasting impact of childhood health and circumstance. Journal of health economics, 24(2), pp.365-389.

Case, A. and Paxson, C., 2011. The long reach of childhood health and circumstance: Evidence from the Whitehall II study. The Economic Journal, 121(554), pp.F183-F204.

Cawley, Caroline, Hemal K. Kanzaria, Barry Zevin, Kelly M. Doran, Margot Kushel, and Maria C. Raven. 2022. "Mortality Among People Experiencing Homelessness in San Francisco During the COVID-19 Pandemic." *JAMA Network Open* 5 (3): e221870. https://doi.org/10.1001/jamanetworkopen.2022.1870.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A. and Cutler, D., 2016. The association between income and life expectancy in the United States, 2001-2014. Jama, 315(16), pp.1750-1766.

Collinson, R., Humphries, J.E., Mader, N.S., Reed, D.K., Tannenbaum, D.I. and Van Dijk, W., 2022. Eviction and poverty in American cities (No. w30382). National Bureau of Economic Research.

Currie, J. and Madrian, B.C., 1999. Health, health insurance and the labor market. Handbook of labor economics, 3, pp.3309-3416.

Currie, J. and Moretti, E., 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. The Quarterly journal of economics, 118(4), pp.1495-1532.

Currie, J. and Almond, D., 2011. Human capital development before age five. In Handbook of labor economics (Vol. 4, pp. 1315-1486). Elsevier.

Currie, J. and Schwandt, H., 2016. Mortality inequality: The good news from a county-level approach. Journal of Economic Perspectives, 30(2), pp.29-52.

Cutler, D.M. and Lleras-Muney, A., 2006. Education and Health: Evaluating Theories and Evidence. National Poverty Center Working Paper Series# 06-19. National Poverty Center, University of Michigan.

Cutler, D.M., Lange, F., Meara, E., Richards-Shubik, S. and Ruhm, C.J., 2011. Rising educational gradients in mortality: the role of behavioral risk factors. Journal of health economics, 30(6), pp.1174-1187.

Cutler, D.M., Lleras-Muney, A. and Vogl, T., 2008. Socioeconomic status and health: dimensions and mechanisms.

Deaton, A.S. and Paxson, C.H., 1998. Aging and inequality in income and health. The American Economic Review, 88(2), pp.248-253.

De Walque, D., 2007. Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. Journal of health economics, 26(5), pp.877-895.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K. and Oregon Health Study Group, 2012. The Oregon health insurance experiment: evidence from the first year. The Quarterly journal of economics, 127(3), pp.1057-1106.

Finlay, Keith, and Katie R. Genadek. 2021. “Measuring All-Cause Mortality With the Census Numident File.” *American Journal of Public Health* 111 (S2): S141–48. https://doi.org/10.2105/AJPH.2021.306217.

Fischer, Pamela J, and William R Breakey. 1991. “The Epidemiology of Alcohol, Drug, and Mental Disorders Among Homeless Persons.” *American Psychologist*.

Garibaldi, B., Conde-Martel, A. and O'Toole, T.P., 2005. Self-reported comorbidities, perceived needs, and sources for usual care for older and younger homeless adults. Journal of general internal medicine, 20(8), pp.726-730.

Gelberg, L., Linn, L.S., Usatine, R.P. and Smith, M.H., 1990. Health, homelessness, and poverty: a study of clinic users. Archives of Internal Medicine, 150(11), pp.2325-2330.

Grimard, F. and Parent, D., 2007. Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?. Journal of health economics, 26(5), pp.896-926.

Grossman, M. 1972. On the Concept of Health Capital and the Demand for Health. Journal of Political Economy, 80(2), 223–255.

Henwood, Benjamin F., Thomas Byrne, and Brynn Scriber. 2015. "Examining Mortality among Formerly Homeless Adults Enrolled in Housing First: An Observational Study." *BMC Public Health* 15 (1): 1209. https://doi.org/10.1186/s12889-015-2552-1.

Hibbs, Jonathan R., Lawrence Benner, Lawrence Klugman, Robert Spencer, Irene Macchia, Anne Mellinger, and Daniel Fife. 1994. "Mortality in a Cohort of Homeless Adults in Philadelphia." *New England Journal of Medicine* 331 (5).

Hwang, Stephen W. 1997. "Causes of Death in Homeless Adults in Boston." *Annals of Internal Medicine* 126 (8): 625. https://doi.org/10.7326/0003-4819-126-8-199704150-00007.

Kasprow, Wesley J., and Robert Rosenheck. 2000. "Mortality Among Homeless and Nonhomeless Mentally Ill Veterans." *The Journal of Nervous and Mental Disease* 188 (3).

Kitagawa, E.M. and Hauser, P.M., 1973. Differential mortality in the United States: A study in socioeconomic epidemiology. Harvard University Press.Koegel, Paul, Greer Sullivan, Audrey Burnam, Sally C. Morton, and Suzanne Wenzel. 1999. "Utilization of Mental Health and Substance Abuse Services Among Homeless Adults in Los Angeles:" *Medical Care* 37 (3): 306–17. https://doi.org/10.1097/00005650-199903000-00010.

Kushel, M. and Moore, T., 2023. Towards a New Understanding: The California Statewide Study of People Experiencing Homelessness.

Layne, Mary, and Deborah Wagner. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." CARRA Working Paper Series #2014-01. Washington, D.C: U.S. Census Bureau. https://www.census.gov/content/dam/Census/library/working-papers/2014/adrm/carra-wp-2014-01.pdf.

Lee, B.A., Tyler, K.A. and Wright, J.D., 2010. The new homelessness revisited. Annual review of sociology, 36, pp.501-521.

Lleras-Muney, A., 2005. The relationship between education and adult mortality in the United States. The Review of Economic Studies, 72(1), pp.189-221.

Metraux, Stephen, Nicholas Eng, Jay Bainbridge, and Dennis P. Culhane. 2011. "The Impact of Shelter Use and Housing Placement on Mortality Hazard for Unaccompanied Adults and Adults in Family Households Entering New York City Shelters: 1990–2002." *Journal of Urban Health* 88 (6): 1091–1104. https://doi.org/10.1007/s11524-011-9602-5.

Meyer, Bruce D, Angela Wyse, and Kevin Corinth. 2022. "The Size and Census Coverage of the U.S. Homeless Population." *NBER Working Paper Series*, no. No. 30163.

Meyer, Bruce D., Angela Wyse, Alexa Grunwaldt, Carla Medalia, and Derek Wu. n.d. "Learning about Homelessness Using Linked Survey and Administrative Data." *NBER Working Paper Series* Working Paper 28861.

Meyer, Bruce D., Angela Wyse, Alexa Grunwaldt, Gillian Meyer, and Derek Wu. 2023. "The Income and Safety Net Participation of the U.S. Homeless Population." *Working Paper*.

National Institute of Justice. 2023. "National Missing and Unidentified Persons System: Cases Created in NamUs: January 2023."

North, Carol S., Karin M. Eyrich, David E. Pollio, and Edward L. Spitznagel. 2004. "Are Rates of Psychiatric Disorders in the Homeless Population Changing?" *American Journal of Public Health* 94 (1): 103–8. https://doi.org/10.2105/AJPH.94.1.103.

O'Flaherty, Brendan. 2019. "Homelessness Research: A Guide for Economists (and Friends)." *Journal of Housing Economics* 44 (June): 1–25. https://doi.org/10.1016/j.jhe.2019.01.003.

Roncarati, Jill S., Travis P. Baggett, James J. O'Connell, Stephen W. Hwang, E. Francis Cook, Nancy Krieger, and Glorian Sorensen. 2018. "Mortality Among Unsheltered Homeless Adults in Boston, Massachusetts, 2000-2009." *JAMA Internal Medicine* 178 (9): 1242. https://doi.org/10.1001/jamainternmed.2018.2924.

Roncarati, Jill S., James J. O'Connell, Stephen W. Hwang, Travis P. Baggett, E. Francis Cook, Nancy Krieger, and Glorian Sorensen. 2020. "The Use of High-Risk Criteria to Assess Mortality Risk among Unsheltered Homeless Persons." *Journal of Health Care for the Poor and Underserved* 31 (1): 441–54. https://doi.org/10.1353/hpu.2020.0032.

Ruiz, John M., Patrick Steffen, and Timothy Smith. 2013. "Hispanic Mortality Paradox: A Systematic Review and Meta-Analysis of the Longitudinal Literature." *American Journal of Public Health* 103 (3).

Schinka, John A., Robert M. Bossarte, Glenn Curtiss, William A. Lapcevic, and Roger J. Casey. 2016. "Increased Mortality Among Older Veterans Admitted to VA Homelessness Programs." *Psychiatric Services* 67 (4): 465–68. https://doi.org/10.1176/appi.ps.201500095.

Schinka, John A., Katherine C. Leventhal, William A. Lapcevic, and Roger Casey. 2018. "Mortality and Cause of Death in Younger Homeless Veterans." *Public Health Reports* 133 (2): 177–81. https://doi.org/10.1177/0033354918755709.

Schultz, T.W., 1962. Reflections on investment in man. Journal of political economy, 70(5, Part 2), pp.1-8.

Smith, J.P., 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. Journal of Economic perspectives, 13(2), pp.145-166.

Trick, William E., Fred Rachman, Keiki Hinami, Jennifer C. Hill, Craig Conover, Lisa Diep, Howard S. Gordon, et al. 2021. "Variability in Comorbidites and Health Services Use across Homeless Typologies: Multicenter Data Linkage between Healthcare and Homeless Systems." *BMC Public Health* 21 (1): 917. https://doi.org/10.1186/s12889-021-10958-8.

U.S. Department of Housing and Urban Development. 2022. "The 2022 Annual Homelessness Assessment Report to Congress, Part 1: Point-in-Time Estimates of Homelessness." Department of Housing and Urban Development.

U.S. Government Accountability Office. 2016. "Missing Persons and Unidentified Remains: Opportunities May Exist to Share Information More Efficiently."

Wyse, Angela, and Bruce D. Meyer. 2023. "Medicaid and the Mortality of the U.S. Homeless Population." *Working Paper*.

9. Appendix

9.1 Potential bias from unidentified decedents

Deceased individuals whose identities cannot be determined do not appear in our Numident death records, potentially understating our mortality estimates. This possibility is likely to occur disproportionately among those experiencing homelessness, as these individuals do not have a fixed address and may be less connected to other people in comparison to housed individuals. At the same time, as Meyer et al. (2023) emphasize, even the homeless are connected to employers and government social services at a high rate, with more than 95 percent either engaged in formal work or receiving social insurance or welfare benefits.[15] To investigate the scope of potential understatement of mortality among the homeless, we explored the extent of unidentified decedents in the U.S., also sometimes referred to as John and Jane Doe deaths.

Two primary federal databases supported by the Department of Justice contain data related to missing and unidentified people in the United States: the National Crime Information Center (NCIC) database and the National Missing and Unidentified Persons System (NamUs). While access to the NCIC database is restricted to authorized agencies, NamUs is open to be viewed by the public. Both, though they are separate and unconnected systems, contain information on long-term unidentified persons, defined as cases where the decedent's identity has not been determined for more than 30 days. Notably, while all cases contained in NCIC should theoretically be captured by NamUs (though this is not always the case), NamUs also accepts and maintains additional records of missing and unidentified persons that may not be found in NCIC because, for instance, they may not yet have been filed with law enforcement, and because NamUs aggregates information from law enforcement, criminal justice agencies, coroners, and families of missing persons alike (U.S. Government Accountability Office 2016). NamUS numbers of unidentified remains tend to be higher, consistent with this broader sourcing, and so are likely a more reliable upper bound.

Ideally, we would like to obtain an estimate of the number of unidentified decedents each year and the share of these who were experiencing homelessness, but neither federal database

[15] Specifically, Meyer et al. (2023) find that 97 percent of Census shelter users and 93 percent of those who are unsheltered are enrolled in at least one safety net program (Medicaid, SNAP, OASDI, SSI, or Veterans' Connected Disability) or were formally employed in the year observed as homeless.

allows us to calculate these numbers directly. In 2023, however, NamUs began publishing a monthly report indicating the number of unidentified persons cases created and resolved each month (National Institute of Justice 2023). In January of that year, 106 cases were created and 41 resolved, indicating a net increase of 65 persons, or 780 per year if extrapolated across twelve months. We obtain an estimate of the share of unidentified decedents who were experiencing homelessness by restricting records in the NamUs database to those in which the word "homeless" appears in the Circumstances of Recovery section of the record.[16] This suggests about 2.7 percent, or 384 of NamUs's stock of 14,382 unidentified decedents (those found in 1915 to 2022) were experiencing homelessness at the time of death. Consistent with our earlier conjecture, this share is more than an order of magnitude higher than the homeless share of the population. Multiplying this share by the annual estimate of 780 unidentified persons gives a back-of-the-envelope estimate of about 21 unidentified homeless decedents each year. This number is dwarfed by the approximately 3,500-7,000 people in our homeless sample who die each year.

Incompleteness of the NamUs database and the presence of homeless decedents missed by our database filtering procedure could cause this number to be an underestimate, but this bias would have to be extremely large to be a cause of concern for the findings in this paper. We may also wish to consider the number of gross, not net, additions to the NamUs database each year, if we think that resolved cases reflect long-deceased individuals, although such an adjustment would increase our estimated annual number of unidentified homeless individuals to just 34. Given the magnitude of available estimates, unidentified deaths seem likely to be a small source of bias in our findings.

[16] These cases generally appear to be ones in which an individual was mentioned to be homeless by other people at the scene familiar with the individual, an individual appeared to be homeless due to appearance and belongings, or a person was found near a homeless encampment.

Tables

Table 1:

Prior Estimates of Homeless Individuals' Relative Mortality Risk					
Author(s) and Year	**Location and Collection Period**	**Mortality Period**	**Sample Demographics and Mortality Data**	**Comparison Group, Mortality Data**	**Estimates (Standardized Mortality Rates)**
Baggett et al. (2013)	Boston, 2003-2008	2003-2008	28,033 sheltered and unsheltered, ages 18-64, the universe served by Boston Health Care for the Homeless Program (BHCHP) between 2003-2008. Mortality data from Massachusetts Department of Public Health death occurrence files.	2003-2008 Massachusetts population. Mortality rates obtained from CDC WONDER.	Race-adjusted totals: Ages 25-44: 8.6 (men), 9.6 (women) Ages 45-64: 4.5 (both genders) Ages 65-84: 1.1 (both genders)
Barrow et al. (1999)	NYC, 1987	1987-1994	1,260 sheltered in 1987, ages 18+, randomly selected from bed rosters in 22 municipally run congregate shelters in NYC and systemically selected from food and clothing lines in 4 shelters. Mortality data from the National Death Index.	1987-1994 U.S. and NYC populations. Mortality rates obtained from the CDC's mortality files.	Total age-adjusted totals in NYC: 2.2 (men), 3.7 (women)
Metraux et. al. (2011)	NYC, 1990-2002	1990-2007	160,525 sheltered adults, ages 18-74, with a record of first entering a homeless shelter run by NYC DHS from 1990-2002 and who had SSNs (universe); among families, one adult selected who was designated by DHS as head of household. Mortality data from Social Security Death Index.	None to housed population; only compares those who are homeless in families to those who are homeless as single adults.	Age- and sex-adjusted totals (no comparisons to general population): Males (family/single): 0.56 Females (family/single): 0.28
Roncarati (2018)	Boston, 2000	2000-2009	445 unsheltered adults in Boston, ages 18-81, seen face-to-face by BHCHP street team in 2000 (universe). Mortality data from Massachusetts Department of Public Health death occurrence files and, at times, the National Death Index.	Massachusetts housed population; sheltered adult homeless cohort. Mortality rates obtained from CDC WONDER.	Age-standardized totals: Relative to MA general population: 9.8 Relative to sheltered homeless: 2.7
Hibbs et. al. (1994)	Philadelphia, 1985-1988	1985-1988	6,308 sheltered and unsheltered, ages 15-74, all served by one or both of two agencies for the homeless (mental health program and Philadelphia Office of Services for Homeless Adults) between 1985 and 1988 (universe). Mortality data from Pennsylvania Department of Health.	Philadelphia housed population. Mortality rates obtained from census data from Pennsylvania Department of Health.	Age-weighted (but not race-weighted) totals: Relative to general Philadelphia Population: 3.5
Hwang et al. (1997)	Boston, 1988-1993	1988-1993	17,292 sheltered and unsheltered, ages 18-64, all served by BHCHP between July 1988 and December 1993. Mortality data from Massachusetts death registry.	Boston housed population. Mortality rate data source for housed population unclear.	Non-adjusted totals: 18-24: 5.9 (men), 11.8 (women) 25-44: 3.0 (men), 3.9 (women) 45-64: 1.6 (men), 1.5 (women)

Tables

Table 2:

Summary Statistics: Demographic Characteristics and Region

	Homeless (Census)		Housed Poor (ACS)		Housed (Census)	
Age in 2010	Ages 18+	Ages 18-54	Ages 18+	Ages 18-54	Ages 18+	Ages 18-54
Mean	45.1	39.6	41.9	33.6	47.3	37.1
Ages 18-24	0.101	0.132	0.236	0.310	0.121	0.181
25-29	0.078	0.103	0.108	0.142	0.088	0.132
30-34	0.078	0.102	0.091	0.120	0.084	0.126
35-39	0.086	0.112	0.084	0.111	0.086	0.129
40-44	0.119	0.155	0.082	0.108	0.091	0.136
45-49	0.151	0.197	0.082	0.108	0.099	0.149
50-54	0.153	0.200	0.077	0.102	0.098	0.148
55-59	0.110		0.063		0.087	
60-64	0.059		0.054		0.074	
65-69	0.029		0.033		0.055	
70 and older	0.037		0.089		0.118	
Gender, Race, and Ethnicity						
Female	0.312	0.327	0.586	0.574	0.519	0.508
White	0.523	0.511	0.655	0.641	0.768	0.735
Black	0.379	0.388	0.212	0.218	0.124	0.137
Other Race	0.098	0.102	0.133	0.141	0.108	0.128
Hispanic	0.153	0.161	0.216	0.230	0.129	0.155
Region						
Northeast	0.230	0.231	0.159	0.153	0.185	0.182
Midwest	0.174	0.176	0.216	0.222	0.221	0.219
South	0.279	0.282	0.408	0.404	0.366	0.367
West	0.318	0.312	0.217	0.221	0.228	0.233
Weighted Count	341,800	261,500	14,110,000	10,740,000	2,182,000	1,454,000
N	181,000	140,000	158,000	110,000	2,000,000	1,313,000

Notes: Weighted counts reflect inverse probability weighting adjustment to account for non-linkage for all three groups. For housed poor, weighted count also reflects survey weights, and for overall housed, weighted count is adjusted to reflect one percent random sampling from the 2010 Census housed population. All reported ages reflect age in 2010.

Tables

Table 3:

Summary Statistics: Disability, Economic Status, Family Connections, and State

	Homeless (Census)		Housed Poor (ACS)		Housed (Census)	
Age in 2010	Ages 18+	Ages 18-54	Ages 18+	Ages 18-54	Ages 18+	Ages 18-54
SSI receipt (2009)	0.205	0.189	0.113	0.093	0.030	0.028
DI receipt (2009)	0.092	0.081	0.066	0.052	0.032	0.023
SSI or DI	0.229	0.206	0.135	0.107	0.049	0.039
Employed in 2009	0.443	0.489				
Top Half of Prior Income	0.494	0.491				
Has Spouse or Former Spouse	0.149	0.143				
Also Recorded in Housing	0.306	0.286				
Has Child	0.266	0.307				
Any Indicator of Family Connection	0.501	0.507				
New York	0.115	0.117	0.066	0.063	0.065	0.065
California	0.185	0.180	0.117	0.118	0.119	0.123
Other State	0.700	0.704	0.817	0.818	0.816	0.812
Sheltered Homeless	0.469	0.492				
Weighted Count	341,800	261,500	14,110,000	10,740,000	2,182,000	1,454,000
N	181,000	140,000	158,000	110,000	2,000,000	1,313,000

Notes: Weighted counts reflect inverse probability weighting adjustment to account for non-linkage for all three groups. For housed poor, weighted count also reflects survey weights, and for overall housed, weighted count is adjusted to reflect one percent random sampling from the 2010 Census housed population. All reported ages reflect age in 2010.

Tables

Table 4:

Cumulative Mortality April 2010-March 2022 (Ages 18-54 in 2010)

Based on Empirical Survivor Function (No Controls)

	Homeless	Housed Poor	Housed
Share died	0.1575	0.0619	0.0385
Probability of dying relative to housed	4.09	1.60	1.00
Probability of dying relative to housed poor	2.55	1.00	0.62

Based on Covariate-Adjusted Survivor Function (Age and Gender Controls)

	Homeless	Housed Poor	Housed
Share died	0.1620	0.1037	0.0503
Probability of dying relative to housed	3.22	2.06	1.00
Probability of dying relative to housed poor	1.56	1.00	0.48

Tables

Table 5:

Average Annual Mortality Hazard by Group in Two Years Before and During COVID-19 Pandemic (Ages 18-54 in 2010)

Empirical Mortality Hazard (No Controls)			
	Homeless	Housed Poor	Housed
April 2018-March 2020	0.0157	0.0061	0.0038
April 2020-March 2022	0.0226	0.0088	0.0054
Change without accounting for aging of population			
Absolute increase	0.0069	0.0027	0.0016
Proportional increase	43.87%	44.84%	41.77%
Change accounting for aging of population			
Absolute increase	0.0052	0.0021	0.0011
Proportional increase	33.25%	33.91%	29.76%
Covariate-Adjusted Mortality Hazard			
	Homeless	Housed Poor	Housed
April 2018-March 2020	0.0163	0.0105	0.0050
April 2020-March 2022	0.0239	0.0154	0.0071
Change without accounting for aging of population			
Absolute increase	0.0076	0.0049	0.0021
Proportional increase	46.39%	46.68%	42.32%
Change accounting for aging of population			
Absolute increase	0.0057	0.0037	0.0015
Proportional increase	35.12%	35.17%	30.13%

Notes: Covariate-adjusted mortality hazard controls for difference in age and gender distribution between groups. Increase accounting for aging of population is equal to the increase without accounting for aging minus eight times the estimated coefficient from a regression of the hazard in the first 20 periods on a time trend, which yields an estimate of the change in the average annual hazard between these two time periods attributable to the aging of our cohort.

Tables

Table 6:

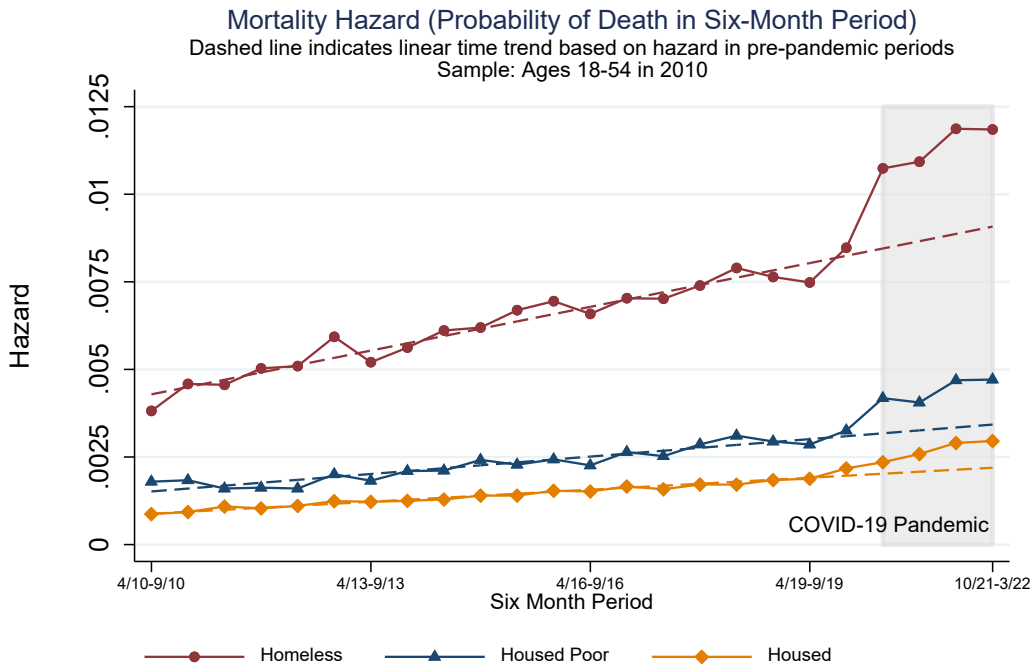
Average Annual Mortality Hazard by Homeless Sub-Group in Two Years Before and During COVID-19 Pandemic (Ages 18-54 in 2010)

Covariate-Adjusted Mortality Hazard				
	Sheltered Males	Unsheltered Males	Sheltered Females	Unsheltered Females
April 2018-March 2020	0.0176	0.0179	0.0140	0.0150
April 2020-March 2022	0.0262	0.0260	0.0189	0.0200
Change without accounting for aging of population				
Absolute increase	0.0086	0.0081	0.0049	0.0050
Proportional increase	49.00%	45.44%	35.17%	33.59%
Change accounting for aging of population				
Absolute increase	0.0067	0.0063	0.0033	0.0034
Proportional increase	38.35%	35.22%	23.71%	22.42%

Note: Covariate-adjusted mortality hazard controls for difference in age and gender distribution between groups. Increase accounting for time trend is equal to the increase without accounting for time trend minus eight times the estimated coefficient from a regression of the hazard in the first 20 periods on a time trend, which yields an estimate of the change in the average annual hazard between these two time periods attributable to the aging of our cohort.

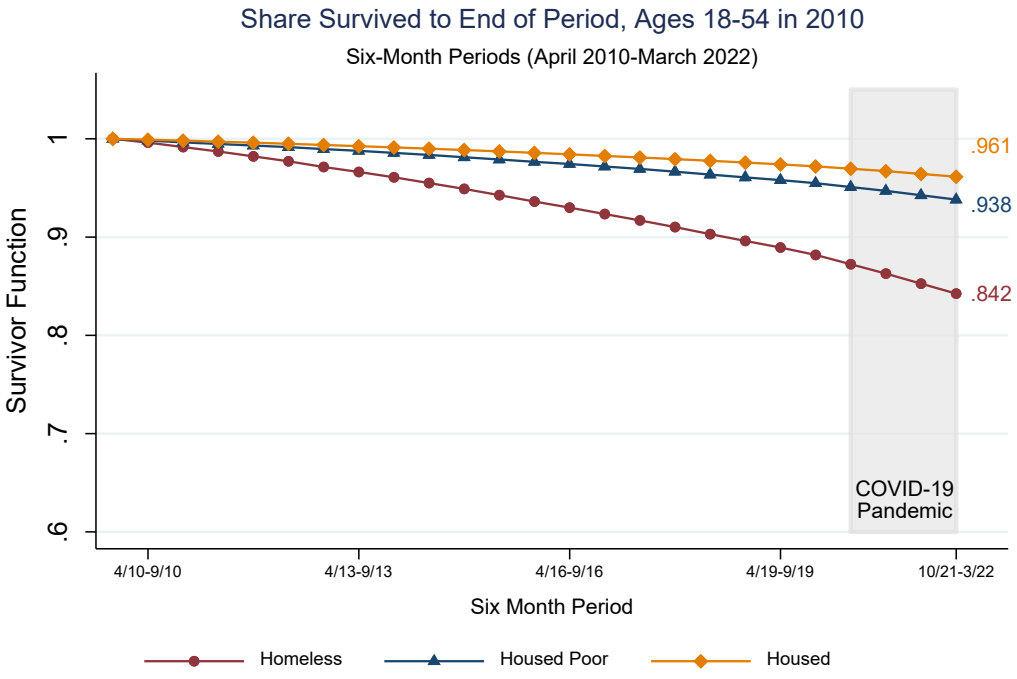
Figures

Figure 1:



Sources: 2010 Decennial Census, 2022 SSA Numident.

Figure 2:



Sources: 2010 Decennial Census, 2022 SSA Numident.

Figures

Figure 3:

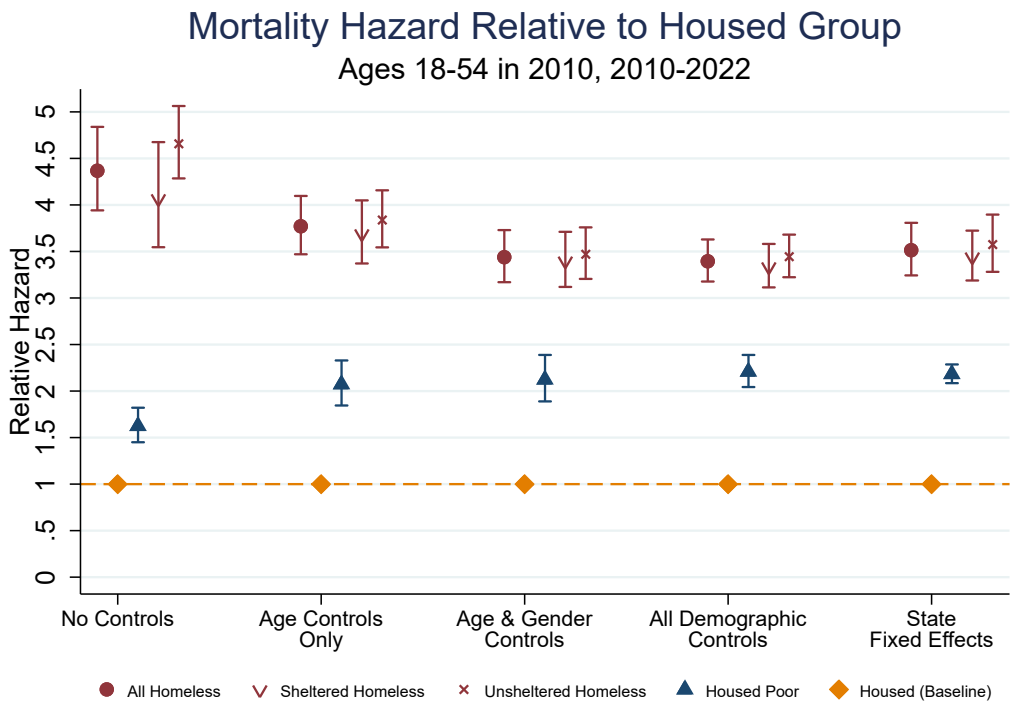
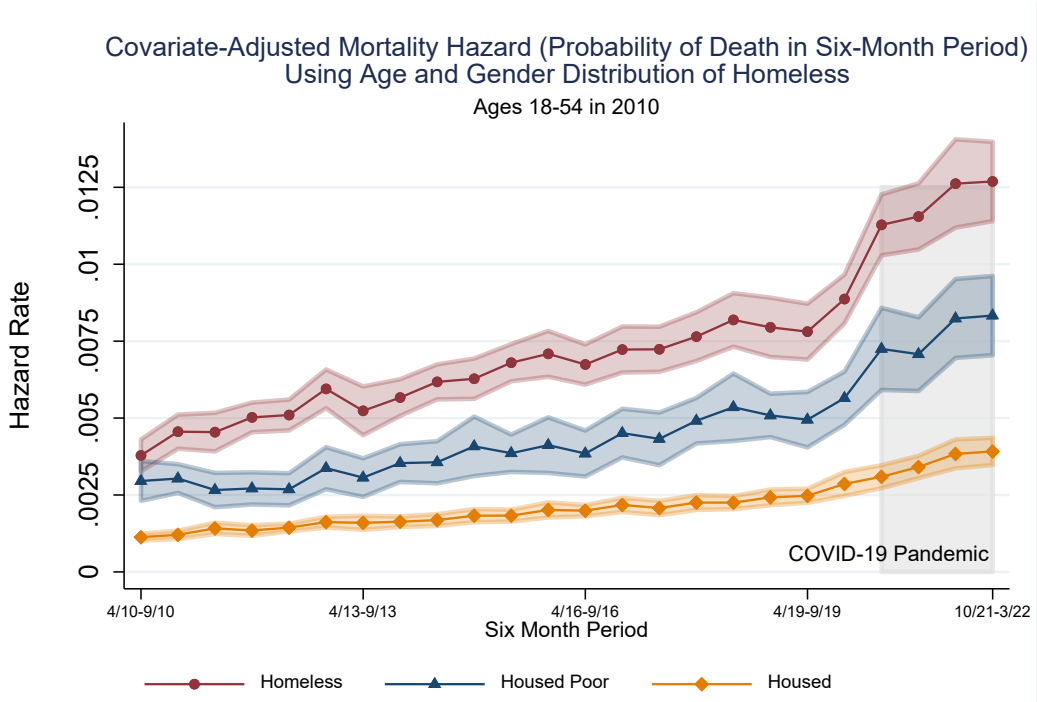


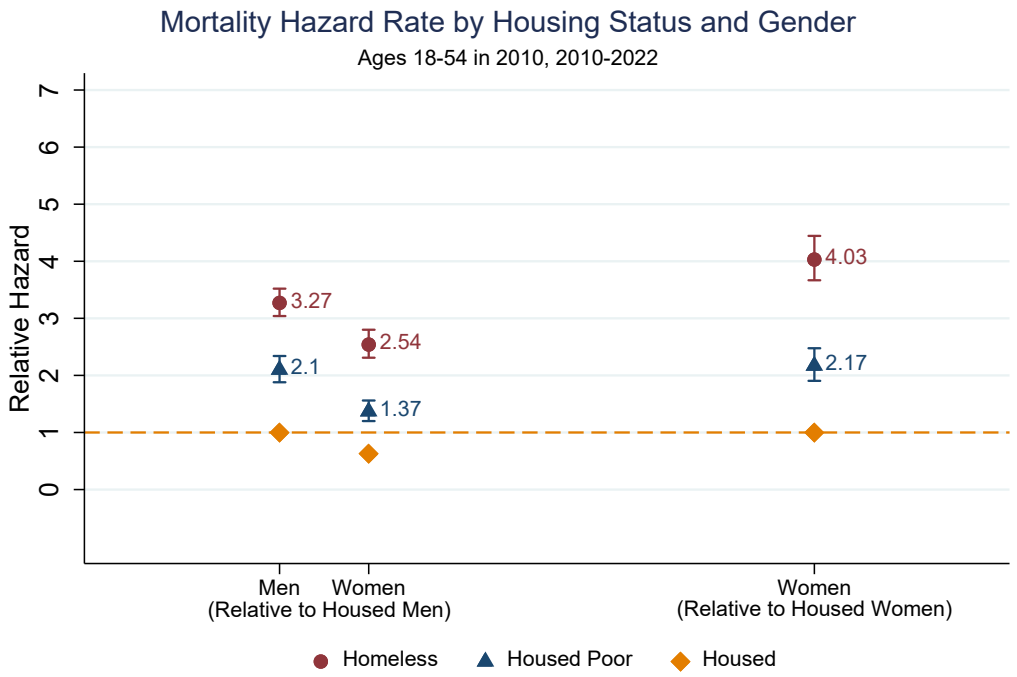
Figure 4:



Sources: 2010 Decennial Census, 2022 SSA Numident.

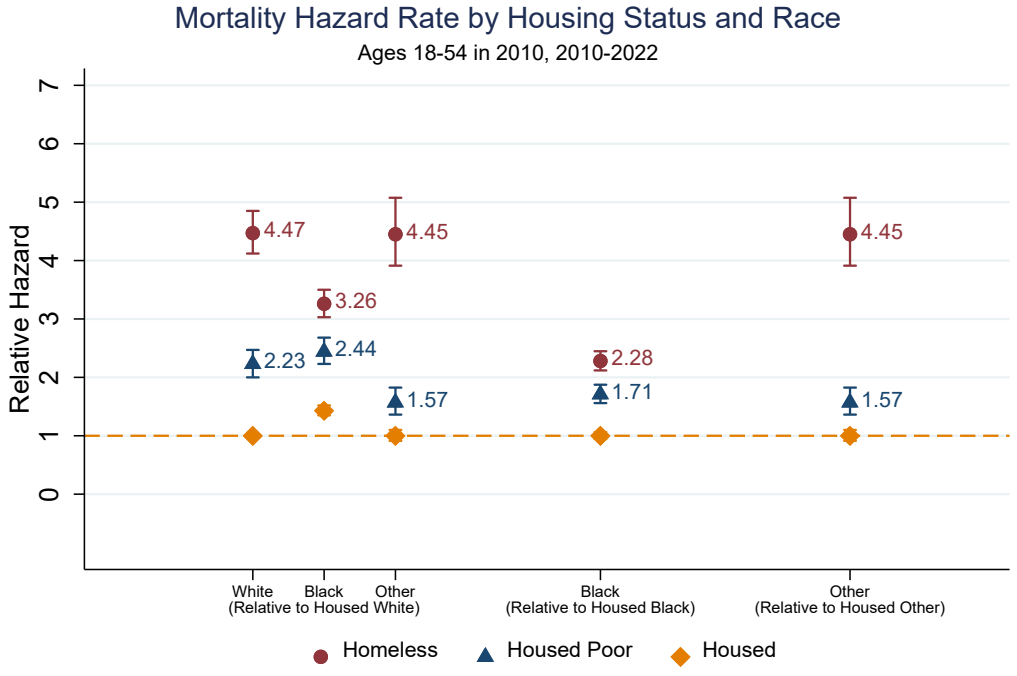
Figures

Figure 5:



Note: Plot displays coefficient from hazard model including age controls.

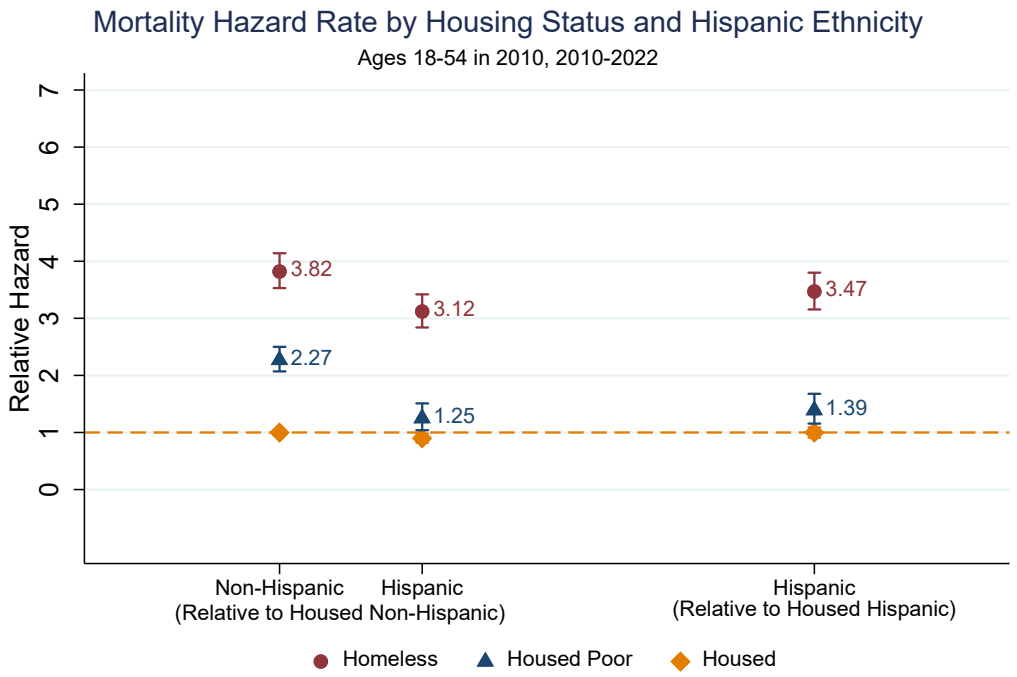
Figure 6:



Note: Plot displays coefficient from hazard model including age controls.

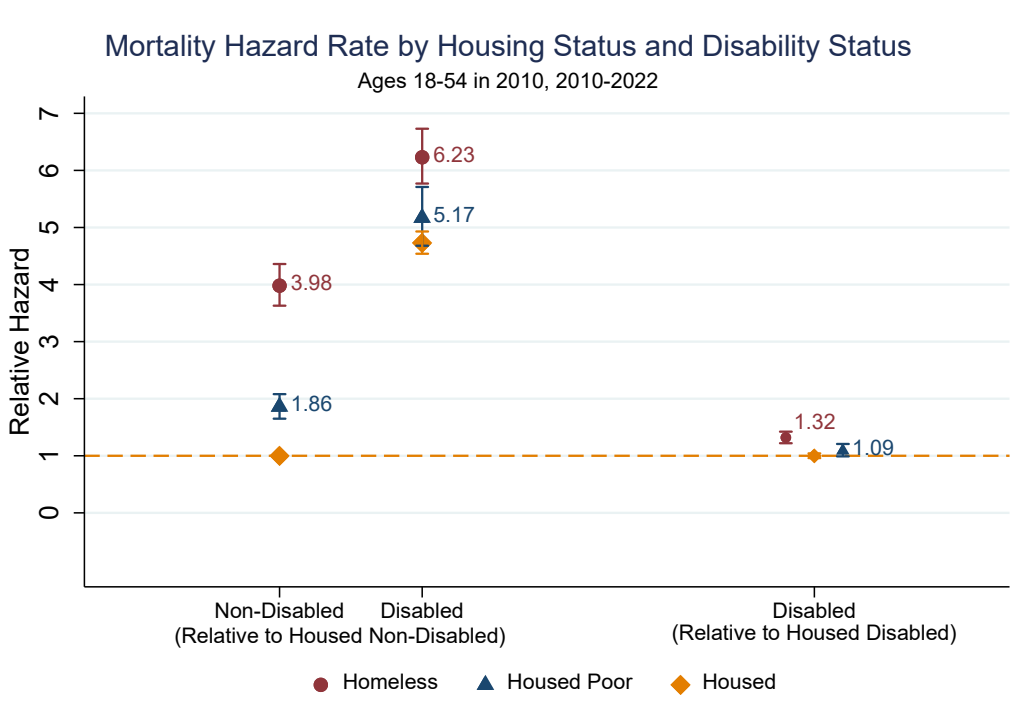
Figures

Figure 7:



Note: Plot displays coefficient from hazard model including age controls.

Figure 8:



Note: Plot displays coefficient from hazard model including age controls.

Figures

Figure 9:

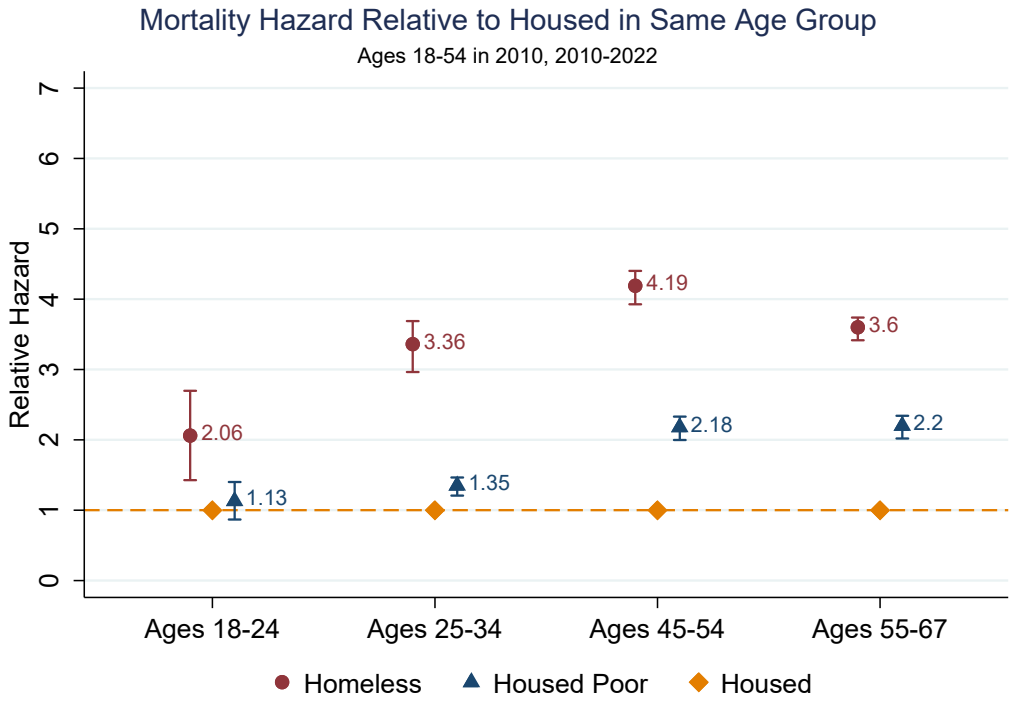
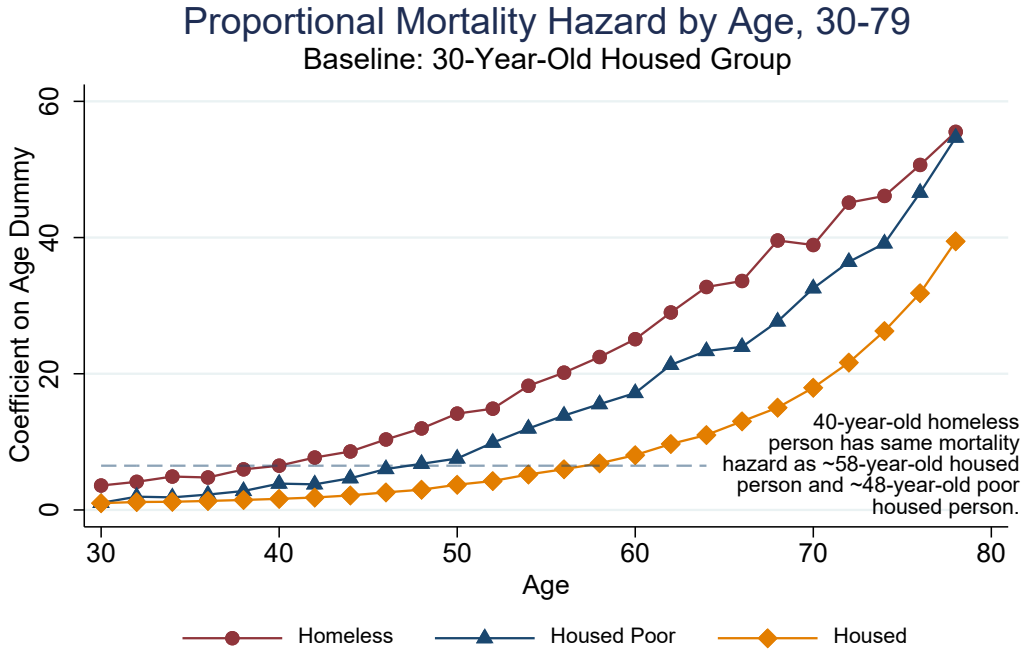


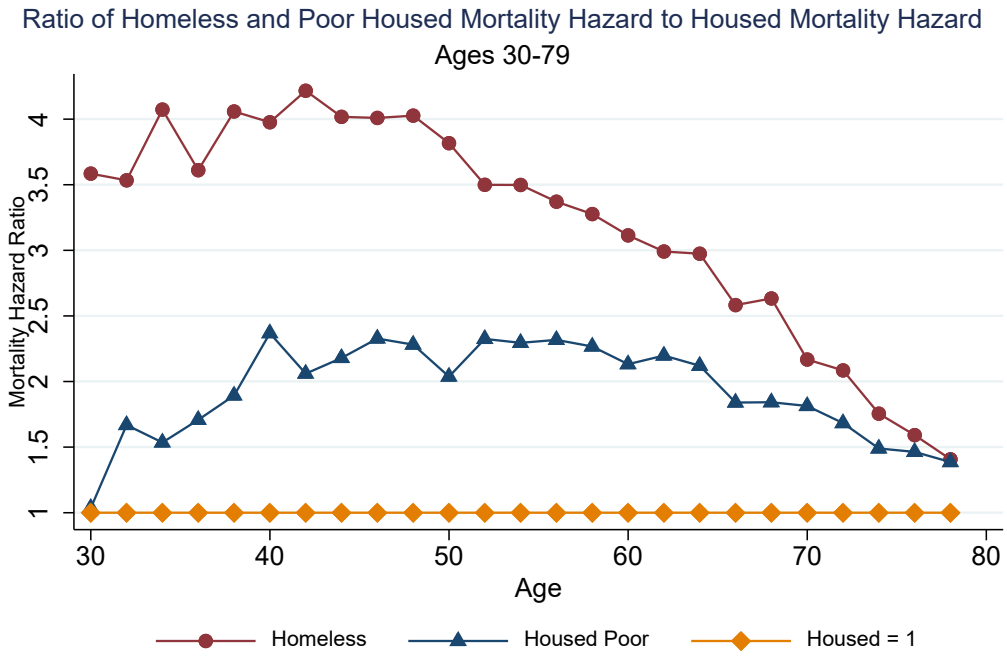
Figure 10:



Note: Figure displays coefficient on two-year age dummy with group interaction in model assuming common baseline hazard and controlling for gender. Samples includes poeple ages 30-79 in 2010-2022.

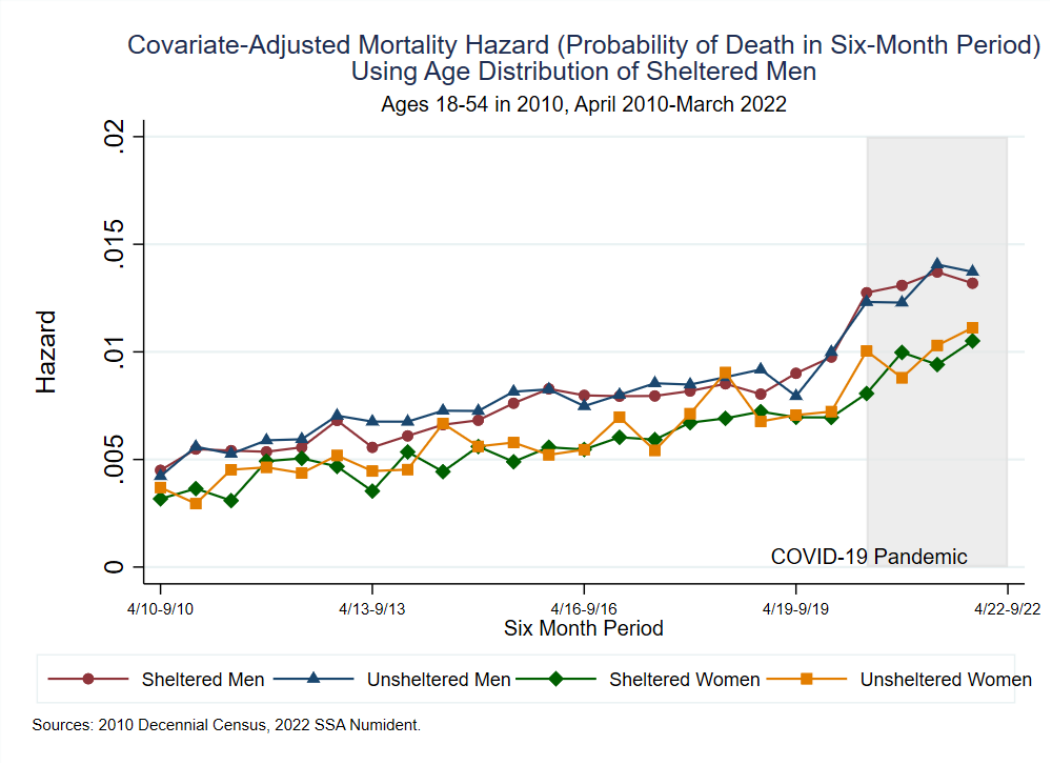
Figures

Figure 11:



Note: Figure dispalys ratio of coefficients on two-year age dummy with group interaction in model assuming common baseline hazard and controlling for gender. Sample includes poeple ages 30-79 in 2010-2022.

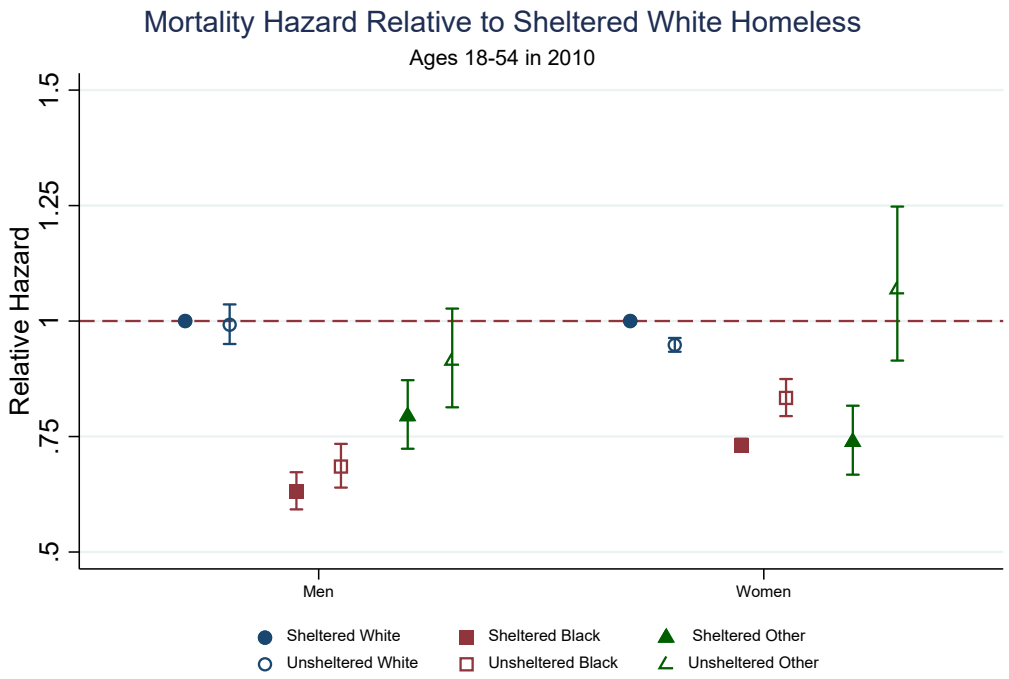
Figure 12:



Sources: 2010 Decennial Census, 2022 SSA Numident.

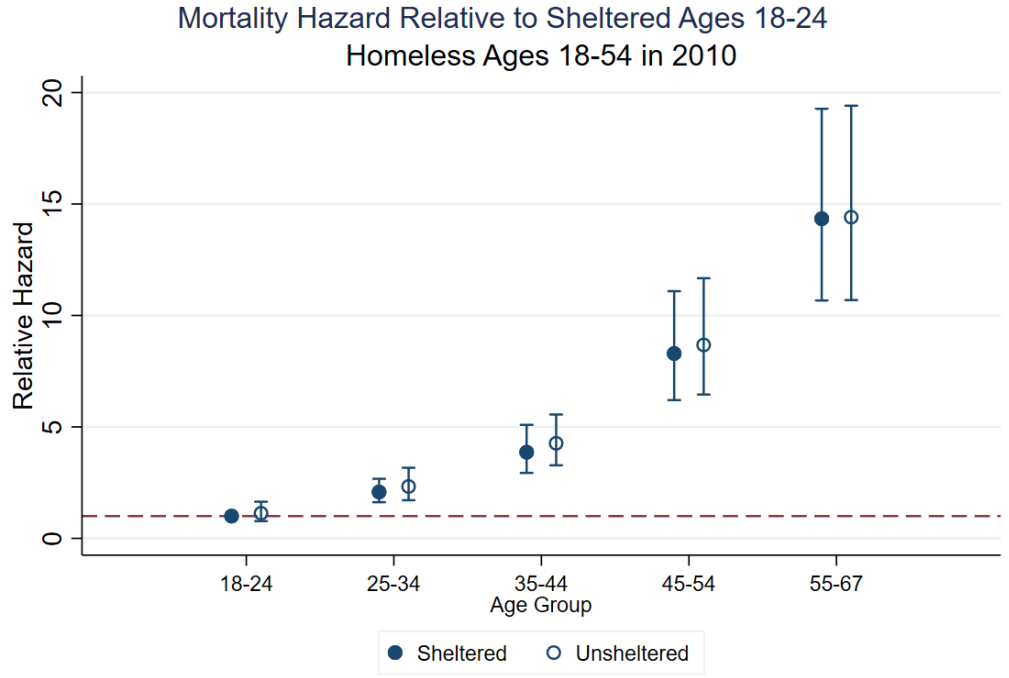
Figures

Figure 13:



Note: Plot displays coefficient from hazard model including age controls.

Figure 14:



Note: Plot displays coefficient from hazard model controlling for gender.

Figures

Figure 15: