NBER WORKING PAPER SERIES

THE CAUSAL EFFECTS OF PLACE ON HEALTH AND LONGEVITY

Tatyana Deryugina
David Molitor

Working Paper 29321
http://www.nber.org/papers/w29321

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2021

We thank Erik Hurst, Nicolai Kuminoff, Jonathan Skinner, Timothy Taylor, and Heidi Williams for helpful suggestions and comments and Celestina Edleman and Yifan Wang for excellent research assistance. This paper was prepared for the Journal of Economic Perspectives. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under award numbers R21AG050795, P01AG005842, and R01AG053350. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Institutes of Health or the National Bureau of Economic Research.

The Causal Effects of Place on Health and Longevity
Tatyana Deryugina and David Molitor
NBER Working Paper No. 29321
October 2021
JEL No. H75,I1,R1

ABSTRACT

Life expectancy varies substantially across local regions within a country, raising conjectures that place of residence affects health. However, population sorting and other confounders make it difficult to disentangle the effects of place on health from other geographic differences in life expectancy. Recent studies have overcome such challenges to demonstrate that place of residence substantially influences health and mortality. Whether policies that encourage people to move to places that are better for their health or that improve areas that are detrimental to health are desirable depends on the mechanisms behind place effects, yet these mechanisms remain poorly understood.

Tatyana Deryugina
Department of Finance
University of Illinois at Urbana-Champaign
515 East Gregory Drive, MC-520
Champaign, IL 61820
and NBER
deryugin@illinois.edu

David Molitor
University of Illinois at Urbana-Champaign
340 Wohlers Hall
1206 S. Sixth Street
Champaign, IL 61820
and NBER
dmolitor@illinois.edu

High life expectancy is a hallmark of economic development. Between 1960 and 2018, life expectancy at birth in the United States increased by nine years, from 69.7 to 78.7 years (Bastian et al. 2020). This average masks large and persistent geographic differences in mortality and life expectancy within the United States (Murray et al. 2006; Chetty et al. 2016; Currie and Schwandt 2016; Dwyer-Lindgren et al. 2017; Mokdad et al. 2018; Woolf and Schoomaker 2019). For example, life expectancy is more than 10 years higher in the top 1 percent of counties by life expectancy compared to the bottom 1 percent of counties (Dwyer-Lindgren et al. 2017). The geographic variation in life expectancy is particularly large among individuals in the lowest quartile of income (Chetty et al. 2016).

In this essay, we begin with an overview of geographic differences in life expectancy across the United States and Europe. We then discuss the problems with seeking to either infer underlying place health effects—defined as the hypothetical effect on health of relocating an individual from one location to another—or explain the causal mechanisms behind the geographic variation by regressing life expectancy on the characteristics of the area and the local population, a "naïve regression" approach. A more promising approach that uses data on movers to identify causal effects of place on health has been developed in recent years, and we draw from this literature to illustrate how one might detect and measure place health effects. Finally, we discuss some possible mechanisms behind place effects.

The extent to which the observed geographic variation in life expectancy across places reflects the causal effects of place of residence is ex-ante uncertain because there are two (not mutually exclusive) ways through which such variation could arise. First, geographic variation in life expectancy could be due to non-random geographic sorting of individuals, with geographic differences reflecting the exogenous characteristics of residents. Second, place of residence could have a causal effect on longevity through a variety of channels, which may operate through largely immutable local characteristics (like climate in that area) or through characteristics amenable to policy (like the local health care system or exposure to pollution). To make matters more complicated, geographic sorting can itself give rise to place effects through peer influence on health-relevant behaviors, yielding geographic differences in life expectancy that are a product of both non-random sorting and the peer influences of individuals who live there. Understanding the contribution of each of these factors is of paramount importance for crafting optimal policy.

When considering how place shapes health, most of the motivation and empirical literature has focused on longevity, arguably because it is easier to measure systematically than other aspects of health. Additionally, life expectancy gains are worth a lot (Murphy and Topel 2006), making longevity a natural first-order concern for researchers and policymakers. We therefore focus most of our discussion on how place of residence affects mortality while recognizing that place could also affect other important dimensions of health.

Regional Differences in Life Expectancy: United States versus Europe

As a basis for gauging the potential role of place in determining life expectancy, we compare how life expectancy varies across local regions in the US and Europe. For the US, we measure life expectancy at birth in 2010–2015 for each county using data from the US Small-Area Life Expectancy Estimates Project (Arias et al. 2018). The average population of the approximately 3,200 US counties and county equivalents is roughly 100,000, but this varies greatly from Los Angeles County with about 10 million people to the smallest counties with no more than a few hundred people.

For Europe, we measure life expectancy at birth for geographies defined by the Nomenclature of Territorial Units for Statistics (NUTS), using the NUTS 3 measure, which is the most geographically refined level in this system and most comparable to US counties. There are about 1,500 NUTS 3 regions in 37 countries across Europe, with population ranging from about 150,000 to 800,000. Life expectancy data are not systematically reported at the NUTS 3 level; we compile the data for 1,057 regions in 22 countries from various sources using the most recent period available for each region (Deryugina and Molitor 2021). In some analyses, we use life expectancy in 2018 at the NUTS 2 level, the next-highest level of geographic aggregation because these data are available for all European countries covered by NUTS except Albania. Online Appendix Sections A.1 and A.2 and Tables A.1 and A.2 describe the life expectancy data and methodology in greater detail.

As a vivid example of geographic differences in mortality across the US, Fuchs (1974) compared mortality rates in Nevada and Utah, neighboring states with similar climates and, at the time, income levels and physicians per capita. Fuchs (1974) noted that, nonetheless, adult mortality rates were substantially higher in Nevada than in Utah, which he attributed to Nevada's high rates of cigarette and alcohol consumption as well as "marital and geographical instability." Even today, the average person born in Utah has a life expectancy that is 1.9 years higher than the average person born in Nevada.

More generally, Figure 1 and column (1) of Table 1 reveal that life expectancy at birth varies widely across US counties, from a low of 69.1 years in East Carroll Parish in northeastern Louisiana to a high of 89.5 years in Cheyenne County, Colorado (Online Appendix Table A.3 lists the top and bottom 10 counties, by life expectancy). County-level life expectancy averages 77.7 years and has a standard deviation of 2.6 years. The top 10 percent of counties have a life expectancy of 81.0 years or more, while the bottom 10 percent have a life expectancy of 74.4 years or less, resulting in an interdecile range of 6.5 years. Although there is some spatial correlation in life expectancy—counties with the lowest life expectancy tend to be in the South—roughly half of the variation in life expectancy across all counties occurs between states, with the other half occurring across counties within states.[1]

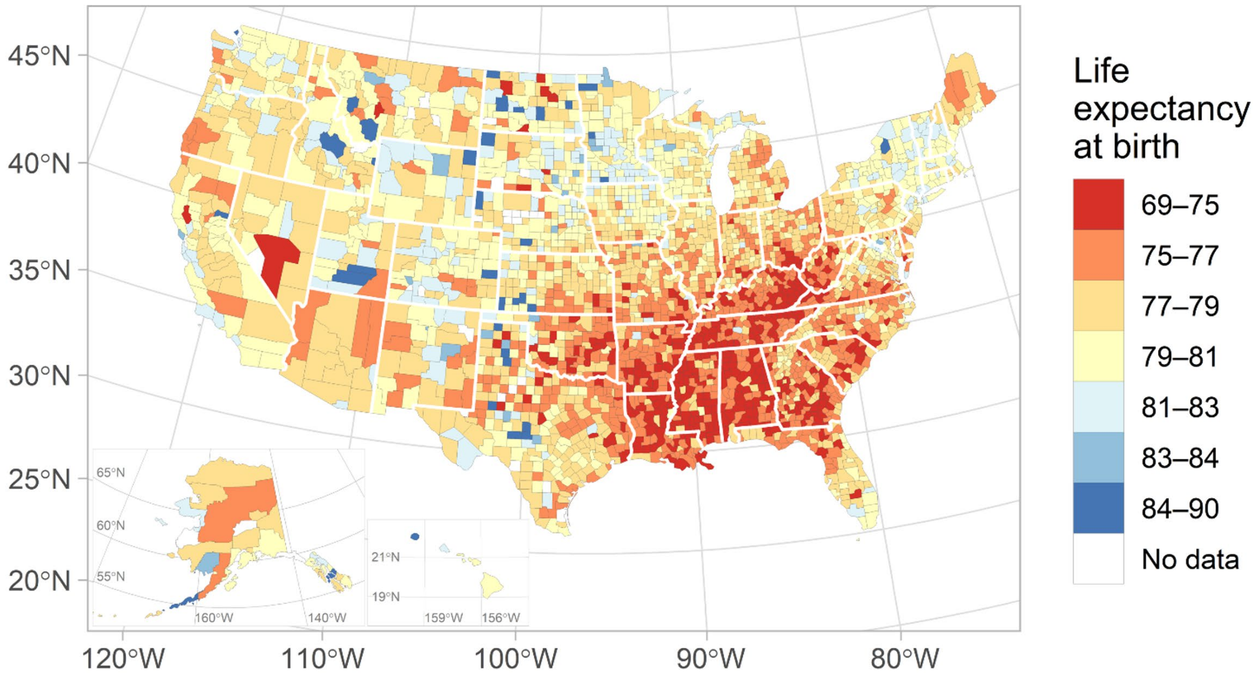
Table 1: Regional life expectancy in the United States and Europe

	(1)	(2)
	United States	Europe
Geographies		
Regional unit of analysis	County	NUTS 3
Number of regions	3,108	1,057
Number of states and District of Columbia (US only)	51	
Number of countries	1	22
Life expectency at birth		
Period	2010–2015	2011–2019
Mean	77.7	80.6
Standard deviation	2.6	2.5
Minimum	69.1	70.4
10th percentile	74.4	76.7
90th percentile	81.0	83.2
Maximum	89.5	84.8

Notes: The table reports summary statistics for the US and European life expectancy data samples.

[1] The R-squared from regressing county-level life expectancy on state fixed effects is 0.46, revealing that just over half (54 percent) of the variation in life expectancy across counties occurs within states, with the remainder occurring across states. See Online Appendix Section A.3 for details of this regression.

Figure 1: US life expectancy

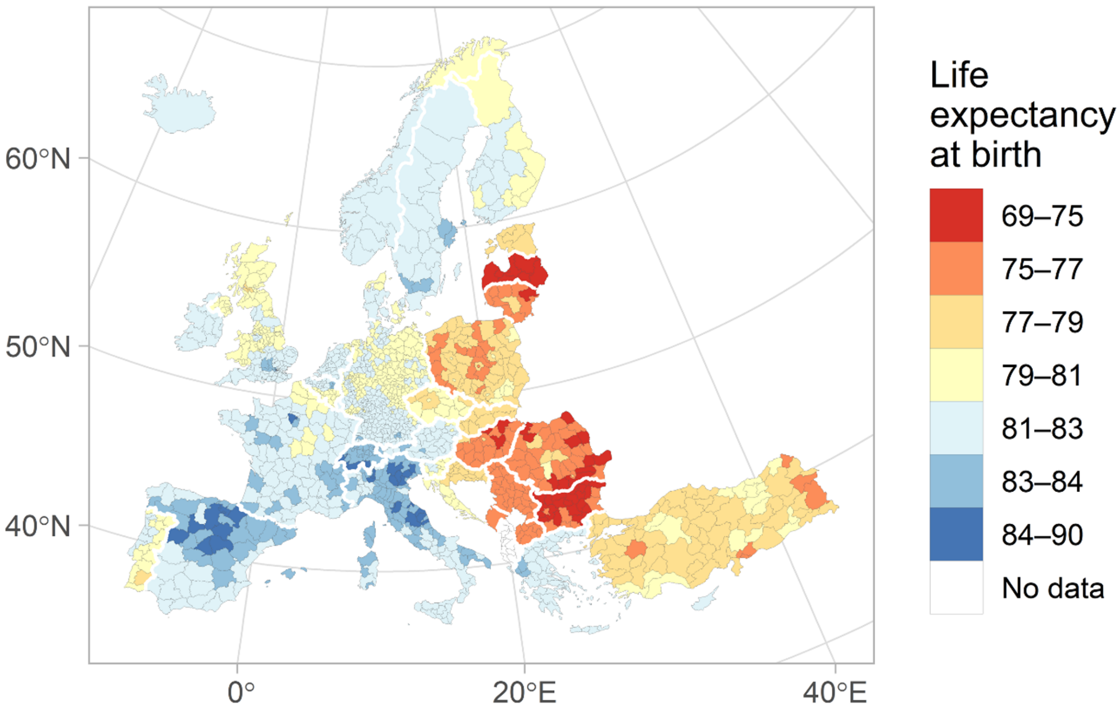


Notes: The figure shows county-level life expectancy at birth in the United States.

Figure 2 and column (2) of Table 1 show the geographic distribution in life expectancy in Europe, which ranges from a low of 70.4 years in Latgale, Latvia, to a high of 84.8 years in a region that makes up a portion of Madrid, Spain. Average life expectancy in the NUTS 3 regions in our sample is 80.6 years, with a standard deviation of 2.5 years. Like the US, the interdecile range is 6.5 years: the top 10 percent of NUTS 3 regions have a life expectancy of 83.2 years or more, while the bottom 10 percent have a life expectancy of 76.7 years or less. Unlike the US, approximately 87 percent of the variation in life expectancy across Europe is explained by between-country variation rather than within-country variation.[2]

[2] The R-squared values from a regression of NUTS 3 or NUTS 2 level life expectancy on country fixed effects are 0.87 and 0.85, respectively, which reveals that 85–87 percent of the variation in life expectancy across European regions is explained by the country of residence. See Online Appendix Section A.3 for details of this regression.

Figure 2: European life expectancy



Notes: The figure shows life expectancy at birth in Europe. Gray borders delineate NUTS 3 regions. Life expectancy is shown at the NUTS 3 whenever available and at the NUTS 2 level otherwise. Online Appendix Figures A.1 and A.2 show separate maps for the NUTS 2 and NUTS 3 samples, respectively.

Three main results emerge from comparing the regional variation in life expectancy in the US and Europe. First, average life expectancy is 2.8 years higher in Europe than in the US. Second, the overall variation in life expectancy, as captured by the standard deviation or interdecile range of the life expectancy distribution, is similar in both contexts. Third, most of the regional variation in life expectancy in Europe is explained by country of residence, whereas in the US most of the variation is within-state.

The reasons behind the large differences in the spatial correlation in local-area life expectancy between Europe and the US are not immediately clear. American states are arguably more similar to each other in terms of policies than are European countries. Thus, country-level place health effects could be more heterogeneous in Europe compared to state-level place health effects in the US, while within-country place health effects could be less heterogeneous than within-state. However, population sorting and individual preferences could also be much more heterogeneous across European countries than across US states. Thus, the life expectancy patterns seen in Europe and the US do not necessarily help rule out or support any particular explanation.

However, the geographic variation in European life expectancy—which, to our knowledge, has not been comprehensively documented at such a granular spatial level until now—demonstrates that large regional differences in life expectancy are not just a US phenomenon.

A Naïve Regression Approach

One approach to investigating whether place affects life expectancy, and if so, how, is to regress local life expectancy on the characteristics of the area and the population. We conduct such an exercise here, similar in approach to Dwyer-Lindgren et al. (2017) except we consider a somewhat different set of local characteristics and weight regressions by each county's population, as life expectancy is likely to be measured with greater error in smaller counties.[3] Because a simple regression of local life expectancy on local area and population characteristics cannot account for a number of important confounders, we dub it a "naïve regression approach."

Table 2, Panel A shows the results of bivariate regressions of US county-level life expectancy on a variety of local health and environmental characteristics. Life expectancy is positively correlated with the percent of population that exercises and is negatively correlated with smoking and obesity rates. The local smoking rate alone explains over 46 percent of the cross-sectional variation in life expectancy, as indicated by the R-squared (column (3)). Obesity and exercise rates individually explain about 42 and 34 percent of the variation, respectively.

Health care quantity—as measured by the number of doctors per capita and the number of hospital beds per capita—each explain 6.0 and 4.4 percent of local life expectancy, respectively. The correlation between life expectancy and the number of hospital beds per capita, however, does not have the expected sign: more hospital beds is associated with *lower* life expectancy. As we discuss further below, this counterintuitive correlation hints at the difficulties inherent in recovering the mechanisms behind place health effects. In this case, for example, more hospital beds could be a response to poor health and elevated health care needs among residents.

[3] We focus on the local characteristics considered by Deryugina and Molitor (2020), except those derived from Medicare claims. See Online Appendix Section A.4 for data and regression details and Deryugina and Molitor (2021) for the data. Online Appendix Table A.5 shows the results of unweighted regressions.

Table 2: The relationship between county-level life expectancy and local characteristics

	(1)	(2)	(3)
County characteristic	Characteristic mean [sd]	OLS coeficient (se)	R-squared
A. Health and environmental characteristics			
Percent smoking	21.29	-0.36	0.462
	[4.05]	(0.02)	
Percent obese	20.10	-0.34	0.423
	[4.12]	(0.02)	
Percent exercising	74.74	0.23	0.341
	[5.44]	(0.02)	
Physicians per 1,000 capita	2.77	0.28	0.060
	[1.94]	(0.05)	
PM 2.5 concentrations	10.38	-0.24	0.049
	[1.94]	(0.08)	
Hospital beds per 1,000 capita	3.40	-0.18	0.044
	[2.55]	(0.03)	
Hot days/year (90°F+)	2.21	-0.01	0.001
	[8.59]	(0.01)	
Hospital quality index	0.78	0.95	0.001
	[0.06]	(1.71)	
B. Economic characteristics			
Median home value ($1,000s)	128.87	0.02	0.490
	[65.91]	(0.00)	
Income per capita ($1,000s)	21.63	0.24	0.344
	[5.28]	(0.01)	
Poverty rate, 65+	0.10	-20.97	0.181
	[0.04]	(3.36)	
Upward income mobility (from p25	-0.03	1.88	0.124
	[0.41]	(0.34)	
Urban population share	0.79	2.47	0.081
	[0.25]	(0.33)	
Crime rate per 1,000	7.62	-0.14	0.052
	[3.49]	(0.03)	
Local gov. spending per capita	2.51	0.46	0.050
	[1.06]	(0.15)	
Upward income mobility (from p75	-0.03	-1.63	0.029
	[0.23]	(0.60)	
Social capital index	-0.46	0.25	0.016
	[1.11]	(0.11)	
Income segregation	0.07	6.61	0.011
	[0.03]	(2.87)	
C. Multivariate comparisons			
All health and environmental characteristics			0.695
All economic characteristics			0.672
All characteristics			0.807

Notes: The table reports results from regressing US county-level life expectancy on local characteristics. Each row in the table corresponds to a separate regression, where the included local characteristic(s) are indicated by the row labels. Observations are weighted by county population. Column (1) shows the mean and standard deviation (in brackets) of the local characteristic. Column (2) reports regression coefficients and robust standard errors (in parentheses). Column (3) reports the R-squared from the regression. Online Appendix Section A.4 provides more details on the data and regressions.

Finally, local environmental quality, as measured by fine particulate matter ($PM_{2.5}$) concentrations, explains almost 5 percent of the geographic variation in life expectancy. Climate, as measured by the number of hot (>90°F) days per year, explains only about 0.1 percent of the geographic variation and has no statistically significant relationship with local life expectancy.

Panel B considers bivariate regressions of local life expectancy on economic characteristics. Median home values explain almost half of the geographic variation in life expectancy, and local income per capita explains over a third. The elderly poverty rate explains about 18 percent of the geographic variation, and upward income mobility from the 25th percentile explains about 12 percent. The share of the population living in an urban area explains about 8 percent of the variation, and per-capita spending by the local government and the local crime rate each explain about 5 percent. Finally, upward income mobility from the 75th percentile and income segregation explain 2.9 and 1.1 percent of the geographic variation, respectively.

Most of the characteristics mentioned above are significantly correlated with each other. In Panel C, we report the R-squared values from regressing life expectancy on bundles of characteristics. All the health and environmental characteristics combined explain 69.5 percent of the cross-sectional variation in life expectancy, while all the economic characteristics explain 67.2 percent. The most complete regression that includes all the characteristics mentioned above explains 80.7 percent of the variation in life expectancy. These results are similar to those of Dwyer-Lindgren et al. (2017), who use a slightly different set of health and socioeconomic characteristics to conclude that these characteristics explain as much as 74 percent of the unweighted county-level variation in life expectancy.

A key threat to concluding from the results in Table 2 that place of residence has a causal effect on health is that people with different health prospects and behaviors may endogenously sort into different locations. As emphasized by Roback (1982), individuals choose where to live based in part on the amenities in each region. If people who are predisposed to good health place a higher value than others on amenities that will extend their life expectancy, such as low levels of air pollution, then regional health differences arise partly because of sorting, and observed health differences will overstate the causal impacts of place. By contrast, if relatively unhealthy individuals place a high value on such amenities, regional differences in health will understate the causal role of place.

One possible approach to account for population sorting is to control for individual characteristics and behavior to see how much of the regional difference in health remain unexplained. Early literature investigating place health effects did exactly this, attributing the residual variation to place-specific health effects (see Macintyre, Ellaway, and Cummins 2002 for an overview). However, this approach to measuring place health effects will not yield correct estimates without extremely restrictive assumptions, such as that certain characteristics reflect only sorting while others only capture the causal effect of place. Yet individual behaviors, characteristics like education and income, and even the demographic composition in a place could reflect the effects of living in that place (Macintyre, Ellaway, and Cummins 2002). As a result, health differences that seem to be explained by individual characteristics could also arise through the causal effects of place on choices, behavior, and aging. Furthermore, if health is transmitted intergenerationally, observed geographic differences in health could reflect sorting of individuals' ancestors, further complicating estimation.

Given the abovementioned issues, what do we learn from naïve regressions like those reported in Table 2? On the one hand, one might argue that many of the correlations between local characteristics and life expectancy reflect place effects, perhaps via peer effects or amenities like a local economy that causes incomes to be higher or lower. On the other hand, if differences in socioeconomic status and behavioral patterns are not *caused* by place of residence, referring to them as "place effects" is a misnomer. For example, these differences could be due partly to sorting and population movement, including from a long time ago, creating cross-county differences that have less to do with place of residence and more to do with genetic predisposition and upbringing not directly related to the geographic location. In this case, exogenously moving individuals to a particular location would not alter their health, and therefore these kinds of geographic differences cannot be considered true "place effects."

This list of concerns about the interpretation of regression results can easily be expanded. Ultimately, one might just conclude that correlations are not causation, and no reliable lessons about place effects of health can ever emerge from a naïve regression approach.

Using Movers to Identify Causal Effects of Place

Gauging the Magnitude of Place Health Effects

Empirical difficulties notwithstanding, economic theory predicts that place health effects are likely to exist. For example, a spatial sorting model in the style of Roback (1982) posits that places differ in their local amenities, such as the climate, pollution levels, and the quantity and type of leisure opportunities. Some amenities, like a mild climate or a good harbor, facilitate certain types of firm production in that location; other amenities, like clean air assured through local government anti-pollution regulations, will increase the cost of production and decrease output, all else equal. In the most basic spatial sorting models, individuals can move costlessly and choose where to live based on local wages, rents, and amenity levels. In equilibrium, utility is equalized across places, but differences in the quantity and productivity of local amenities give rise to spatial heterogeneity in wages and rents. The resulting variation in amenities and disposable income in turn implies that longevity is likely to be influenced by one's place of residence. But are the health effects of place generated in this way likely to be large in magnitude?

Gauging the magnitude of place health effects ex-ante is difficult for several reasons. Accounting for the distinct role of sorting in local life expectancy is challenging, not least because individuals sorting based on an area's amenities may (directly or indirectly) give rise to place health effects and do so in offsetting ways. For example, while regions with high amenities for certain kinds of production will have higher wages, sorting of people into that area will raise rents in those regions (and perhaps suppress real wages for workers in other industries), potentially counteracting the direct effects of higher incomes on health to some extent. In practice, the direct and indirect effects are unlikely to offset fully in all places, especially because people do not value all amenities solely for their effects on longevity. A related challenge is that there may be peer effects in health-relevant behaviors like smoking or exercise that exacerbate the impact of any population sorting on the equilibrium levels of spatial heterogeneity in life expectancy. Finally, heterogeneity in local amenities as well as heterogeneity in individuals' preferences, productivity, or information can also yield heterogeneity in place health effects. If someone likes a place for its hiking trails and someone else likes it for its lively nightlife, for example, the causal effects of that place on the life expectancy of these two individuals may be of opposite signs.

Thus, even the simplest sorting models predict the existence of place health effects but also make clear that measuring them empirically is challenging. Of course, the strict assumptions of such models are not likely to be satisfied: people do not have full information about health (or other) prospects of potential locations, nor is it costless to move. But the reality of differences in amenities across locations, and the likelihood that these differences will in some way cause differences in health, remains.

As a conceptual experiment, place health effects could be quantified and disentangled from sorting effects by randomly assigning individuals to different places of residence and measuring differences in their subsequent life expectancy. While conducting such an experiment is impractical for many reasons, looking at movers coming from the same place and ending up in different locations, for example, can help separate the causal effects of place from various confounders. To our knowledge, all quasi-experimental papers that speak to the causal effect of place on health leverage movers in some form. An earlier literature compares the health of movers to non-movers, while a more recent one compares different groups of movers.

Spatial equilibrium models caution against a possible pitfall from using movers to identify the causal effects of place on health: if individuals sort into locations in equilibrium, then the movers may be no less selected than individuals who already reside in a particular location. However, an advantage of a design that uses movers is that, as long as movers are observed for a reasonable period of time before a move, such sorting can generally be evaluated and potentially accounted for. We next discuss the research designs of quasi-experimental studies of place health effects in more detail and summarize the conclusions the literature has reached thus far.

Comparing Movers to Non-Movers

In some studies of health effects related to moving, researchers have sought to address identification problems by looking for factors predictive of certain individuals moving but plausibly exogenous with respect to future health. If these two conditions are satisfied, the predictors of moving can be used either as instruments or in a reduced-form way to estimate the causal effects of moving on health. For example, Gibson et al. (2013) exploit a migration lottery to estimate the causal effect of migration from Tonga to New Zealand on blood pressure and the prevalence of hypertension. In a study of the long-run mortality effects of the early twentieth-century Great Migration of African Americans from mostly rural locations in the Deep South to mostly urban locations in the North, Black et al. (2015) use the proximity of individuals' birthplaces to railroad lines as an instrument for migration. Johnson and Taylor (2019) build on this identification strategy and use the timing of railroad construction as well as patterns of postal mail flows to estimate the mortality effects of the mid-twentieth-century migration from rural locations in the Northern Great Plains states to urban locations in the American West and Midwest.

These earlier studies focus on estimating the health effects of migration rather than of place, so their estimates will reflect both the health effects of the act of migrating itself (e.g., due to losses of community ties) as well as the average effect on health of living in the destination regions. More generally, these studies are unable to pin down the exact mechanisms through which migration affects longer-run health. For example, they cannot determine the extent to which the specific composition of origin and destination regions matters for the estimated effects (e.g., the origins are mostly rural and destinations mostly urban in the cases of Black et al. 2015 and Johnson and Taylor 2019).

Nonetheless, these studies provide suggestive evidence that health effects of place exist and are non-trivial in magnitude. Gibson et al. (2013) estimate that Tonga-to-New-Zealand migration raises blood pressure and increases hypertension prevalence by 11 percentage points or about one-third of the mean among lottery losers. Black et al. (2015) find that, conditional on surviving to age 65, leaving the Deep South lowered life expectancy by at least 1.5 years. They also show that the movers smoked and drank significantly more than those who did not migrate. Correspondingly, movers to the North are substantially more likely to die from respiratory cancer, chronic obstructive pulmonary disease, and chronic liver disease and cirrhosis. Likewise, Johnson and Taylor (2019) find that the mid-twentieth-century US migration from rural to urban areas increased mortality and provide suggestive evidence that this is due to increased smoking and alcohol consumption.

Other indirect evidence that local conditions matter for health comes from papers that use movers to study how local conditions affect health care provision and other non-health outcomes that could ultimately affect health. For example, Song et al. (2010) show that when Medicare recipients move between regions, rates of medical diagnoses change. Finkelstein, Gentzkow, and Williams (2016) study Medicare recipients who move between areas and show that place of residence affects movers' medical spending. Molitor (2018) looks at cardiologists who move and finds that, on average, their own practice patterns change by 60–80 percent of the difference in local norms between their new and original practice regions. Mover designs have also shown that local conditions can affect levels of education and earnings (Chetty, Hendren, and Katz 2016; Nakamura, Sigurdsson, and Steinsson 2017; Chyn 2018; Chetty and Hendren 2018). To the extent that each of these factors matters for health, we might therefore surmise that place of residence will have health consequences via such channels. As noted earlier, however, such analysis would need to account for both direct and indirect health effects of living in a higher-income and higher-cost-of-living area, for example.

Comparing Movers to Other Movers

Studies of place health effects that do not use movers cannot thoroughly assess the degree of sorting into a location and therefore cannot control for it without restrictive assumptions (e.g., that sorting only operates through immutable characteristics like race or age). Studies that compare movers to non-movers can make progress on this dimension but nonetheless cannot separate place health effects from the health effects of moving itself.

A substantially more credible research design for estimating place health effects comprises comparing movers to each other, something that several recent studies have done. This research design is based on the insight that if two otherwise identical individuals initially living in the same place simultaneously move to different destinations, then subsequent differences in their health will be due to place effects. Medicare recipients who move between counties offer a promising source of evidence in this area, in part because they can maintain their health insurance coverage when they move. Additionally, Medicare administrative data include the vast majority of US elderly and long-term disabled individuals and provide detailed health utilization records for many of them, allowing researchers to control for differences in observable characteristics before the move and to assess the extent of non-random sorting into destination regions. However, researchers must still overcome the challenge that individuals who move to different destination regions may not be identical and their destination choices may not be exogenous with respect to other, unobserved, determinants of health. This hurdle has proven formidable, and the literature is still in its nascency.

How can researchers overcome the difficulty of movers sorting non-randomly into destinations? Helpfully, identification of place health effects in studies that compare movers to each other does not require that movers choose their destination region completely at random. Instead, identification requires that, conditional on available controls, movers' choice of destination is unrelated to any other future determinant of the health outcome of interest. Studies interested in relating health outcomes to some specific characteristic of place—such as the local mortality or obesity rate—require an even weaker identification assumption to interpret that correlation as proxying for the causal effect of place on health: the destination characteristic of interest must be unrelated to unobserved determinants of future health. For example, movers selecting destinations based on whether they have relatives living there does not confound the research design as long as either the presence of relatives *or* the health effects of living near relatives are orthogonal to the destination characteristic being used as the proxy for place health effects. As we discuss in the next section, however, while a correlation between a destination characteristic and changes in movers' health can be interpreted as demonstrating that place has a causal effect on health if the abovementioned assumption holds, the relationship cannot be interpreted as the causal effect of *that particular characteristic* without additional assumptions.

Directly testing the identification assumptions discussed above is infeasible because one can never be sure of observing all determinants of future health. One indirect test involves estimating whether a destination characteristic of interest is correlated with pre-existing trends in the health outcome(s) of interest. Research using outcomes other than mortality can assess the likelihood of such endogenous sorting by explicitly estimating such trends among movers before the move. For example, Baum et al. (2020) use administrative records from the Veterans Health Administration to study how a mover's probability of having uncontrolled chronic conditions (hypertension, diabetes, obesity, or depression) is affected by the local prevalence of such conditions. The authors show that movers to regions that differ in the prevalence of uncontrolled chronic conditions do not exhibit differential pre-trends in such conditions before the move. Thereafter, moving to a ZIP code with a greater prevalence of a given chronic condition increases the probability of being diagnosed with that condition within three years of the move. The magnitude of the estimated effects varies from 3.1 percent of the change in the local prevalence for obesity to 27.5 percent of the change in the local prevalence for hypertension.

Because one must be alive to move, a direct test of parallel trends in mortality before a move is not possible, and other approaches to assess and control for any differential sorting must be used. Assessing sorting on *predictors* of mortality is one such approach. Deryugina and Molitor

(2020) study how the mortality of Medicare beneficiaries displaced by Hurricane Katrina relates to mortality in their destination county. They construct an index of predicted mortality for each mover from extensive measures of chronic conditions and spending histories and show that movers' predicted mortality is uncorrelated with mortality in their destination. There is, however, an almost one-for-one relationship between movers' realized mortality and the mortality of residents in their destination county, demonstrating that where movers relocated had a causal effect on their longevity.

A sophisticated approach to control for sorting is developed by Finkelstein, Gentzkow, and Williams (2021), who use the relocation of Medicare beneficiaries to estimate the causal effects of place on mortality. Their definition of place consists of "Commuting Zones," which are aggregations of counties chosen to approximate local labor markets.[4] The authors control for sorting using a generalization of the method developed by Oster (2019), which uses variation in an observable variable (in their study, the correlation between choice of destination and observed health characteristics) to adjust for variation in an unobservable variable (in their study, the correlation between choice of destination and unobserved health characteristics). Finkelstein, Gentzkow, and Williams (2021) can estimate place-specific mortality effects and can therefore directly study heterogeneity in place health effects, unlike Baum et al. (2020) and Deryugina and Molitor (2020). Finkelstein, Gentzkow, and Williams (2021) estimate that equalizing health-related place effects across US Commuting Zones would reduce the geographic variation in life expectancy of 65-year-olds by 15 percent.

As discussed earlier, place effects may be heterogeneous across places and across individuals. Consistent with this prediction, Finkelstein, Gentzkow, and Williams (2021) show that estimated place effects on longevity vary widely across the US. Thus, different research designs may arrive at different conclusions because of differences in where the in-sample movers are relocating. It is also possible that the health effects of a given place vary across individuals themselves: for example, Chetty et al. (2016) find that the largest regional health disparities occur among the poorest individuals (those in the bottom 5 percent of income), suggesting place matters more for this group. The presence of heterogeneous place effects implies that who the marginal mover is could affect a study's estimates.

[4] In 2000, the US had 709 Commuting Zones.

Given the small number of studies leveraging movers to estimate the causal effects of place on health and their somewhat heterogeneous methods, a direct comparison of most existing results is difficult. Finkelstein, Gentzkow, and Williams (2021) relate the place effects they estimate to local average life expectancy and find a positive correlation: moving to a place where life expectancy is one year higher causes the mover to live 0.23 years longer, on average. This relationship is substantially smaller than the almost one-for-one relationship estimated by Deryugina and Molitor (2020). If place effects are heterogeneous, which seems likely, then the differences between these two studies could be due to differences in destination regions or in the composition of movers. For example, Hurricane Katrina almost certainly displaced many people who would otherwise not have moved, whereas the movers exploited by Finkelstein, Gentzkow, and Williams (2021) are typical elderly movers. Of course, it is also possible that one or both studies failed to properly account for non-random sorting into destination regions.

While recent research using state-of-the-art movers' design points to a causal relationship between place of residence and health and longevity, these findings must be verified and extended. Both Finkelstein, Gentzkow, and Williams (2021) and Deryugina and Molitor (2020) use Medicare data, and therefore their sample of movers consists of older individuals and, in the case of Deryugina and Molitor (2020), the long-term disabled. Baum et al. (2020) has a sample of US veterans that is younger, on average, but is overwhelmingly male. The extent to which place of residence matters for the health of younger individuals, especially younger women, therefore remains an important question for future research.

Another important shortcoming of the papers estimating place health effects is that none speaks to the welfare impacts of migration, which could differ qualitatively from estimated health effects both because there are costs to moving and because there can be benefits of living in a place other than its effects on health. Whether the observed migration was welfare-improving on net may vary by context. In the case of Finkelstein, Gentzkow, and Williams (2021) and Baum et al. (2020), for example, migration is likely voluntary and thus plausibly welfare-improving. By contrast, many movers studied by Deryugina and Molitor (2020) were forced to move by Hurricane Katrina and may have suffered a welfare reduction on net. But there exists no direct evidence on whether encouraging migration to places with favorable health effects would be welfare-improving.

Channels Through Which Place May Affect Health and Longevity

Prior sections have hinted at the various channels through which place may affect health. We now consider them systematically and in more detail. Understanding these channels will inform whether or how policies can be designed to improve population health. For example, policy implications are rather different if place health effects are driven by immutable local characteristics, such as climate, compared to if they are driven by peer effects or by public policies.

The ideal experiment to understand the mechanisms behind measured place health effects involves exploiting an exogenous change in some local characteristic and estimating the subsequent change in local life expectancy. Such experiments—whether natural or implemented by a researcher—are rare. Estimating the presence and magnitudes of specific mechanisms has thus proven difficult, and the current evidence in this area is mostly suggestive. In addition to the identification assumptions required to establish that place of residence has a causal effect on health, establishing a causal relationship between any specific local characteristic and health in an observational or experimental setting requires an additional assumption: the local characteristic must not be correlated with any other unobserved local determinant of health. Given the variety of local characteristics that may matter for health, it is unlikely that this assumption is valid for any existing study of place health effects.

For example, greater economic activity could both raise residents' incomes and increase air pollution. Even if one can establish that living in that particular area raises life expectancy on net, separating the contribution of higher income from that of higher air pollution is challenging because both are generated by difficult-to-quantify "economic activity." A naïve regression of life expectancy on local air pollution may even yield counterintuitive positive correlations. Such difficulties are not limited to cross-sectional studies: Deryugina and Molitor (2020) indeed find that higher local concentrations of $PM_{2.5}$ are associated with lower mover mortality. Similarly, Finkelstein, Gentzkow, and Williams (2021) estimate that places that are good for longevity tend to have fewer hospital beds per capita.

Identification challenges notwithstanding, a variety of studies that do and do not exploit movers have examined how local characteristics correlate with life expectancy and health and have largely been careful not to interpret them as causal. In part due to statistical power considerations and high degrees of correlation between some local characteristics, most research that uses movers has considered local characteristics separately rather than jointly. Thus, just as estimated place health effects potentially reflect the influence of a bundle of characteristics, the specific local characteristic(s) identified as predictive of place health effects could be proxies for the influence of a group of correlated characteristics.

Some insight about the mechanisms behind place health effects can also be gleaned from exploiting plausibly exogenous region-wide changes in policy, such as smoking or health care regulations. Yet another approach is to use experimental or quasi-experimental methods to study the causal health effects of factors that vary across individuals rather than regions. Numerous such studies exist. While they may be indirectly informative about the mechanisms that could generate the observed place effects, they cannot speak to them directly because of their piecemeal approach. This is because the regional distribution of positive and negative contributors to life expectancy could be such that some factors counteract each other and only a few are important for explaining place effects on aggregate.

There are five broad, interrelated mechanisms that could be generating observed place health effects: socioeconomic status, peer effects, health care delivery, the local environment, and public policy. They are interrelated because of their potential to influence each other; for example, elevating the socioeconomic status of some of an area's residents could affect others through peer effects or through greater demand for pollution reduction. There may also be peer effects among health care providers, which influences the types and intensity of health care provided in different areas (e.g., Molitor 2018). We next discuss each of them in turn and summarize the available evidence on their importance, drawing both from studies that do and do not use movers.

The first mechanism is socioeconomic status, such as income and wealth, employment conditions, and education. While these channels are unlikely to explain the place effects among elderly and disabled individuals, such as those studied by Deryugina and Molitor (2020) and Finkelstein, Gentzkow, and Williams (2021), they may be particularly important components of place health effects for younger and working-age individuals. If income affects health and if places estimated to be good for health also tend to increase younger movers' income and employment, then estimates of place effects based on moves later in life will capture only part of the overall effect of place on longevity. More generally, the effects of place later in life may correspond less than one-to-one with regional mortality outcomes, even if all regional differences in mortality are due to place of residence. An opportunity for future research is therefore to measure how place of residence earlier in life matters for health and longevity.

A considerable body of evidence not directly related to place effects does suggest that socioeconomic status plays a key role in building and maintaining health (Grossman 1972). For example, Frijters, Haisken-DeNew, and Shields (2005) and Lindahl (2005) find that plausibly exogenous income shocks improve self-reported health, and Schwandt (2018) finds that negative wealth shocks due to stock market fluctuations impair physical and mental health and increase mortality. Job separations have been linked to elevated mortality risk for decades post-separation (Sullivan and Von Wachter 2009), and young people first entering the labor market during a recession face higher mortality risks later in life (Schwandt and Von Wachter 2020). The level and quality of education individuals receive can also influence mortality (Buckles et al. 2016; Galama, Lleras-Muney, and van Kippersluis 2018).

Even though the movers in their sample are largely not in the labor force and have completed their formal education, both Deryugina and Molitor (2020) and Finkelstein, Gentzkow, and Williams (2021) find that moving to area with higher socioeconomic status is beneficial for survival. Although these findings could reflect the influence of other local characteristics that are simply correlated with socioeconomic status, living in a higher socioeconomic status area might also offer indirect health benefits if higher socioeconomic status causes such areas to develop amenities that are beneficial for health. For example, proximity to grocery stores or restaurants with nutritious food may facilitate healthy living, and areas with higher socioeconomic status may attract more such establishments. The empirical evidence on this specific mechanism is mixed. Allcott et al. (2019) find that the nutritional quality of purchased groceries is not affected by moves to neighborhoods with greater availability of healthy groceries, and Hut (2020) finds no relationship between average nutrition quality of purchased groceries in a destination and changes in movers' nutritional quality for at least two years following the move. But Currie et al. (2010) find that the presence of a fast-food restaurant near a school raises the probability of obesity among the students. If a similar dynamic operates for adults and if demand for fast food is lower in higher-income or higher-education areas, then the presence of healthier restaurant foods may be one mechanism behind place effects.

The second potential mechanism behind place health effects is peer effects. Moving can change one's peers and in this way give rise to peer effects in health behaviors, ultimately affecting a mover's health. Studies in this domain face well-known identification challenges (Manski 1993; Angrist 2014). Most of the research regarding peer effects on health-related behaviors has been done on students and young people and none has been directly related to place effects. Sacerdote (2001) finds that a randomly assigned college dormitory roommate's drinking behavior does not influence one's own drinking behavior, but overall dormmates' drinking behavior does, suggesting the existence of peer effects at higher levels on this dimension. Of course, such peer effects may look very different outside of a college dormitory. Fletcher (2010) combines an instrumental variables strategy with fixed effects to show that classmates' smoking behavior affects one's own. Card and Giuliano (2013) also find peer effects in smoking among youths. Angrist (2014) notes that the best-identified studies have largely found effects small in magnitude, but this, of course, does not rule out their existence in the context of place effects. Additionally, there may be peer effects along other health-relevant dimensions, such as preventive care utilization or regular health screenings, for which there is virtually no well-identified empirical evidence. Overall, whether peer effects are present among older adults and whether they are large enough to generate meaningful differences in health behaviors on aggregate remains an open question.

Both studies of rural-urban movers discussed earlier conclude that movers increase their consumption of alcohol and tobacco, potentially explaining their decrease in life expectancy (Black et al. 2015; Johnson and Taylor 2019). Baum et al. (2020) find that post-move changes in the prevalence of uncontrolled hypertension, obesity, diabetes, and depression are each significantly correlated with the destination region's prevalence of the same condition.[5] These correlations are consistent with peer effects, but other explanations are possible. Among rural-urban movers, it could be that alcohol and tobacco consumption increased because of higher incomes after moving to the city rather than driven by peer effects. Among movers more generally, health habits of both movers and local residents could simultaneously be affected by a variety of living conditions, including local prices and policy, giving rise to the observed correlations.

[5] Deryugina and Molitor (2020) find that rates of smoking, obesity, and exercise at the destination location are significantly associated with movers' subsequent mortality. Similarly, Finkelstein, Gentzkow, and Williams (2021) find a positive relationship between the effect of a place on mortality and smoking and obesity rates and a negative relationship between place effects and exercise rates.

A third potential mechanism is the quality or quantity of health care delivery, which could have both short- and long-term health effects on residents of all ages. Per-capita health care spending varies substantially across US regions, making it an ex-ante plausible determinant of local life expectancy. Correlational studies have found that, on average, regions with higher levels of per-capita health spending have no better health outcomes than lower-spending regions (for a discussion, see Skinner and Fisher 2010). Finkelstein, Gentzkow, and Williams (2021) find that positive place effects are correlated with higher quality and quantity of health care, but Deryugina and Molitor (2020) find no relationship between movers' mortality and local medical spending or health care quality. The potential for reverse causality and other possible confounders makes it difficult to conclude that either of these correlations are causal. Other studies using quasi-experimental evidence suggest that, at least in some settings, higher health care spending is beneficial for health. Using different identification strategies, Doyle (2011) and Doyle et al. (2015) find that patients randomly hospitalized in a higher-spending region or hospital, respectively, are less likely to die. Doyle, Graves, and Gruber (2017) confirm these findings for inpatient spending but also show that higher outpatient spending by hospitals is associated with lower survival of patients who are randomly transported there.

Health care access also varies geographically in the US and could have meaningful impacts on health. Finkelstein et al. (2012) find that randomly selected recipients of Medicaid reported better physical and mental health after a year with health insurance but find no clinical evidence of better health. Miller, Johnson, and Wherry (2019) show that mortality among the near-elderly fell by almost 10 percent in states that participated in the Affordable Care Act Medicaid expansion compared to states that did not. Abaluck et al. (2020) find that, conditional on being insured, specific health insurance plans affect beneficiaries' mortality rates. Combined with geographic differences in plans' availability, this study suggests another possible mechanism behind place effects.

The fourth potential mechanism behind observed place health effects is environmental quality, which varies considerably across the US. Numerous studies have shown that air pollution has a causal effect on both infant and older adult mortality.[6] Similarly, both abnormally cold and

[6] See, for example, Chay and Greenstone (2003a,b); Currie and Neidell (2005); Currie and Walker (2011); Chen et al. (2013); Knittel, Miller, and Sanders (2016); Barreca, Neidell, and Sanders (2017); Deschênes, Greenstone, and Shapiro (2017); Ebenstein et al. (2017); Deryugina et al. (2019).

abnormally hot temperatures have been shown to increase the mortality rates of the elderly (Barreca et al. 2016; Deschênes and Greenstone 2011; Heutel, Miller, and Molitor 2020). Finkelstein, Gentzkow, and Williams (2021) find that positive place effects are correlated with lower pollution levels and a more moderate climate. However, Heutel, Miller, and Molitor (2020) show that cooler (warmer) places are more adapted to cold (hot) temperatures, implying that even though abnormal temperature raise mortality, the total contribution of the local temperature climate to regional differences in life expectancy may be small.

The fifth mechanism is public policy, which can affect everything from socioeconomic status to air pollution levels. Policy can also affect life expectancy through many channels not discussed above, such as by influencing individuals' smoking and drinking behavior directly or by providing a variety of social safety nets to low-income households. The most direct evidence on the role of policy in influencing the geographic variation in life expectancy comes from Montez et al. (2020), who relate state-level changes in life expectancy to changes in policy over the period 1970–2014. They examine 135 different policies, categorizing each as liberal (defined as increased regulation of the economy by the state or increased protection of marginalized groups) or conservative and then creating 18 time-varying policy indices that group related policies together. The authors find that more liberal policies tend to be associated with improved life expectancy for both men and women and that this relationship is particularly strong for the regulation of private labor, immigration, civil rights, and the environment. Higher tobacco taxes are also associated with increased state-level life expectancy over this time period.

While research on the mechanisms behind place effects has produced some suggestive correlations, the fact that some local characteristics have consistently been shown to be correlated with the life expectancy of movers and non-movers alike does not imply that the literature has successfully identified the mechanisms behind measured place effects because of the likelihood of unobserved confounders. Related areas of research offer stronger evidence on several potential mechanisms, such as income and health care access, but cannot be used to quantify the magnitude of place effects without restrictive assumptions.

Further complicating the study of mechanisms is that their health effects could be heterogeneous. Consistent with this idea, Chetty et al. (2016) document that the standard deviation of commuting-zone life expectancy is 1.4 for men in the bottom income quartile but only 0.70 for men in the top income quartile. While they do find that some of the important correlates of life expectancy are similar for low- and high-income individuals—such as smoking, exercise, and obesity rates—correlations between life expectancy and other local characteristics are sometimes significantly different for these two groups. Similarly, Montez et al. (2019) find that, conditional on birth state and basic demographics, there is little variation in state-level life expectancy for those with at least one year of college but a substantial amount of variation for those without a high school degree. Due to the difficulties with interpreting cross-sectional analyses causally, such patterns of course do not prove that place effects and mechanisms are heterogeneous, but they do provide suggestive evidence that this is the case.

Conclusion

The observed geographic dispersion in life expectancy and evidence from movers between areas strongly suggest that where one lives matters for when one dies. Determining whether place health effects are large or trivially small, however, has not been accomplished until very recently. New evidence comparing movers to other movers to estimate place health effects make it reasonable to conclude that, at least for some groups, place of residence has a sizable effect on health. However, more research is needed to build on these findings and, in particular, to understand the effect of place at younger ages on long-term longevity. Although there are many plausible mechanisms through which these place effects may materialize, the question of what it is exactly that causes some places to be better for health than others has so far not been answered directly by any existing study. Given the conceptual need to have local characteristics be as good as randomly assigned, studies that use quasi-experimental regional variation are necessary to make progress on this dimension.

Can public policy take advantage of place health effects in a way that would improve public health? One possible conclusion to draw from the emergent literature is that helping individuals relocate to places that are better for health could be welfare-improving. An advantage of such a policy is that one only needs to know which places are conducive to good health rather than understand the exact mechanisms behind place health effects. However, given the observed reluctance of individuals to move to higher-wage areas (Kennan and Walker 2011), a program that offers subsidies to those who relocate to more favorable locations is likely to be very costly.

It is also unclear whether individuals who are most likely to move as a result of any given policy are those who would benefit most from positive place effects. If individuals are not taking advantage of place health effects due to imperfect information, however, a welfare improvement at a fairly low cost may be possible. If it is social or family ties that bind individuals to a particular location, then any program that aims to relocate individuals to healthier places would need to be designed in a way that coordinates relocation (or perhaps improves communication and travel) of related individuals or of social networks. Viewing geographic differences in health outcomes through the lens of Roback's (1982) spatial sorting model also offers a reminder that ending up in "unhealthy" places will be at least partially the result of choices that include an array of observed and unobserved factors. Without understanding why individuals have not already relocated themselves to places that would benefit their health, any policy that attempts to influence relocation runs the risk of reducing overall welfare.

An alternative policy goal would be to target health-improving policies to areas that have been shown to be detrimental to health. Without greater understanding of the mechanisms behind place health effects, however, it is unclear which local characteristics such a policy should try to improve. Additionally, given scarce social resources, it is worth considering whether policies that target some other existing inequality (perhaps in wealth or income) would be superior to policies that target life expectancy more directly. Indeed, it may be that targeting income or wealth inequality would reduce inequalities not only in life expectancy but also in other, non-health, dimensions.

References

Abaluck, Jason, Mauricio M. Caceres Bravo, Peter Hull, and Amanda Starc. 2020. "Mortality Effects and Choice across Private Health Insurance Plans." NBER Working Paper 27578.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. 2019. "Food Deserts and the Causes of Nutritional Inequality." *The Quarterly Journal of Economics* 134 (4): 1793–844.

Angrist, Joshua D. 2014. "The Perils of Peer Effects." *Labour Economics* 30: 98–108.

Arias, Elizabeth, Loraine A. Escobedo, Jocelyn Kennedy, Chunxia Fu, and Jodi A. Cisewski. 2018. "US Small-Area Life Expectancy Estimates Project: Methodology and Results Summary." National Center for Health Statistics.

Barreca, Alan I., Matthew Neidell, and Nicholas J. Sanders. 2017. "Long-Run Pollution Exposure and Adult Mortality: Evidence from the Acid Rain Program." NBER Working Paper 23524.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro. 2016. "Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century." *Journal of Political Economy* 124 (1): 105–59.

Bastian, Brigham, Betzaida Tejada-Vera, Elizabeth Arias, et al. 2020. "Mortality Trends in the United States, 1900–2018." National Center for Health Statistics.

Baum, Aaron, Juan Wisnivesky, Sanjay Basu, Albert L. Siu, and Mark D. Schwartz. 2020. "Association of Geographic Differences in Prevalence of Uncontrolled Chronic Conditions with Changes in Individuals' Likelihood of Uncontrolled Chronic Conditions." *JAMA* 324 (14): 1429–38.

Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor. 2015. "The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South." *American Economic Review* 105 (2): 477–503.

Buckles, Kasey, Andreas Hagemann, Ofer Malamud, Melinda Morrill, and Abigail Wozniak. 2016. "The Effect of College Education on Mortality." *Journal of Health Economics* 50: 99–114.

Card, David, and Laura Giuliano. 2013. "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends." *Review of Economics and Statistics* 95 (4): 1130–49.

Chay, Kenneth, and Michael Greenstone. 2003a. "Air Quality, Infant Mortality, and the Clean Air Act of 1970." NBER Working Paper 10053.

Chay, Kenneth, and Michael Greenstone. 2003b. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118 (3): 1121–67.

Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. 2013. "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 110 (32): 12936–41.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–62.

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

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106 (4): 855–902.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children." *American Economic Review* 108 (10): 3028–56.

Currie, Janet, and Hannes Schwandt. 2016. "Mortality Inequality: The Good News from a County-Level Approach." *Journal of Economic Perspectives* 30 (2): 29–52.

Currie, Janet, and Matthew Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics* (3): 1003–30.

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics* 3 (1): 65–90.

Currie, Janet, Stefano DellaVigna, Enrico Moretti, and Vikram Pathania. 2010. "The Effect of Fast Food Restaurants on Obesity and Weight Gain." *American Economic Journal: Economic Policy* 2 (3): 32–63.

Deryugina, Tatyana, and David Molitor. 2020. "Does When You Die Depend on Where You Live? Evidence from Hurricane Katrina." *American Economic Review* 110 (11): 3602–33.

Deryugina, Tatyana, and David Molitor. 2021. "Data and Code for: The Causal Effects of Place on Health and Longevity." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor].

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif. 2019. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction." *American Economic Review* 109 (12): 4178–4219.

Deschênes, Olivier, and Michael Greenstone. 2011. "Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the US." *American Economic Journal: Applied Economics* 3 (4): 152–85.

Deschênes, Olivier, Michael Greenstone, and Joseph S. Shapiro. 2017. "Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program." *American Economic Review* 107 (10): 2958–89.

Doyle, Joseph J. 2011. "Returns to Local-Area Health Care Spending: Evidence from Health Shocks to Patients Far from Home." *American Economic Journal: Applied Economics* 3 (3): 221–43.

Doyle, Joseph J., John A. Graves, Jonathan Gruber, and Samuel A. Kleiner. 2015. "Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns." *Journal of Political Economy* 123 (1): 170–214.

Doyle, Joseph J., John A. Graves, and Jonathan Gruber. 2017. "Uncovering Waste in US Healthcare: Evidence from Ambulance Referral Patterns." *Journal of Health Economics* 54: 25–39.

Dwyer-Lindgren, Laura, Amelia Bertozzi-Villa, Rebecca W. Stubbs, Chloe Morozoff, Johan P. Mackenbach, Frank J. van Lenthe, Ali H. Mokdad, and Christopher J.L. Murray. 2017. "Inequalities in Life Expectancy Among US Counties, 1980 to 2014: Temporal Trends and Key Drivers." *JAMA Internal Medicine* 177 (7): 1003–11.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114 (39): 10384–89.

Finkelstein, Amy. 2009. "E-ZTAX: Tax Salience and Tax Rates." *The Quarterly Journal of Economics* 124 (3): 969–1010.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *The Quarterly Journal of Economics* 131 (4): 1681–726.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2021. "Place-Based Drivers of Mortality: Evidence from Migration." *American Economic Review* 111 (8): 2697–735.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127 (3): 1057–106.

Fletcher, Jason M. 2010. "Social Interactions and Smoking: Evidence Using Multiple Student Cohorts, Instrumental Variables, and School Fixed Effects." *Health Economics* 19 (4): 466–84.

Frijters, Paul, John P. Haisken-DeNew, and Michael A. Shields. 2005. "The Causal Effect of Income on Health: Evidence from German Reunification." *Journal of Health Economics* 24 (5): 997–1017.

Fuchs, Victor R. 1974. *Who Shall Live? Health, Economics, and Social Choice*. New York: Basic Books.

Galama, Titus J., Adriana Lleras-Muney, and Hans van Kippersluis. 2018. "The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence." NBER Working Paper 24225.

Gibson, John, Steven Stillman, David McKenzie, and Halahingano Rohorua. 2013. "Natural Experiment Evidence on the Effect of Migration on Blood Pressure and Hypertension." *Health Economics* 22 (6): 655–72.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2): 223–55.

Heutel, Garth, Nolan H Miller, and David Molitor. 2020. "Adaptation and the Mortality Effects of Temperature across US Climate Regions." NBER Working Paper 23271.

Hut, Stefan. 2020. "Determinants of Dietary Choice in the US: Evidence from Consumer Migration." *Journal of Health Economics* 72, 102327.

Johnson, Janna E., and Evan J. Taylor. 2019. "The Long-Run Health Consequences of Rural-Urban Migration." *Quantitative Economics* 10 (2): 565–606.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica* 79 (1): 211–51.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders. 2016. "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health." *Review of Economics and Statistics* 98 (2): 350–66.

Lindahl, Mikael. 2005. "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income." *Journal of Human Resources* 40 (1): 144–68.

Macintyre, Sally, Anne Ellaway, and Steven Cummins. 2002. "Place Effects on Health: How Can We Conceptualise, Operationalise and Measure Them?" *Social Science & Medicine* 55 (1): 125–39.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies* 60 (3): 531–42.

Miller, Sarah, Norman Johnson, and Laura R. Wherry. 2019. "Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data." NBER Working Paper 26081.

Mokdad, Ali H., Katherine Ballestros, Michelle Echko, Scott Glenn, Helen E. Olsen, Erin Mullany, Alex Lee, et al. 2018. "The State of US Health, 1990–2016: Burden of Diseases, Injuries, and Risk Factors among US States." *JAMA* 319 (14): 1444–72.

Molitor, David. 2018. "The Evolution of Physician Practice Styles: Evidence from Cardiologist Migration." *American Economic Journal: Economic Policy* 10 (1): 326–56.

Montez, Jennifer Karas, Anna Zajacova, Mark D. Hayward, Steven H. Woolf, Derek Chapman, and Jason Beckfield. 2019. “Educational Disparities in Adult Mortality across US States: How Do They Differ, and Have They Changed Since the Mid-1980s?” *Demography* 56 (2): 621–44.

Montez, Jennifer Karas, Jason Beckfield, Julene Kemp Cooney, Jacob M. Grumbach, Mark D. Hayward, Huseyin Zeyd Koytak, Steven H. Woolf, and Anna Zajacova. 2020. “US State Policies, Politics, and Life Expectancy.” *The Milbank Quarterly* 98 (3): 668–99.

Murphy, Kevin M., and Robert H. Topel. 2006. “The Value of Health and Longevity.” *Journal of Political Economy* 114 (5): 871–904.

Murray, Christopher J.L., Sandeep C. Kulkarni, Catherine Michaud, Niels Tomijima, Maria T. Bulzacchelli, Terrell J. Iandiorio, and Majid Ezzati. 2006. “Eight Americas: Investigating Mortality Disparities across Races, Counties, and Race-Counties in the United States.” *PLoS Medicine* 3(9): e260.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. 2017. “The Gift of Moving: Intergenerational Consequences of a Mobility Shock.” NBER Working Paper 22392.

Oster, Emily. 2019. “Unobservable Selection and Coefficient Stability: Theory and Evidence.” *Journal of Business & Economic Statistics* 37 (2): 187–204.

Roback, Jennifer. 1982. “Wages, Rents, and the Quality of Life.” *Journal of Political Economy* 90 (6): 1257–78.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *The Quarterly Journal of Economics* 116 (2): 681–704.

Schwandt, Hannes. 2018. “Wealth Shocks and Health Outcomes: Evidence from Stock Market Fluctuations.” *American Economic Journal: Applied Economics* 10 (4): 349–77.

Schwandt, Hannes, and Till M. Von Wachter. 2020. “Socioeconomic Decline and Death: Midlife Impacts of Graduating in a Recession.” NBER Working Paper 26638.

Skinner, Jonathan, and Elliott Fisher. 2010. “Reflections on Geographic Variations in US Health Care.” Dartmouth Institute for Health Policy and Clinical Practice Working Paper.

Song, Yunjie, Jonathan Skinner, Julie Bynum, Jason Sutherland, John E. Wennberg, and Elliott S. Fisher. 2010. “Regional Variations in Diagnostic Practices.” *New England Journal of Medicine* 363 (1): 45–53.

Sullivan, Daniel, and Till Von Wachter. 2009. “Job Displacement and Mortality: An Analysis Using Administrative Data.” *The Quarterly Journal of Economics* 124 (3): 1265–306.

Woolf, Steven H, and Heidi Schoomaker. 2019. “Life Expectancy and Mortality Rates in the United States, 1959–2017.” *JAMA* 322 (20): 1996–2016.

Online Appendix

The Causal Effects of Place on Health and Longevity

Tatyana Deryugina, University of Illinois and NBER
David Molitor, University of Illinois and NBER

September 2021

A.1 US Life Expectancy Data

The US life expectancy sample consists of all counties in the US for which life expectancy data is available from the U.S. Small-Area Life Expectancy Estimates Project (USALEEP) (Arias et al., 2018). The USALEEP data report life expectancy at birth for most US census tracts for the period 2010–2015. We aggregate the tract-level measures to the county level using a weighted average, where the weights are the total population for the census tract, as reported in the 2010 Decennial Census (Manson et al., 2020). The highest and lowest 10 counties, ranked by life expectancy, are reported in Appendix Table A.3.

A.2 European Life Expectancy Data

For the European life expectancy sample, we consider all regions covered by the Nomenclature of Territorial Units for Statistics (NUTS) 2021 classification, a hierarchical system that sequentially divides countries into NUTS 1, NUTS 2, and NUTS 3 regions.[1] The NUTS system covers 37 European countries, including those in the European Union (EU) and the United Kingdom, EU candidate countries, and European Free Trade Association countries.[2]

We collect data on life expectancy at birth at both the NUTS 2 and NUTS 3 levels. From the European Statistical Office (Eurostat), we obtain life expectancy at birth in 2018 at the NUTS 2 level for all countries covered by NUTS except Albania. Life expectancy data are not systematically reported at the NUTS 3 level. We compile the data for 1,057 regions in 22 countries from various sources using the most recent period available for each region. We collected total life expectancy at birth, whenever available. In cases where life expectancy was only reported separately for men and women, we defined total life expectancy to be the average of these two measures. Appendix Table A.1 lists the source and time period of NUTS 3 life expectancy data for each of these countries. Appendix Table A.2 summarizes the

[1] Details of the NUTS 2021 classification can be found at `https://ec.europa.eu/eurostat/web/nuts/background` (accessed November 13, 2021).

[2] A description of the statistical regions in the NUTS classification can be found at `https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-092` (accessed November 13, 2021).

availability of life expectancy by country at the NUTS 2 and NUTS 3 levels. The highest and lowest 10 NUTS 3 regions, ranked by life expectancy, are reported in Appendix Table A.5.

A.3 Analysis of Regional Variation

To estimate the share of US county-level variation in life expectancy that occurs within versus across states, we run the following regression.

$$[\textit{life expectancy}]_i = [\textit{state fixed effects}] + \epsilon_i, \tag{A–1}$$

where observations, indexed by i, are at the county level. States include all 50 states and the District of Columbia. The R-squared (R^2) from this regression captures the share of the county-level variance in life expectancy that is explained by the state fixed effects, i.e., the share of the variation that occurs across states. The remainder of the variance, given by $1 - R^2$, describes the share of the county-level variation in life expectancy that occurs within states.

Similarly, we estimate the share of European regional life expectancy that occurs across countries using the R^2 from the regression

$$[\textit{life expectancy}]_i = [\textit{country fixed effects}] + \epsilon_i. \tag{A–2}$$

We estimate this regression separately for observations i defined at the NUTS 2 and NUTS 3 levels. The countries included in the NUTS 2 and NUTS 3 life expectancy samples are reported in Appendix Table A.2.

A.4 Local correlates of US county-level life expectancy

To estimate the correlates of US county-level life expectancy, we run the following regression for each local characteristic considered in the analysis:

$$[\textit{life expectancy}]_i = [\textit{local characteristic}]_i + \epsilon_i, \tag{A–3}$$

where observations, indexed by i, are at the county level and local characteristics are those considered by Deryugina and Molitor (2020), Figure 6, except for characteristics derived from Medicare claims. The characteristics are derived from various sources and are intended to capture a broad range of environmental, economic, and public health conditions. Below, we reproduce the list of characteristics from the Online Appendix of Deryugina and Molitor (2020), organized by data source. We also reproduce the description of how each variable was constructed.

- Census
  - Income per capita
  - Poverty rate, 65+
  - Median home value
  - Urban population share
- Area Resource Files
  - Physicians per capita
  - Hospital beds per capita
- CMS Hospital Compare
  - Hospital quality index
- Behavioral Risk Factor Surveillance System
  - Percent obese
  - Percent smoking
  - Percent exercising
- Chetty and Hendren (2018)
  - Upward income mobility (from p25)
  - Upward income mobility (from p75)
  - Social capital index
  - Crime rate
  - Local government spending per capita
  - Income segregation
- Climate
  - $PM_{2.5}$ concentrations
  - Hot days/year (90°F+)

Census We measure income, poverty, home values, urban population share, and total population for each county using 2000 Decennial Census data, which we obtain from the IPUMS National Historical Geographic Information System (NHGIS) (Manson et al., 2017).[3] The table and dataset names we refer to below are from the NHGIS.

We measure income as per capita income in 1999 (table *NP082A* of dataset *2000_SF3a*). We measure the poverty share among the 65+ population as the number of individuals aged 65 or older with income in 1999 below the poverty level (table *NP087C* of dataset *2000_SF3a*) as a share of the 65+ population for whom poverty status can be determined (table *NP087C* of dataset *2000_SF3a*). We measure median home values as the median value of owner-occupied housing units (table *NH085A* of dataset *2000_SF3a*). Finally, we use the total population of a county (table *NP001A* of dataset *2000_SF1a*) as the denominator for physicians and hospital beds per capita.

Area Resource File (ARF) We obtain the number of physicians and hospital beds for each county in 2004 from the ARF. For the number of physicians, we use variable *F12129-04*, the total number of active MDs (federal and non-federal) in 2004 from the AMA Physician Master File, as provided in the 2005 release of the ARF (U.S. Department of Health and Human Services, 2005). The variable *F08921-04* reports the total number of hospital beds in 2004 from the AHA Survey Database and is provided in the 2009 release of the ARF (U.S. Department of Health and Human Services, 2009).

We calculate the number of physicians per capita by dividing the total number of active MDs by the total population in the county (from census data, described above). Likewise, we calculate hospital beds per capita by dividing the total number of hospital beds by the county population.

Behavioral Risk Factor Surveillance System (BRFSS) We measure obesity, smoking, and exercise behavior using the BRFSS, a telephone survey that collects information on health-related behaviors and chronic conditions (Centers for Disease Control and Prevention, 1995–2004).[4] We pool survey responses for the period 1995–2004.

We calculate percent smoking in each county as the percent of survey respondents for whom the reported smoking status is either "current, daily" or "current, other than daily". We calculate percent obese in each county as the percent of survey respondents who report a body mass index of 30 or greater. We calculate percent exercising in each county as the

[3]Data were downloaded from `https://data2.nhgis.org/main` (accessed October 1, 2019).
[4]Data were downloaded from `https://www.cdc.gov/brfss/annual_data/annual_data.htm` (accessed October 1, 2019).

percent of survey respondents who report participating in any physical activities or exercises other than their regular job in the past month.

CMS Hospital Compare We measure hospital quality within each county using data from the CMS Hospital Compare Process of Care Scores for 2004, which we obtain from Sacarny (2018). We focus on process of care measures for heart attack (AMI), heart failure (HF), and pneumonia (PN), and restrict to metrics that are reported in at least 1,750 counties. This restriction selects a total of 13 metrics, consisting of four AMI metrics (*ami1_share*, *ami2_share*, *ami5_share*, *ami6_share*), three HF metrics (*hf1_share*, *hf2_share*, *hf3_share*), and six PN metrics (*pn1_share*, *pn2_share*, *pn3_share*, *pn4_share*, *pn5_share*, *pn6_share*).

For each process of care metric, we calculate the share of patients in each county who receive appropriate care according to that metric, among hospitals for whom the metric is reported. We combine these 13 process of care metrics into a single hospital quality index, defined as the county-level mean across all metrics (this mean will be missing if any of the underlying metrics are missing for that county). Thus, this hospital quality index can be loosely interpreted as the share of AMI/HF/PN patients receiving appropriate care in the county.

Chetty and Hendren (2018) We obtain county-level measures of upward income mobility, social capital, crime, local government spending, and income segregation from Chetty and Hendren (2018). For measuring upward income mobility, we use the variables *pct_causal_p25_kr26* and *pct_causal_p75_kr26* from Online Data Table 2, "Preferred Estimates of Causal Place Effects by County."[5] The measures of upward income mobility capture the percentage change in income at age 26 from spending one more year of childhood in the county, for children whose parents were at the 25th or 75th percentiles, respectively, of the US household income distribution.

The measures of social capital, crime, local government spending, and income segregation come from Online Data Table 4, "Complete County-Level Dataset: Causal Effects and Covariates." Specifically, we use the variables *scap_ski90pcm*, *crime_total*, *subcty_total_expend_pc*, and *cs00_seg_inc*.[6]

[5] A description of the variables in Online Data Table 2 can be found at `https://opportunityinsights.org/wp-content/uploads/2018/04/online_table2-2.pdf` (accessed October 1, 2019).

[6] A description of the variables in Online Data Table 4 can be found at `https://opportunityinsights.org/wp-content/uploads/2018/04/online_table4-2.pdf` (accessed October 1, 2019).

Climate We measure fine particulate ($PM_{2.5}$) air pollution concentrations and the number of extremely hot days using data recorded by ground monitor stations. We measure the average $PM_{2.5}$ concentration in a county for the period 2006–2013. We obtain $PM_{2.5}$ air pollution data from EPA's Air Quality System database, which provides hourly data at the pollution-monitor level for pollutants that are regulated by the Clean Air Act (U.S. Environmental Protection Agency, 2006–2013). We aggregate monitor readings to the daily level by averaging across hourly observations and then construct daily ZIP code level pollution measures by calculating the inverse distance-weighted average across all monitors located within 20 miles of the ZIP code centroid. We then average these daily values over the period 2006–2013. Finally, we aggregate ZIP code level average pollution concentrations to the county level by averaging across all ZIP codes matched to a county based on the county recorded for the plurality of Medicare beneficiaries living in that ZIP code.

Our source for daily temperature variables is the Global Historical Climatology Network GHCN-Daily database, which provides weather measurements from land surface stations across the United States (National Oceanic and Atmospheric Administration, 2006–2013). For the period 2006–2013, we calculate daily high and low temperatures for each ZIP code as the inverse distance-weighted average of all available daily maximum and minimum temperatures, respectively, for GHCN stations within a 20-mile radius of the ZIP code centroid. The daily average temperature for a ZIP code is calculated as the midpoint of the daily high and low temperatures. We calculate the number of days per year in which the average daily temperature exceeded 90°F in a ZIP code, and then aggregate to the county level using the same ZIP code to county crosswalk used to construct the pollution measure.

Table A.1: Sources of life expectancy at NUTS 3 level by country

Country	NUTS 3 Data Source	NUTS 3 Data Report Year
Austria	Statistics Austria	2019
Bulgaria	Republic of Bulgaria National Statistical Institute	2019
Cyprus	Statistical Service of Cyprus	2019
Denmark	Statistics Denmark	2019
Finland	Statistics Finland	2019
France	French National Institute of Statistics and Economic Studies	2019
Germany	Rau and Schmertmann (2020)	2015–2017
Hungary	Hungarian Central Statistical Office	2019
Italy	Italian National Institute of Statistics	2019
Latvia	Central Statistical Bureau of Latvia	2019
Liechtenstein	World Bank Open Data	2018
Lithuania	Lietuvos statistika	2019
Luxembourg	World Bank Open Data	2018
Montenegro	World Bank Open Data	2018
Norway	Statistics Norway	2011–2015
Poland	Statistics Poland	2019
Portugal	Statistics Portugal	2017–2019
Romania	Romania National Institute of Statistics	2019
Slovenia	Statistical Office of the Republic of Slovenia	2019
Spain	Spain National Statistics Institute	2018
Sweden	Statistics Sweden	2015–2019
Turkey	Turkish Statistical Institute	2017

Notes: The table reports the source and report year, by country, for the NUTS 3 life expectancy sample.

Table A.2: Availability of life expectancy data by European country

	(1)	(2)	(3)	(4)
Country	# NUTS 2	# NUTS 2 with life expectancy data	# NUTS 3	# NUTS 3 with life expectancy data
Albania	3	0	12	0
Austria	9	9	35	35
Belgium	11	11	44	0
Bulgaria	6	6	28	28
Croatia	4	4	21	0
Cyprus	1	1	1	1
Czech Republic	8	8	14	0
Denmark	5	5	11	11
Estonia	1	1	5	0
Finland	5	5	19	19
France	27	24	101	101
Germany	38	38	401	397
Greece	13	13	52	0
Hungary	8	8	20	20
Iceland	1	1	2	0
Ireland	3	3	8	0
Italy	21	21	107	107
Latvia	1	1	6	6
Liechtenstein	1	1	1	1
Lithuania	2	2	10	10
Luxembourg	1	1	1	1
Malta	1	1	2	0
Montenegro	1	1	1	1
Netherlands	12	12	40	0
North Macedonia	1	1	8	0
Norway	7	6	13	11
Poland	17	17	73	73
Portugal	7	7	25	25
Romania	8	8	42	42
Serbia	4	4	25	0
Slovakia	4	4	8	0
Slovenia	2	2	12	12
Spain	19	19	59	54
Sweden	8	8	21	21
Switzerland	7	7	26	0
Turkey	26	25	81	81
United Kingdom	41	41	179	0
Total	334	326	1,514	1,057

Notes: Columns (1) and (3) report the number of NUTS 2 and NUTS 3 regions, respectively, in the NUTS 2021 classification system. Columns (2) and (4) report the number of NUTS 2 and NUTS 3 regions, respectively, for which we have life expectancy data. Maps of life expectancy by NUTS 2 and NUTS 3 region are shown in Online Appendix Figures A.1 and A.2, respectively.

Table A.3: US counties with the highest and lowest life expectancy

	(1)	(2)	(3)	(4)
County name	County FIPS code	State	Life expectancy at birth	Rank
Regions with the highest life expectancy				
Cheyenne County	08017	CO	89.5	1
Wayne County	49055	UT	89.3	2
Haskell County	20081	KS	88.6	3
Stanton County	20187	KS	87.9	4
Custer County	16037	ID	87.7	5
Sherman County	48421	TX	87.3	6
Crook County	56011	WY	87.1	7
Granite County	30039	MT	87.0	8
Aleutians East Borough	02013	AK	86.9	9
Concho County	48095	TX	86.8	10
Regions with the lowest life expectancy				
Walker County	01127	AL	71.4	3,099
Madison Parish	22065	LA	71.4	3,100
Emporia city	51595	VA	71.4	3,101
Estill County	21065	KY	71.3	3,102
Sussex County	51183	VA	71.0	3,103
Tallahatchie County	28135	MS	70.8	3,104
Powell County	21197	KY	70.8	3,105
Breathitt County	21025	KY	70.2	3,106
McIntosh County	40091	OK	69.7	3,107
East Carroll Parish	22035	LA	69.1	3,108

Notes: The table reports the top 10 and bottom 10 counties in the US county life expectancy sample ($N = 3,108$), ranked by life expectancy at birth.

Table A.4: European NUTS 3 regions with the highest and lowest life expectancy

	(1)	(2)	(3)	(4)
NUTS 3 region name	NUTS 3 code	Country	Life expectancy at birth	Rank
Regions with the highest life expectancy				
Madrid	ES300	Spain	84.8	1
Salamanca	ES415	Spain	84.7	2
Soria	ES417	Spain	84.6	3
Hauts-de-Seine	FR105	France	84.5	4
Prato	ITI15	Italy	84.5	5
Perugia	ITI21	Italy	84.5	6
Paris	FR101	France	84.5	7
Pordenone	ITH41	Italy	84.4	8
Firenze	ITI14	Italy	84.4	9
Araba/Álava	ES211	Spain	84.3	10
Regions with the lowest life expectancy				
Разград	BG324	Bulgaria	73.1	1,048
Добрич	BG332	Bulgaria	73.1	1,049
Vidzeme	LV008	Latvia	73.0	1,050
Монтана	BG312	Bulgaria	73.0	1,051
Видин	BG311	Bulgaria	73.0	1,052
Сливен	BG342	Bulgaria	72.8	1,053
Враца	BG313	Bulgaria	72.8	1,054
Kurzeme	LV003	Latvia	72.7	1,055
Zemgale	LV009	Latvia	72.3	1,056
Latgale	LV005	Latvia	70.4	1,057

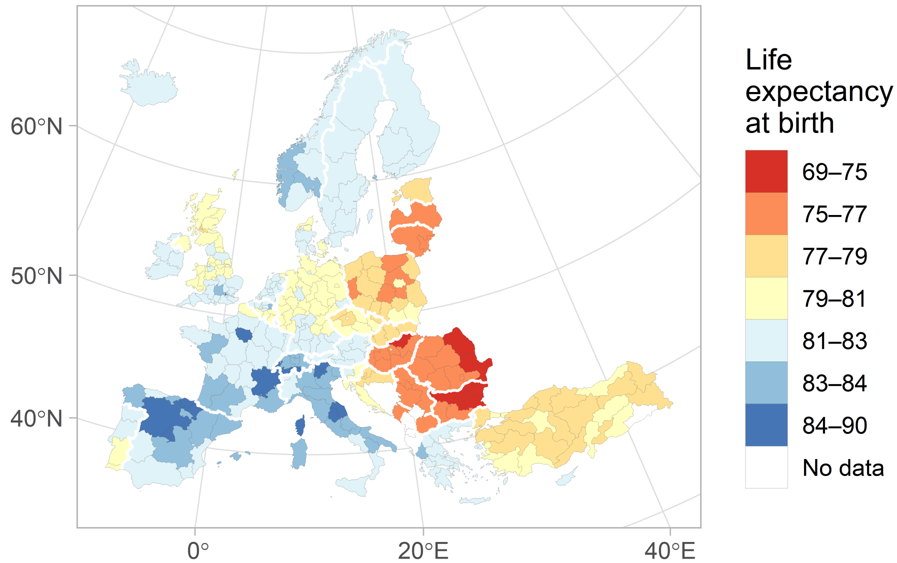
Notes: The table reports the top 10 and bottom 10 regions in the European NUTS 3 life expectancy sample ($N = 1,057$), ranked by life expectancy at birth.

Table A.5: Local correlates of US county-level life expectancy

	(1)	(2)	(3)
County characteristic	Characteristic mean [sd]	OLS coeficient (se)	R-squared
A. Health and environmental characteristics			
Percent smoking	22.83	-0.23	0.258
	[5.20]	(0.01)	
Percent obese	22.80	-0.19	0.228
	[5.92]	(0.02)	
Percent exercising	72.95	0.23	0.444
	[6.84]	(0.01)	
Physicians per 1,000 capita	1.33	0.16	0.011
	[1.69]	(0.03)	
PM 2.5 concentrations	10.08	-0.47	0.153
	[2.01]	(0.03)	
Hospital beds per 1,000 capita	3.77	0.02	0.001
	[5.39]	(0.01)	
Hot days/year (90°F+)	0.90	-0.03	0.002
	[3.78]	(0.01)	
Hospital quality index	0.76	5.07	0.033
	[0.08]	(0.69)	
B. Economic characteristics			
Median home value ($1,000s)	81.64	0.02	0.154
	[42.26]	(0.00)	
Income per capita ($1,000s)	17.53	0.27	0.173
	[3.94]	(0.01)	
Poverty rate, 65+	0.12	-21.18	0.244
	[0.06]	(0.80)	
Upward income mobility (from p25)	0.23	2.27	0.236
	[0.53]	(0.07)	
Urban population share	0.41	0.12	0.000
	[0.31]	(0.16)	
Crime rate per 1,000	5.78	-0.15	0.046
	[3.83]	(0.01)	
Local gov. spending per capita	2.15	0.39	0.049
	[1.45]	(0.18)	
Upward income mobility (from p75)	0.14	0.95	0.007
	[0.21]	(0.23)	
Social capital index	-0.01	0.84	0.192
	[1.34]	(0.03)	
Income segregation	0.02	0.15	0.000
	[0.03]	(1.48)	
C. Multivariate comparisons			
All health and environmental characteristics			0.649
All economic characteristics			0.539
All characteristics			0.774

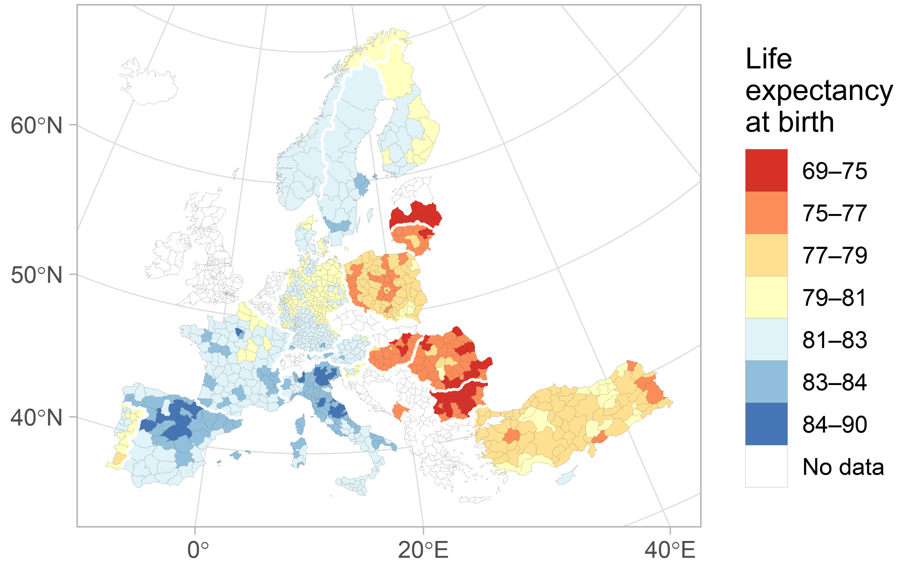
Notes: The table reports results from regressing US county-level life expectancy on local characteristics. Each row in the table corresponds to a separate regression, where the included local characteristic(s) are indicated by the row labels. Observations are unweighted. Column (1) shows the mean and standard deviation (in brackets) of the local characteristic. Column (2) reports regression coefficients and robust standard errors (in parentheses). Column (3) reports the R-squared from the regression. Online Appendix Section A.4 provides more details on the data and regressions.

Figure A.1: European Life Expectancy by NUTS 2 Region



Notes: The figure shows life expectancy at birth in Europe at the NUTS 2 level, based on the NUTS 2 life expectancy sample described in Online Appendix Table A.2.

Figure A.2: European Life Expectancy by NUTS 3 Region



Notes: The figure shows life expectancy at birth in Europe at the NUTS 3 level, based on the NUTS 3 life expectancy sample described in Online Appendix Table A.2.

References

Arias, Elizabeth, Loraine A Escobedo, Jocelyn Kennedy, Chunxia Fu, and Jodi A Cisewski. 2018. "US small-area life expectancy estimates project: Methodology and results summary." https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NVSS/USALEEP (accessed January 11, 2021).

Centers for Disease Control and Prevention. 1995–2004. "Behavioral Risk Factor Surveillance System." https://www.cdc.gov/brfss/annual_data/annual_data.htm (accessed October 1, 2019).

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-level estimates: Dataset." *The Quarterly Journal of Economics*, 133(3): 1163–1228. https://opportunityinsights.org/paper/neighborhoodsii (accessed July 13, 2020).

Deryugina, Tatyana, and David Molitor. 2020. "Does when you die depend on where you live? Evidence from Hurricane Katrina." *American Economic Review*, 110(11): 3602–3633.

Manson, Steven, Jonathan Schroeder, David Van Riper, Steven Ruggles, et al. 2017. "IPUMS National Historical Geographic Information System: Version 12.0 [Database]." *Minneapolis: University of Minnesota*, 39.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2020. "IPUMS National Historical Geographic Information System: Version 15.0 [dataset]." *Minneapolis, MN: IPUMS.*

National Oceanic and Atmospheric Administration. 2006–2013. "Global Historical Climatology Network." ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/ (accessed October 1, 2019).

Rau, Roland, and Carl P Schmertmann. 2020. "District-level life expectancy in Germany." *Deutsches Ärzteblatt International*, 117(29-30): 493. http://german-district-mortality.schmert.net/ (accessed September 13, 2021).

Sacarny, Adam. 2018. "CMS Hospital Compare Data 2004–2016." https://github.com/asacarny/hospital-compare (accessed October 1, 2019).

U.S. Department of Health and Human Services, Health Resources and Services Administration. 2005. "Area Resource File, 2005." National Bureau of Economic Research [data host]. http://data.nber.org/homedata (accessed July 13, 2020).

U.S. Department of Health and Human Services, Health Resources and Services Administration. 2009. "Area Resource File, 2009." National Bureau of Economic Research [data host]. http://data.nber.org/homedata (accessed July 13, 2020).

U.S. Environmental Protection Agency. 2006–2013. "Air Quality System." https://aqs.epa.gov/aqsweb/airdata/download_files.html (accessed October 1, 2019).