Human Capital Spillovers and Health:

Does Living Around College Graduates Lengthen Life?*

Jacob Bor[†], David Cutler[‡], Edward Glaeser[‡], Ljubica Ristovska[§]

August 2024

[CLICK HERE FOR THE MOST RECENT VERSION]

Abstract

Equally educated people are healthier if they live in more educated places. Every 10 percentage point increase in an area's share of adults with a college degree is associated with a decline in all-cause mortality of 7%, controlling for individual education, demographics, and area characteristics. Area human capital is also associated with lower disease prevalence and improved self-reported health. The association between area education and health increased greatly between 1990 and 2010. Increased spatial sorting does not drive these spillovers; there is little evidence that sicker people move disproportionately into less educated areas. Differences in health-related amenities, ranging from hospital quality to pollution, explain no more than 17% of the area human capital spillovers on health. Over half of the correlation between area human capital and health is a result of the correlation between area human capital and smoking and obesity. More educated areas have stricter regulations regarding smoking and more negative beliefs about smoking. These translate over time into a population that smokes noticeably less and that is less obese, leading to increasing divergence in health outcomes by area education.

[†] Boston University School of Public Health, jbor@bu.edu
[‡] Harvard University and NBER, Cutler: dcutler@harvard.edu, Glaeser: eglaeser@harvard.edu
[§] Yale University, ljubica.ristovska@yale.edu
* Cutler thanks the National Institute on Aging, and Glaeser thanks the Taubman Center for State and Local Government for financial support.

Introduction

The health of a region is closely tied to its education level. Ezzati et al. (2008) show that county life expectancy rises by 1.3 years as the share of adults with a college degree increases by 10 percentage points. Much of this county-level correlation between education and health reflects the well-known individual-level relationship between years of schooling and mortality (Elo and Preston, 1996; Cutler and Lleras-Muney, 2006; Grossman, 2006; Grossman, 2008; Meara, Richards, and Cutler, 2008; Cutler, Lange, Meara, Richards-Shubik, and Ruhm, 2011; Cutler and Lleras-Muney, 2012; Grossman, 2015), but that may not be all of it. This paper asks whether human capital spillovers in health, akin to human capital spillovers in earnings (Rauch, 1993; Moretti, 2004a; Moretti, 2004b; Canton, 2007; Rosenthal and Strange, 2008; Iranzo and Peri, 2009), help explain the relationship between area education and mortality, and why that might be.

The link between area education and health was particularly apparent during COVID. The number of COVID-related deaths declined by 35% for each 10 percentage point increase in college graduates in an area, ten times larger than the difference expected due to compositional effects alone.[1] The COVID-19 pandemic was but one example of a larger trend. Large and growing geographic disparities in health across the US are central aspects of American life (Murray et al., 2005; Murray et al., 2006; Krieger et al., 2008; Ezzati et al., 2008; Kulkarni et al., 2011; Chetty et al., 2016; Dwyer-Lindgren, 2017; Finkelstein, Gentzkow, and Williams, 2021). The life expectancy gap between counties in the 1st vs. 99th percentile increased from 8.3 years in 1980 to 10.7 years in 2014 (Dwyer-Lindgren, 2017). Experimental and quasi-experimental methods have established that place of residence causally impacts both physical and mental health, although there

[1] Case and Deaton (2023) report a COVID mortality rate of 57 per 100,000 for those with a college degree and 164 per 100,000 for those without, for an average of 132 per 100,000. Based on this, a 10 percentage point increase in college graduates would be expected to reduce deaths by 28.8 per 100,000, or roughly 22%.

is little consensus about why place is so powerful (Katz, Kling, and Liebman, 2001; Kling, Liebman, and Katz, 2007; Doyle, 2011; Ludwig et al., 2011; Ludwig et al., 2012; Ludwig et al., 2013; Deryugina and Molitor, 2020; Finkelstein, Gentzkow, and Williams, 2021). The increasing variation in education levels over the past four decades (Berry and Glaeser, 2005; Moretti, 2013; Diamond, 2016) combined with human capital spillovers in health may help us understand widening geographic health disparities.

To estimate the association between area education and health, we combine U.S. Census and American Community Survey data for 1990, 2000, and 2010 with complete mortality records containing cause of death information and individual education from the Multiple Cause Mortality Files. [2] Controlling for individual-level educational attainment, a 10 percentage point increase in the percentage of college graduates in an area – a move from roughly the 25th percentile to the 75th percentile – is associated with a 7% lower all-cause mortality rate. It is present across all education and demographic groups and is strongest for Hispanics and people under 65 and for people residing in rural and less educated areas. Area human capital is also strongly correlated with non-fatal health outcomes and self-reported health.

Particularly importantly, the magnitude of the relationship between area education and longevity has increased over time.[3] The relationship between area education and mortality was 56% greater in 2010 than in 1990. Examining data by cause of death shows that while area human capital spillovers exist across almost all causes, they are increasing over time only for medically amenable causes of death (which includes respiratory conditions and heart disease deaths), deaths due to cancer, chronic respiratory disease, external causes, and drug overdoses.

[2] The U.S. Standard Certificate of Death only included information on the decedent's education after 1989.
[3] Human capital earnings spillovers also appear to be increasing over time (Glaeser et al., 2004).

After documenting these facts, we examine three potential reasons for the correlation between area-level education and health: i) spatial sorting, where healthier individuals move to high human capital areas; ii) area-level amenities that influence health and that are correlated with higher education, for example a better and safer environment and more medical care; and (iii) individuals in better-educated areas choosing fewer health-harming behaviors.

We reject the spatial sorting hypothesis using data from the Health and Retirement Study for individuals 51 years of age and older and the National Longitudinal Survey of Young Women and Men for younger individuals. Less healthy people, as measured by predicted mortality, move to areas with approximately equal levels of human capital as healthier people.

In contrast, we find that differences across areas in health-related behaviors such as smoking and obesity explain approximately 60% of the correlation between area human capital and mortality, even after controlling for individual education. This result is consistent with our finding using data from the Behavioral Risk Factor Surveillance System and the Current Population Survey that, after controlling for individual-level education, a 10 percentage point increase in area human capital is associated with a 13% decrease in the probability of smoking, a 7% decrease in the probability of having no physical activity, and a 12% decrease in the probability of being very obese. These changes also show up in both smoking initiation and cessation. Young adults in more educated areas are less likely to begin smoking than equally educated young adults in low human capital areas. Further, conditional on being a smoker, individuals in better educated areas are much more likely to quit smoking in their 30s and 40s than similarly educated individuals in low human capital areas. These findings are true cross-sectionally and in panel data, controlling for area and time fixed effects. Accumulated over several decades, lower smoking initiation and higher quit rates in more educated areas have led to a population that is increasingly characterized by never smokers and former smokers compared to current smokers, resulting in growing gaps in mortality across areas and the greater correlation between area education and longevity over time.

Other health-related amenities such as pollution, crime, access to medical care, and quality of medical care do not explain much of the gradient in health with area education. We estimate that at most 17% of the human capital spillover on health is due to these external factors, driven largely by greater use of preventative care. As we cannot observe or even imagine the full range of omitted variables that could potentially impact health, we cannot reject the view that omitted variables are more important than what we estimate, but we have no evidence of any observed variable significantly explaining health disparities across the US.

In the final part of the paper, we consider two potential channels linking area human capital spillovers and smoking: smoking regulations and social norms. We find that individuals in areas with a 10 percentage point higher share of college graduates are 1.8 percentage points more likely to be employed at workplaces with a smoking ban in all work and public areas, even after controlling for individual-level education. Thus, smoking regulation could explain some of the spillover effect we find.

We examine differences in social norms through questions on peoples' beliefs about smoking's harms and attitudes towards smoking regulation. Analyzing data from the National Health Interview Survey (1987, 1992, and 2000), we find that controlling for individual education, a 10 percentage point increase in the percent of college graduates is associated with an 11% increase in the probability of agreeing with the statement that smoking is harmful to fetal development and a 15% increase in the probability of agreeing that most lung cancer deaths stem from smoking. It is also associated with an 8% increase in the probability of supporting smoking bans in bars, restaurants, and work areas. Directly controlling for smoking regulations and beliefs about smoking in regressions of smoking on area human capital and individual education suggests that regulations and beliefs can explain up about 15% (17%) of the correlation between area human capital and smoking rates (quitting rates) and almost *a quarter* of the correlation between area human capital and smoking initiation.

Overall, we conclude that there is a nexus between smoking beliefs, smoking behavior, and area mortality that helps understand the growing correlation between life expectancy and education across areas.

Data on Mortality and Area Characteristics

Mortality

We obtained microdata on all deaths of U.S. residents in 1990, 2000, and 2010 from the National Center for Health Statistics Multiple Cause Mortality Files (MCMF) restricted access files, which include county of residence.[4] MCMF data are compiled from death certificates and include underlying cause of death as well as age, sex, and educational attainment of the deceased (since 1988). Educational attainment on death certificates is typically reported by next-of-kin.[5]

We aggregate total deaths into county-age-sex-race-education cells. We excluded the 3% of deaths that occurred among individuals younger than 25, as education is not reliably completed before that age. Cells were defined by 5-year age categories (25-29, 30-34, …, 85+), five levels of educational attainment based on completed years of school (<12, 12, 13-15, 16, 17+), gender (M,

[4] All data sources used in our paper are summarized in Appendix A.
[5] Some concerns have been raised regarding the accuracy of such reporting relative to self-reports, particularly the overstatement of high school graduation rates (Shai and Rosenwaike, 1989; Sorlie and Johnson, 1996; Rosamond et al., 1997; Rostron et al., 2010).

F), and race/ethnicity (White non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic).[6,7] We excluded deceased individuals with missing data on age (0.02%), county (0.16%), and education (10.5%) since we cannot match these deaths to a population denominator when calculating mortality rates. To ensure comparability of geographic units across years and for matching with population denominators, we aggregated counties into consistent public use microdata areas (CONSPUMA), representing the most detailed geographic areas that can be consistently identified between 1980 and 2011.[8] There are 486 CONSPUMAs in the US.

We excluded any area-year-age-sex-race cells where the percent of deaths with missing education was more than 25%, which eliminated 1.9% of adult deaths in our sample. Our regression analysis controls for the percent of death certificates without education in each area-age-sex-race cell.

Mortality rates were calculated by merging death counts for area-age-sex-race-education cells with corresponding population counts from the 1990 and 2000 U.S. Decennial Census (5% sample) and the pooled 2009-2011 American Community Survey (ACS) for 2010 (as in Wheeler, 2007).[9]

Appendix Table C1 shows summary statistics for the mortality data. Our final dataset, pooled across 1990, 2000, and 2010, contains 369,707 area-year-age-sex-race-education cells and

---

[6] Due to differences in data encoding over time in the MCMF data, in 1990 and 2000, we considered individuals with four years of high school as having completed high school, regardless of whether they were awarded a diploma. In 2010, we considered those with 12 years of education and no high school diploma as not having completed high school. Associate degrees were included in the 13-15 (some college) education category.

[7] Rostron et al. (2010) show that college completion rates (schooling equal to 16 years) may be underreported in favor of 'some college' (schooling equal to 13-15 years) in death certificates, which may lead us to underestimate the mortality rate among college graduates. As a result, we check robustness of our main results to measuring area human capital as the percent of individuals with more than high school education in an area.

[8] Appendix Figure C1 shows a map of the CONSPUMAs across the U.S. For counties included in multiple CONSPUMAs, we use the CONSPUMA containing most of the county's population. Only 36 out of 3141 counties in 1990, 42 in 2000, and 44 in 2010 (out of 3,219) map to multiple CONSPUMAs.

[9] We use the 3-year ACS because the 2010 Decennial Census did not include data on individual education.

covers 5,934,489 deaths across all years, which represents 84% percent of deaths for people aged 25 and older.[10] The overall death rate was 1,196 deaths per 100,000, roughly 1 percent per year. Cause of death is also identified on death certificates. We classified causes of death as medically amenable, smoking-related, obesity-related, or due to external causes based on the literature (see Appendix B for details).[11] The same cause of death may be in multiple categories (e.g., heart disease is both smoking and obesity related).

Data on Non-Fatal Health Outcomes

We obtained individual-level data on health conditions such as cancer, lung disease, diabetes, heart disease, and stroke from the Health and Retirement Study (HRS), a biennial, longitudinal survey of people aged 51 and older over the 1992-2008 period. We also used microdata on self-reported general health and number of days over the last 30 days where poor physical or mental health interfered with daily activities from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS).[12] Both data sources contain individual education and demographics. Each respondent in the HRS and BRFSS data was mapped to a CONSPUMA.[13] As with the mortality data, we restrict the BRFSS sample to individuals aged 25 years and older.

Data on Migration

We use the HRS along with the National Longitudinal Survey of Youth (NLSY) to assess migration by health status. The HRS sample is the same as above and includes data on cross-

[10] The largest drop in sample size is due to the exclusion of decedents with missing education.

[11] A cause of death can fall into multiple categories; for example, heart disease is both smoking-related and obesity-related. Approximately 56% of deaths were due to causes classified as smoking-related, 41% as obesity-related, 41% as medically amenable, and 6% due to external causes.

[12] The 1990 BRFSS does not contain data on self-reported health or number of days where poor health interfered with daily activities. It also does not have consistent geographic identifiers.

[13] We utilize a restricted-use HRS file with county identifiers. County identifiers are included in the BRFSS but are suppressed for areas with fewer than 50 respondents.

county migration between survey waves. The NLSY sample was aged 26-38 in 1990 and 46-58 in 2010. Young men were asked in 1969-1971 and 1976 whether they had moved to a different standard metropolitan statistical area (SMSA) or county since the last interview, and young women were asked annually or every two years between 1968-2001 whether they had moved to a different SMSA or county since the last interview.

Data on Health-Related Behaviors

We used self-reported data on smoking status, body mass index (BMI), and physical activity from the 1999-2001 and 2009-2011 BRFSS data. Since the BRFSS does not contain these measures in the 1990s, we supplement the BRFSS data with data on individual education, demographics, and smoking status from the Tobacco Use Supplement in the Current Population Survey (CPS) from 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015.[14] These data contain identifiers for counties with a population of 100,000 or greater. As with other data sources, we include only individuals 25 or older and match available counties to CONSPUMAs.

Area Characteristics

We merged in several CONSPUMA-level attributes to our micro data, as summarized in Appendix A. Area human capital was defined as the percent of area residents aged 25 or older with at least a college degree, using Census data from 1990 and 2000 and ACS data from 2009-2011. We also use these sources to calculate area-level percent Black and Hispanic, and industry shares of employment.

Area population size and land area come from the Area Resource Files provided by the Bureau of Health Workforce for 1990, 2000, and 2010, which we use to compute population

[14] The CPS data does not ask about height and weight and thus we cannot calculate BMI in this data set.

density. We also obtained counts of non-federal physicians and hospital beds at the county level for 1990, 2000, and 2010 from the Area Resource Files. From the Dartmouth Health Atlas, we obtained county-level data on the average annual percent of Medicare enrollees having at least one annual ambulatory visit to a primary care clinician and the average percent of female Medicare enrollees aged 67-69 having at least one mammogram over a two-year period for years 2003-2015.

County-level reported homicides come from the Uniform Crime Reports. For each of 1990, 2000, and 2010, we averaged homicides in the three years centered around the decade (e.g., 1989-91 for 1990) to improve precision.

Satellite data on air pollution for 1999-2001 and 2009-2011 are from van Donkelaar (2019) and capture the concentration of suspended particulate matter of diameter 2.5µm or less (PM-2.5). For 1989-1991, we obtain data on PM-10 measurements from the Environmental Protection Agency for counties with particulate matter monitoring agencies. We follow the methodology from Meng et al. (2019) to generate predicted PM-2.5 measurements for 1989-1991 using the PM-10 and PM-2.5 data.

Hospital quality comes from the Hospital Compare Database provided by the Centers for Medicare and Medicaid Services for 2003-2008, which contains measures reflecting the usage of inexpensive, easy-to-implement practices representing the standard of care.[15] For each hospital, measures were first converted into z-scores for each condition. We then average these z-scores across conditions, which provides a single hospital-specific metric for 3,879 hospitals, which we treat as roughly representing hospital quality for 2010. Finally, we calculate area-level hospital quality scores, weighting the hospital quality of all hospitals in the area by the number of discharges per hospital.

[15] For example, one measure is the percent of patients presenting with an acute myocardial infarction who are given aspirin upon arrival.

Area Human Capital and Mortality

Figure 1 shows ventiles of the relationship between area human capital and mortality across area-age-sex-race-education cells for each of 1990, 2000, and 2010, adjusting for age and sex but not for individual education.[16] The figure shows a negative relationship between education and mortality which is increasing over time. In 1990, a 10 percentage point increase in the area-level share of adults with a college degree – i.e., moving from the bottom quartile to the top quartile of the 2010 distribution of area human capital – was associated with a decline of 44 deaths per 100,000 ($p < 0.01$), a 3.8% reduction in all-cause mortality. By 2010, the change was 97 deaths per 100,000, representing an 8.2% reduction in all-cause mortality.

Area education is correlated with individual education, and individual education is related to health.[17] A central question is whether these aggregate findings simply reflect the well-known relationship between individual education and health. Table 1 shows a variety of analyses separating individual and area-level education. Each column reports results of a regression model relating cell-level mortality rates to cell and area characteristics, using data for all area-year-age-sex-race-education cells. We limit controls to demographic and geographic characteristics that are unlikely to be part of the causal pathway between area human capital and health: in addition to the controls reported in Table 1, we include controls for 5-year age-sex-race/ethnicity interactions, as well as year. We also control for the percent of death certificates in the cell with missing education data, population and population density (both log-transformed), and employment shares by

[16] Since we do not control for individual education, data in this figure includes deaths with missing education information. Excluding deaths with missing education yields similar results (see Appendix Figure C2).

[17] The literature on the relationship between individual education and health is vast. For a comprehensive review of the theoretical background, as well as descriptive and quasi-experimental evidence on the relationship between education and health, see, for instance, Grossman (2006), Cutler and Lleras-Muney (2006), Grossman (2008), Cutler and Lleras-Muney (2012), Grossman (2015), and Galama et al. (2017).

industry at the area level.

Column 1 of the table examines the effect of individual education alone. Controlling for other cell-level and area-level covariates, the correlation between individual education and mortality is enormous. Individuals without a high school degree experience 730 additional deaths per 100,000 (50%) relative to individuals with graduate education. Mortality risk declines with each additional level of educational attainment. The second column shows the relationship between mortality and area human capital without individual education controls. These results are closely related to Figure 1 and show that a 10 percentage point increase in the area-level percent of the adult population with a college degree is associated with 103 fewer deaths per 100,000.

The third column presents the primary motivating fact for the paper. Even controlling for individual education, a 10 percentage point increase in the share of college graduates in an area is associated with 64 fewer deaths per 100,000 ($p < 0.01$), a 5.4% decrease relative to average mortality. The difference in the coefficients on the share of college graduates between the second and third columns of Table 1 implies that controlling for individual education explains just 38% of the relationship between area human capital and mortality shown in Figure 1.

Column 4 of Table 1 allows for the relationship between area human capital and mortality to vary by year. As with Figure 1, the relationship between area human capital and mortality increases over time. Previous studies have found widening mortality disparities across individuals with different levels of education over time (Meara, Richards, and Cutler, 2008; Cutler, Lange, Meara, Richards-Shubik, & Ruhm, 2011; Olshansky et al., 2012; Masters et al., 2012; Hayward et al., 2015; Sasson 2016; Bor et al., 2017, Case and Deaton, 2023). Our paper demonstrates that there exists a similarly increasing impact of area human capital on mortality over time.

Columns 5 and 6 match the specifications in columns 3 and 4 but include state-by-year fixed effects, which account for time-varying state-level characteristics that may be correlated with both area human capital and health (e.g., changing state-level health or education policies such as Medicaid coverage, tobacco taxes, smoking regulations, etc.). The impact of area human capital on mortality falls in these specifications but remains statistically significant and increasing over time.[18,19,20,21] Thus, differences in state-level policies cannot be the sole factor driving the correlation between area human capital and health nor the increase in this effect over time; rather, local relationships exist.[22]

Finally, columns 7 and 8 include area fixed effects, which control for time-invariant area-level characteristics that may be correlated with area human capital and health. Within areas, there is a similar correlation between human capital and health as in our baseline specifications in columns 5 and 6.[23]

Appendix Figure C3 presents estimates from our baseline regression from column 5 in

---

[18] This is consistent with Karas Montez et al. (2022) findings on the relationship between state policies and the mortality of working-age adults.

[19] These results are robust, but smaller in magnitude if area human capital is measured as percent of adults with more than a high school degree. Under this definition of area human capital, a 10 percentage point increase in the area human capital is statistically significantly associated with 69.6 fewer deaths per 100,000. This suggests that our results are not driven by measurement error in education on the death certificates.

[20] Education is multidimensional, and measures of area education levels such as percent college educated may omit dimensions of education such as quality, resources, assistance, and intensity. To our knowledge, there are no comprehensive education quality measures across areas for 1990-2000. For 2010, we use the publicly available College Scorecard data to control for the most common measure of value added of higher education in our regressions – earnings of college graduates. We find similar results if we control for mean or median earnings of entry cohorts of universities located in a CONSPUMA 6, 8, or 10 years after graduating.

[21] We obtain similarly increasing spillovers of area education over time if we allow the coefficient of all cell-level characteristics, including individual education, to vary over time.

[22] We obtain similar results when including a full set of individual-level age-sex-race-education interactions – a 10 percentage point increase in the percent college graduates in an area is associated with a decrease of 81.8 deaths per 100,000. Even if we control for the changing relationship between individual education and mortality over time by including fixed effects for year interacted with individual education in a separate specification, we find that a 10 percentage point increase in the percent college graduates in an area is associated with a decrease of 81.1 deaths per 100,000.

[23] Our within-area results are consistent with our findings that (a) if we control for the lag of area human capital, contemporaneous area human capital is strongly negatively correlated with all-cause mortality, and (b) changes in area human capital are strongly negatively correlated with all-cause mortality after controlling for initial area human capital levels. Altogether, these results suggest that it is the change in area human capital levels between years driving the correlation between area human capital and mortality rather than the initial levels of area human capital.

Table 1 but with mortality rates separated by cause of death (Appendix Table C2 has full regression results). Area human capital is negatively correlated with mortality rates across all cause-of-death groupings we analyzed. Appendix Table C3 shows that this correlation strengthens over time for deaths due to medically amenable causes, chronic respiratory disease, cancer, and external causes.

Heterogeneity in the Relationship Between Education and Mortality

Appendix Figure C4 shows estimates of area human capital using the same regression as in column 5 of Table 1, estimated separately by age, gender, individual education, and race.[24] In absolute terms, the relationship between area human capital and mortality is larger for older than younger individuals. However, in relative terms, the relationship between area human capital and health is stronger for younger individuals. The relationship between area human capital and mortality is slightly stronger for men and more educated individuals relative to the respective means. Coefficient estimates are similar for white, Black, and Hispanic individuals in absolute terms, but the relationship is strongest for Hispanic individuals relative to the average mortality rate. Appendix Table C5 shows that this relationship increased over time for younger and older individuals, women and men, the more and less educated, and white individuals.

Appendix Figure C5 examines how the impact of area education varies with area characteristics. The figure shows the coefficient on the interaction of area human capital with being above or below the median on four area characteristics: area human capital, percent of the area population that is Black, percent of the population that is Hispanic, and population density. There is a negative correlation between area human capital and mortality across all area characteristics, but it is stronger in areas that are less educated and rural (as of 1990). However, the coefficients are not statistically different across the groups (conditional on year).

[24] Appendix Table C4 has the detailed regression results.

Area Human Capital and Non-Fatal Health Outcomes

Appendix Table C6 shows that area human capital is negatively and statistically associated with new lung and heart disease onset cases but is not associated with the onset of other conditions measured in the HRS. The magnitude of the effects is large. A 10 percentage point increase in area human capital reduces the probability of lung and heart disease by 9 and 6 percent respectively. Area human capital is also strongly positively associated with self-reported overall health and negatively associated with self-reported poor health days, as measured in the BRFSS.

Testing Explanations: Sorting, Behaviors, and Amenities

We propose three potential explanations for the relationship between area education and health: spatial sorting that is correlated with health status; sorting by amenities that are correlated with health; and changes in individual behavior. We consider them in turn.

Spatial sorting

The spatial sorting hypothesis posits that healthy individuals move to areas with higher human capital, or less healthy individuals move to areas with lower human capital, perhaps because house prices or other amenities differ. In this hypothesis, area human capital need not have any direct effect on population health; rather, the two are related because of migration.

We start with data from the HRS to test the sorting hypothesis. We create a measure of health status as the predicted probability of death in the next two years, given information on demographics and health conditions.[25] We then estimate a probit model for migration to another

[25] Specifically, we estimate a probit model for mortality that includes indicators for whether the respondent was working, baseline risk factors such as high blood pressure, ever and current smoker, BMI, and medical history (ever had heart disease, lung disease, cancer, stroke, arthritis, psychological conditions, hospitalizations). The sample for the model is 1992 to 2008.

county in the next two years, using our baseline health measure as the main explanatory variable and individual demographic and area-level controls as in the previous analyses as controls.

Column 1 in Table 2 shows that people who are less healthy are more likely to move across counties than people who are healthier.[26] The second column assesses whether healthier migrants sort into higher educated counties. Restricting the sample to movers, we estimate a regression that relates baseline health status to the difference in area education between the destination county and the origin county. There is no statistically significant association between baseline health status and the human capital differential between origin and destination counties. Conditional on moving, healthier people are not more likely to move to better educated counties.

We also look at younger ages just preceding the HRS using data from the NLSY. We use a similar approach to the HRS and predict the probability of dying between the current and next interview using a probit model relating death to characteristics in the current interview wave. Because the survey asked different health-related questions for men vs. women, we use different predictors for the two groups and report results separately. For men, the controls include 5-year age categories by race/ethnicity interactions, individual education, year, presence of any health limitations interfering with work, school, or other activities, and the type and duration of health limitations. Additional controls for women include BMI, smoking status, whether they currently have angina, hypertension, congestive heart failure, whether they have ever had an acute myocardial infarction or cancer, and presence of any health limitations affecting school, work, or other activities. We then relate baseline health to the probability of moving to a new SMSA or county before the next interview, using a probit model.

Column 3 of Table 2 shows that young women in worse baseline health are more likely to

[26] Finkelstein, Gentzkow, and Williams (2016 and 2021) report similar findings using Medicare data.

move to a new SMSA or county. This is not true among young men, as shown in column 5, but this estimate is noisy. We do not observe geographic identifiers in the NLSY, but column 4 further shows that among young women who move to a new SMSA or county, those of worse baseline health are more likely to remain in an SMSA or move to an SMSA from a location that is not an SMSA. Since SMSAs on average are higher human capital areas relative to non-SMSAs, this is inconsistent with the idea that those with better baseline health are more likely to move to high human capital areas. We thus take the NLSY results as suggestive evidence that the sorting hypothesis does not hold for younger adults either.[27,28]

Health Behaviors

We now turn to our second hypothesis: which area human capital affects health-related behaviors. We focus on the two behavioral health risk factors that contribute the most to mortality in the US – smoking and obesity (Mokdad et al., 2004; Cutler and Lleras-Muney, 2010).

We use data on smoking status and obesity from the BRFSS and smoking status from the CPS, each matched to area characteristics measured in the decennial census or ACS wave most closely preceding the given year.[29] We use similar regression models as for our baseline mortality regressions in column 5 of Table 1 but instead of mortality as the dependent variable, we use whether the individual was a current, former, or never smoker, three categories of obesity (very

---

[27] As an alternative approach to address spatial sorting, we also examined an instrumental variables approach used in Moretti (2004b), Wheeler (2008), and Diamond (2016). These papers use two instruments for the level of human capital in an area: (1) the presence of a land-grant college in the area established by the federal Morrill Act of 1862 and (2) a shift-share instrument using the demographic structure of an area from 1980 in combination with secular national changes in educational attainment biased towards more college over time. However, both instruments yielded F-statistics less than 10 in the first stage after including individual education, indicating weak instruments.

[28] Our empirical tests cannot reject spatial sorting prior to the span of our data or in the long-term. For instance, if (i) innately healthier individuals move to certain areas prior to our time period, (ii) innate health is transmitted across generations, and (iii) areas that had a large inflow of innately healthier individuals also established institutions or a culture of education that persists through the present, this would also yield a positive correlation between health and area human capital. While it is plausible that innate health may be transmitted intergenerationally, we are unaware of any papers showing that historically healthier areas have better education systems today.

[29] E.g., 1995 data is merged to area data from the 1990 census, 2003 to area characteristics from the 2000 census, and 2014 to area data from the 2009-2011 ACS).

obese [BMI>35], obese [35>=BMI>30], and overweight [30>=BMI>25] vs. normal or underweight), and whether the individual reported mainly being physically inactive (vs. being physically active). We also use probit instead of OLS for estimation.

Figure 2 shows the results.[30] Area human capital is strongly negatively correlated with the probability of being a current smoker and being obese. The coefficient using the CPS data implies that individuals in areas with 10 percentage points more college graduates are 2.2 percentage points less likely to be current smokers, equivalent to a 13.5% decrease in the probability of smoking relative to the average smoking rate. People in high human capital areas are also more likely to have never smoked and to quit smoking conditional on ever starting smoking. Area human capital is also statistically significantly associated with a lower probability of being overweight or obese – a 10 percentage point increase in the percentage of college graduates in an area is associated with a 12.4% lower likelihood of being very obese and a 4.1% lower probability of being obese. People are also less likely to engage in no physical activity in areas with higher human capital. These findings closely align with causal neighborhood effects on obesity from the Moving to Opportunity experiment (Ludwig et al., 2013). Appendix Tables C9 and C10 show that the relationship between area human capital and health-related behaviors has slightly strengthened over time.

Appendix Figure C6 plots the coefficients on area human capital interacted with age from regression models using our standard set of controls, where the outcomes are smoking non-initiation (never smoking) and the smoking quit rate (formerly smoking conditional on ever smoking). Spillovers of area human capital on preventing smoking initiation begin in young adulthood. The effects are largest from ages 25-49. The biggest spillovers on quitting smoking occur in the late 30s and early 40s. At this pivotal age, 10-15 years of prior smoking is sufficient

[30] See Appendix Table C7 and Appendix Table C8 for detailed regression results.

to harm health, but quitting ameliorates the risk of these harms manifesting in 10-20 years, when these individuals are in their 50s and 60s (Durham and Adcock, 2015). Higher quit rates and lower initiation of smoking across the age spectrum will yield strengthening relationships between health and area human capital over the next 10-20 years, as people age into late middle age and retirement years.[31]

To summarize how much these variables can explain of the effect of area education on mortality, we re-estimate our central model in column 5 of Table 1, including measures of smoking and obesity in the area as controls. For smoking, we use the BRFSS data as a baseline and supplement it with CPS data prior to 1999. Area-level data on obesity comes from the 1999-2001 and 2009-2011 BRFSS and is only available for those years.[32] Even with these noisy measures of smoking and obesity, Table 3 shows that controlling for differential smoking rates explains about 38%-46% of the effect of area human capital on all-cause mortality, depending on the year (as demonstrated by the difference in the coefficients between column 1 and column 2 for 1990-2010 and column 3 and 4 for 2000-2010). Controlling for both smoking and obesity explains 59% of the relationship between area human capital and mortality in 2000 and 2010, mostly driven by differences in smoking. Smoking and obesity are particularly good for explaining the correlation between area human capital and deaths due to cancer, chronic lower respiratory disease, stroke, and drug-related deaths, as shown in Appendix Table C11.

Other Health-Related Amenities

We next turn to whether some of the remaining differences in mortality across areas can

[31] Consistent with the hypothesis of accumulating effects, in a regression including lagged and contemporaneous area human capital, initial levels of human capital are strongly negatively correlated with mortality for young individuals (<65 years) but not their older counterparts.

[32] Since not all areas are represented in the CPS and BRFSS, we estimate the models including smoking, obesity, and physical activity among cells where we have available data on these behaviors. Thus, the number of observations and average mortality rates reported in Table 3 are lower than the ones reported in Appendix Table C1.

be explained by correlations between area education and health amenities that may not operate primarily through health behaviors. We focus on three external stressors – air pollution, crime, and healthcare quality and access – while acknowledging that other environmental factors beyond these may affect health.

Significant literature shows that exposure to air pollution increases lung disease incidence or aggravation of existing lung disease, cancer, and premature death (Environmental Protection Agency, 2023). High levels of air pollution may also discourage outdoor exercise and thus indirectly impact mortality through obesity. Higher area human capital could also be associated with less pollution because air quality may be priced into property values, leading to selection of the better educated (and wealthy) into such areas. While homicides are a crude measure that may not capture all aspects of crime, they are more reliably reported than other crimes (Bureau of Justice Statistics, 1994). Crime could also decrease health through indirect channels; for example, unsafe streets could increase stress, lead residents to stay inside and get less exercise, or make it difficult to obtain necessary health care or management of chronic conditions.

Additionally, pollution and crime might be lower in more educated areas for similar reasons that demand for high-quality medical care might be higher – people in higher human capital areas may pay more for public goods addressing environmental stressors and may possess the political clout to regulate crime and pollution. We also control for differences in healthcare quality captured by hospital quality since most deaths occur in hospitals/nursing homes and mortality may be particularly sensitive to this dimension of healthcare quality.

Figure 3 builds on the regressions with smoking and obesity controls from column 5 of Table 3 and examines whether controlling for differences in health-related amenities across areas can explain the correlation between area human capital and mortality above and beyond what is explained by health-related behaviors. The first set of models in Figure 3 start from the estimates in column 6 of Table 3, which adjusts for smoking and obesity, and sequentially add in pollution, crime, number of physicians, and number of hospital beds. The second set of models start from the model including all previously mentioned health amenities as controls and adds demand for preventative care as a control. The last set of models uses data for 2010 only and controls for health care supply and quality.[33]

In total, external factors such as pollution and homicide rates explain a small share of the relationship between area human capital and mortality after controlling for smoking and obesity. Both pollution and homicide are positively correlated with mortality, but neither explains much of the relationship between area human capital and mortality. Similarly, measures of health care supply and quality, such as the number of physicians, hospital beds, and health care quality, which are also correlated with mortality, cannot explain the effect of area human capital on mortality beyond what is already explained by differences in smoking. Preventative care measures, particularly the percent of women who timely go for mammogram screenings, explain 32% of the relationship beyond what is already explained by smoking and obesity (i.e., around 12% of the baseline relationship between area human capital and mortality reported in Table 1). Health-related amenities in total explain 17% of the relationship between area human capital and mortality conditional on individual education.

Understanding Health-Related Behaviors

Regulation

[33] Appendix Table C12 shows the corresponding regression results in table form. Appendix Figure C7 shows the same results but building on column 2 in Table 3, which does not include obesity as a control and thus allows us to include data from 1990.

The importance of health-related behaviors for the link between area education and mortality suggests a deeper look at the relationship between behaviors and area education. One possible reason the two might be related is through prices and regulation. More educated areas may be more likely to support legislation and regulations aimed at improving health, such as tobacco taxes, clean indoor air laws, and workplace smoking bans.[34] Tobacco taxes and state clean indoor air laws and regulations will typically vary by state and year since they are set federally or at the state level. Since we include state-by-year fixed effects in our specifications, these variables are already accounted for. Thus, we focus on private workplace smoking bans, which can be implemented as company policy independent from law and thus may vary within states and years. The CPS data described above ask questions on workplace smoking policies for indoor workers. We focus on whether the workplace has an official smoking policy in place and whether the workplace bans smoking in all public and work areas.[35]

Table 4 shows the impact of area human capital on these policies. Controlling for own education, individuals in more educated areas are more likely to work at places with a complete ban on smoking in all public and work areas. A worker in an area with a 10 percentage point higher share of college graduates is 1.8 percentage points more likely to be employed at places with a complete smoking ban.

Peer Effects and Social Norms

A second theory is that area human capital drives peer effects, leading to the development of different social norms in high and low human capital areas. For instance, the proximity of more

[34] A growing body of literature suggests that education has a causal effect on voter turnout, political involvement, and information on politics in the U.S. (Dee, 2004; Milligan et al., 2004; Borgonovi et al., 2010; Sondheimer and Green, 2010).

[35] Several papers discuss the effectiveness of these bans in reducing smoking. For example, Evans, Farrelly, and Montgomery (1999) show that compared to a firm with few restrictions on smoking, adopting a smoke-free policy at a workplace reduces the probability of smoking by 5.7 percentage points and decreases the daily number of cigarettes smoked by 14% on average.

educated individuals undertaking healthy behaviors may encourage individuals across the education distribution to undertake healthy behaviors themselves. Differences in information and beliefs about the harmful effects of smoking and obesity, which may correlate with area human capital, may also be driving these differences in smoking behavior across areas. These changes in social norms may result from public health messaging campaigns or may arise more organically as peer effects.

While we cannot directly assess peer effects, we can examine informational spillovers and attitudes towards smoking. The 1987, 1992, and 2000 National Health Interview Surveys (NHIS) asked individuals about their agreement with a series of statements about the effects of smoking on health: smoking by pregnant women is harmful for the baby, someone else's smoke is harmful, and most lung cancer deaths are caused by smoking, among others. We consider how these are related to area education.[36]

Table 5 shows that NHIS respondents living in counties with a 10 percentage point higher percentage of college graduates are 11% more likely to agree with the statement that smoking is harmful for pregnant women's babies and 15% more likely to agree that most lung cancer deaths stem from smoking, controlling for individual education. These results are statistically significant at the 10% level and 5% level, respectively. As the next columns show, individuals in more educated areas are also more likely to support smoking bans in bars, restaurants, and work areas, based on data from the CPS, reflecting attitudes towards smoking and regulation of it.

We lastly examine whether smoking regulations and beliefs and attitudes about smoking mediate the correlation between area human capital and smoking behavior. These results are not

[36] Each year in the NHIS data was merged to area characteristics measured in the decennial census immediately prior to the given year (i.e., 1987 is merged to area characteristics from 1980, 1992 to area data from 1990, 2000 to area data from 2000). We use counties instead of CONSPUMAs in the NHIS.

causal since regulations and beliefs are not randomly assigned across areas but are indicative of the relationship between beliefs about the harms of smoking or local smoking regulations and smoking initiation/cessation across areas. Differences in the coefficient on the percent college graduates across the columns of Appendix Table C12 suggest that believing that smoking should be banned everywhere can explain 12%, 15%, and 22% of the correlation between area human capital and current smoking, former smoking, and never smoking, respectively. Workplace smoking bans explain an additional 2-5% of the correlation. Shifting social norms around smoking are thus associated with a substantial share of the divergence in smoking and subsequent mortality between high- and low-human capital areas.

Conclusion

This paper documents a strong and robust relationship between area human capital and mortality, even after controlling for individual education. This relationship existed in 1990 and strengthened in the next two decades.

More than half of the correlation between area human capital and mortality can be explained by differences in smoking rates and obesity rates across areas. We further examine two channels for these spillovers and find empirical evidence for both regulatory policies that increase the price of unhealthy behaviors, such as workplace smoking bans; and peer effects in beliefs about the harms of smoking. In contrast, health-related amenities such as pollution, homicides, health care quality, and quantity can explain at most 17% of the correlation between area human capital and mortality. We find no evidence for spatial sorting driven by health – healthier individuals are no more likely to move to more educated areas than less healthy individuals.

Our paper points towards mechanisms that can help explain why locations have such a powerful impact on health, shedding light on local policies that do not target health directly but may affect it indirectly. Even without direct effects of local and place-based labor or educational policies on health, any welfare and general equilibrium analysis of such policies may need to incorporate spillovers on health. Health-related behaviors are particularly sensitive to human capital spillovers among younger individuals, implicating the role of changing social norms around smoking and obesity across generations in the widening geographic gaps in health between high and low human capital areas.

References

Berry, Christopher R., and Edward L. Glaeser, “The divergence of human capital levels across cities,” Papers in Regional Science, 84 (2005), 407–444.

Bor, Jacob, Gregory H. Cohen, and Sandro Galea, “Population health in an era of rising income inequality: USA, 1980–2015,” The Lancet, 389 (2017), 1475–1490.

Borgonovi, Francesca, Béatrice d’Hombres, and Hoskins Bryony, “Voter Turnout, Information Acquisition and Education: Evidence from 15 European Countries,” The B.E. Journal of Economic Analysis & Policy, 10 (2010), 1–34 (De Gruyter).

Bureau of Justice Statistics; Office of Justice Programs; U.S. Department of Justice. Fact Sheet: Drug-Related Crime. September 1994.

Bureau of Health Workforce. Area Resource Files, 1990, 2000, 2010.

Cacioppo, John T., James H. Fowler, and Nicholas A. Christakis, “Alone in the crowd: the structure and spread of loneliness in a large social network,” Journal of Personality and Social Psychology, 97 (2009), 977–991.

Canton, Erik, “Social Returns to Education: Macro-Evidence,” De Economist, 155 (2007), 449–468.

Case, Anne, and Angus Deaton, “Accounting for the Widening Mortality Gap between American Adults with and without a BA,” Brookings Papers on Economic Activity, Fall (2023), 1–44.

Centers for Medicare and Medicaid Services. Hospital Compare Database 2008-2012.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler, “The Association Between Income and Life Expectancy in the United States, 2001-2014,” JAMA, 315 (2016), 1750.

Christakis, Nicholas A., and James H. Fowler, “The Spread of Obesity in a Large Social Network over 32 Years,” New England Journal of Medicine, 357 (2007), 370–379.

——, “The Collective Dynamics of Smoking in a Large Social Network,” New England Journal of Medicine, 358 (2008), 2249–2258.

Cohen-Cole, Ethan, and Jason M. Fletcher, “Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic,” Journal of Health Economics, 27 (2008), 1382–1387.

Cutler, David, and Edward Glaeser, “Social Interactions and Smoking,” (University of Chicago Press, 2010).

Cutler, David M, Fabian Lange, Ellen Meara, Seth Richards-Shubik, and Christopher J Ruhm, “Rising Educational Gradients in Mortality: The Role of Behavioral Risk Factors,” Journal of health economics, 30 (2011), 1174–1187.

Cutler, David M., and Adriana Lleras-Muney, “Education and Health: Evaluating Theories and Evidence,” Working Paper Series, Working Paper, 2006 (National Bureau of Economic Research).

——, “Understanding differences in health behaviors by education,” Journal of Health Economics, 29 (2010), 1–28.

——, “Education and Health: Insights from International Comparisons,” Working Paper Series, Working Paper, 2012 (National Bureau of Economic Research).

Dee, Thomas S., “Are there civic returns to education?,” Journal of Public Economics, 88 (2004), 1697–1720.

Deryugina, Tatyana, and David Molitor, “Does When You Die Depend on Where You Live? Evidence from Hurricane Katrina,” American Economic Review, 110 (2020), 3602–3633.

Diamond, Rebecca, “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000,” American Economic Review, 106 (2016), 479–524.

van Donkelaar, Aaron, Randall V. Martin, Chi Li, and Richard T. Burnett, “Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors,” Environmental Science &

Technology, 53 (2019), 2595–2611.

Doyle, Joseph J., “Returns to Local-Area Healthcare Spending: Evidence from Health Shocks to Patients Far From Home,” American economic journal. Applied economics, 3 (2011), 221–243.

Durham, A.L., and I.M. Adcock, “The relationship between COPD and lung cancer,” Lung Cancer 90 (2015), 121–127.

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, “Inequalities in Life Expectancy Among US Counties, 1980 to 2014: Temporal Trends and Key Drivers,” JAMA Internal Medicine, 177 (2017), 1003–1011.

Elo, Irma T., and Samuel H. Preston, “Educational differentials in mortality: United States, 1979–1985,” Social Science & Medicine, 42 (1996), 47–57.

Environmental Protection Agency (EPA). Air Quality Data 1989-1991, 1999-2001, 2009-2011.

Environmental Protection Agency (EPA). Air Quality Guide for Particle Pollution, 2023.

Evans, William N, Matthew C Farrelly, and Edward Montgomery, “Do Workplace Smoking Bans Reduce Smoking?,” American Economic Review, 89 (1999), 728–747.

Ezzati, Majid, Ari B. Friedman, Sandeep C. Kulkarni, and Christopher J. L. Murray, “The Reversal of Fortunes: Trends in County Mortality and Cross-County Mortality Disparities in the United States,” PLOS Medicine, 5 (2008), e66.

Finkelstein A, Gentzkow M, Williams H. Sources of Geographic Variation in Health Care: Evidence From Patient Migration. The Quarterly Journal of Economics 2016;131(4):1681–726.

——, “Place-Based Drivers of Mortality: Evidence from Migration,” American Economic Review, 111 (2021), 2697–2735.

Flegal, Katherine M., Barry I. Graubard, David F. Williamson, and Mitchell H. Gail, “Cause-Specific Excess Deaths Associated With Underweight, Overweight, and Obesity,” JAMA, 298 (2007), 2028.

Fowler, James H., and Nicholas A. Christakis, “Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study,” BMJ (Clinical research ed.), 337 (2008a).

Fowler, J.H., and N.A. Christakis, “Estimating Peer Effects on Health in Social Networks : A Response to Cohen-Cole and Fletcher; Trogdon, Nonnemaker, Pais,” Journal of health economics, 27 (2008b), 1400–1405.

Galama, Titus J., Adriana Lleras-Muney, and Hans van Kippersluis, “The Effect of Education on Health and Mortality: A Review of Experimental and Quasi-Experimental Evidence,” SSRN Scholarly Paper, 2017 (Rochester, NY).

Glaeser, Edward L., Albert Saiz, Gary Burtless, and William C. Strange, “The Rise of the Skilled City [with Comments],” Brookings-Wharton Papers on Urban Affairs, (2004), 47–105 (Brookings Institution Press).

Greenwood, Jeremy, Nezih Guner, Georgi Kocharkov, and Cezar Santos, “Marry Your Like: Assortative Mating and Income Inequality,” American Economic Review Papers and Proceedings, 104 (2014), 348–353.

Grossman, Michael, “Chapter 10 Education and Nonmarket Outcomes,” in Handbook of the Economics of Education (Elsevier, 2006).

——, “The Relationship between Health and Schooling,” Eastern Economic Journal, 34 (2008), 281–292 (Palgrave Macmillan Journals).

——, “The Relationship between Health and Schooling: What’s New?,” Nordic Journal of Health Economics, 3 (2015), 7–17.

Hayward, Mark D., Robert A. Hummer, and Isaac Sasson, “Trends and group differences in the association between educational attainment and U.S. adult mortality: Implications for understanding education’s causal influence,” Social Science & Medicine, Special Issue: Educational Attainment and Adult Health: Contextualizing Causality, 127 (2015), 8–18.

Inter-university Consortium for Political and Social Research. United States Department of Justice.

Federal Bureau of Investigation. Uniform Crime Reports, 1989-1991, 1999-2001, 2009-2011.

Iranzo, Susana, and Giovanni Peri, “Schooling Externalities, Technology, and Productivity: Theory and Evidence from U.S. States,” The Review of Economics and Statistics, 91 (2009), 420–431.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment*,” The Quarterly Journal of Economics, 116 (2001), 607–654.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, “Experimental Analysis of Neighborhood Effects,” Econometrica, 75 (2007), 83–119.

Krieger, Nancy, David H Rehkopf, Jarvis T Chen, Pamela D Waterman, Enrico Marcelli, and Malinda Kennedy, “The Fall and Rise of US Inequities in Premature Mortality: 1960–2002,” PLoS Medicine, (T. E. Novotny, ed.), 5 (2008), e46.

Kulkarni, Sandeep C., Alison Levin-Rector, Majid Ezzati, and Christopher Jl Murray, “Falling behind: life expectancy in US counties from 2000 to 2007 in an international context,” Population Health Metrics, 9 (2011), 16.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu, “Neighborhood effects on the long-term well-being of low-income adults,” Science (New York, N.Y.), 337 (2012), 1505–1510.

——, “Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity,” American Economic Review, 103 (2013), 226–231.

Ludwig, Jens, Lisa Sanbonmatsu, Lisa Gennetian, Emma Adam, Greg J. Duncan, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, Stacy Tessler Lindau, Robert C. Whitaker, and Thomas W. McDade, “Neighborhoods, Obesity, and Diabetes — A Randomized Social Experiment,” New England Journal of Medicine, 365 (2011), 1509–1519.

Masters, Ryan K., Robert A. Hummer, and Daniel A. Powers, “Educational Differences in U.S. Adult Mortality: A Cohort Perspective,” American Sociological Review, 77 (2012), 548–572.

Meara, Ellen R., Seth Richards, and David M. Cutler, "The Gap Gets Bigger: Changes In Mortality And Life Expectancy, By Education, 1981–2000," Health Affairs, 27 (2008), 350–360.

Mednick, Sara C., Nicholas A. Christakis, and James H. Fowler, "The Spread of Sleep Loss Influences Drug Use in Adolescent Social Networks," PLoS ONE, (K. Hashimoto, ed.), 5 (2010), e9775.

Meng, Jun, Chi Li, Randall V. Martin, Aaron van Donkelaar, Perry Hystad, and Michael Brauer, "Estimated Long-Term (1981–2016) Concentrations of Ambient Fine Particulate Matter across North America from Chemical Transport Modeling, Satellite Remote Sensing, and Ground-Based Measurements," Environmental Science & Technology, 53 (2019), 5071–5079.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos, "Does education improve citizenship? Evidence from the United States and the United Kingdom," Journal of Public Economics, 88 (2004), 1667–1695.

Mokdad, Ali H., James S. Marks, Donna F. Stroup, and Julie L. Gerberding, "Actual Causes of Death in the United States, 2000," JAMA, 291 (2004), 1238–1245.

Montez, Jennifer Karas, Nader Mehri, Shannon M. Monnat, Jason Beckfield, Derek Chapman, Jacob M. Grumbach, Mark D. Hayward, Steven H. Woolf, and Anna Zajacova, "U.S. state policy contexts and mortality of working-age adults," PLOS ONE, 17 (2022).

Moretti, Enrico, "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions," American Economic Review, 94 (2004a), 656–690.

——, "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data," Journal of Econometrics, Higher education (Annals issue), 121 (2004b), 175–212.

——, "Real Wage Inequality," American Economic Journal: Applied Economics, 5 (2013), 65–103.

Murray, Christopher J. L, Sandeep C Kulkarni, Catherine Michaud, Niels Tomijima, Maria T Bulzacchelli, Terrell J Iandiorio, and Majid Ezzati, "Eight Americas: Investigating Mortality

Disparities across Races, Counties, and Race-Counties in the United States," PLoS Medicine, (T. Novotny, ed.), 3 (2006), e260.

Murray, Christopher J. L., Sandeep Kulkarni, and Majid Ezzati, "Eight Americas: New Perspectives on U.S. Health Disparities," American Journal of Preventive Medicine, 29 (2005), 4–10 (Elsevier).

National Center for Health Statistics (NCHS). Multiple Cause of Death Files, 1989-2010.

Nolte, Ellen, and C. Martin McKee, "Measuring The Health Of Nations: Updating An Earlier Analysis," Health Affairs, 27 (2008), 58–71 (Health Affairs).

Olshansky, S. Jay, Toni Antonucci, Lisa Berkman, Robert H. Binstock, Axel Boersch-Supan, John T. Cacioppo, Bruce A. Carnes, Laura L. Carstensen, Linda P. Fried, Dana P. Goldman, James Jackson, Martin Kohli, John Rother, Yuhui Zheng, and John Rowe, "Differences In Life Expectancy Due To Race And Educational Differences Are Widening, And Many May Not Catch Up," Health Affairs, 31 (2012), 1803–1813.

Rauch, James, "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," Journal of Urban Economics, 34 (1993), 380–400 (Elsevier).

Rosamond, Wayne D., Herman A. Tyroler, Lloyd E. Chambless, Aaron R. Folsom, Lawton Cooper, and David Conwill, "Educational Achievement Recorded on Certificates of Death Compared with Self-Report," Epidemiology, 8 (1997), 202–204 (Lippincott Williams & Wilkins).

Rosenquist, J. N., J. H. Fowler, and N. A. Christakis, "Social network determinants of depression," Molecular Psychiatry, 16 (2011), 273–281.

Rosenthal, Stuart S., and William C. Strange, "The attenuation of human capital spillovers," Journal of Urban Economics, 64 (2008), 373–389.

Rostron, Brian L., John L. Boies, and Elizabeth Arias, "Education reporting and classification on death certificates in the United States," Vital and Health Statistics. Series 2, Data Evaluation and Methods Research, (2010), 1–21.

Ruggles S, Flood S, Foster S, Goeken R, Pacas J, Schouweiler M, Sobek M. IPUMS USA: Version 11.0 [dataset]. Minneapolis, MN: IPUMS, 2021.

Sasson, Isaac, “Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990-2010,” Demography, 53 (2016), 269–293.

Shai, D., and I. Rosenwaike, “Errors in reporting education on the death certificate: some findings for older male decedents from New York State and Utah,” American Journal of Epidemiology, 130 (1989), 188–192.

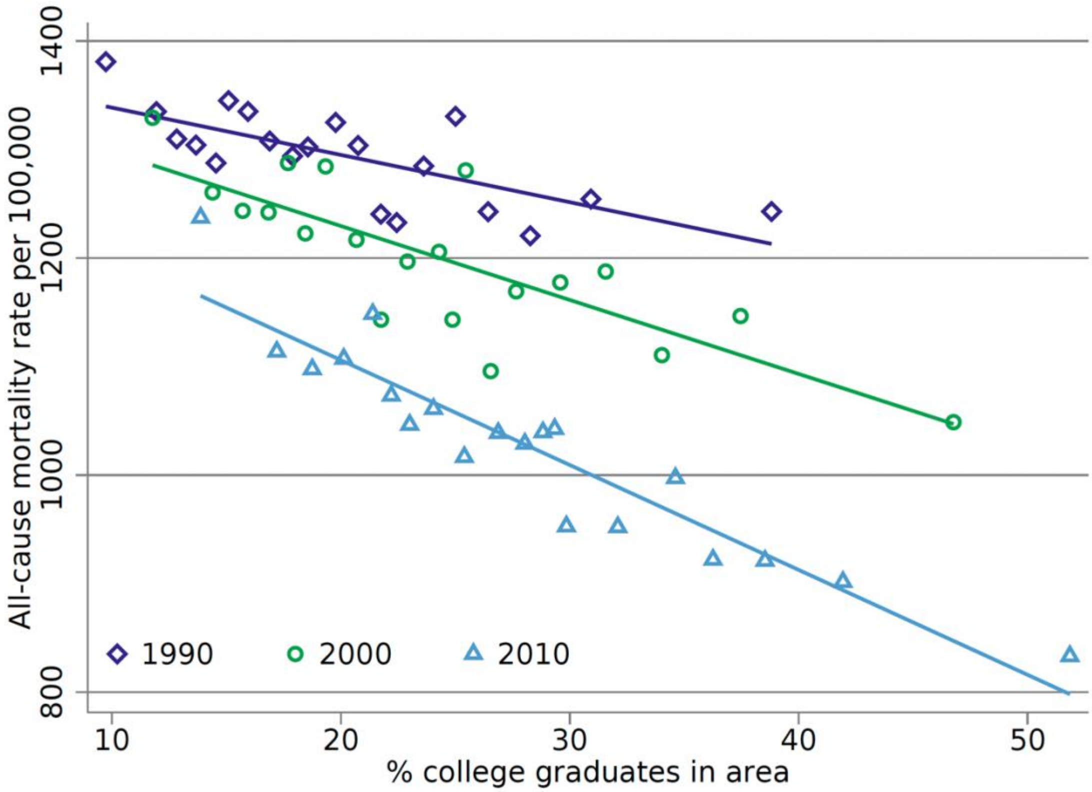
Sondheimer, Rachel Milstein, and Donald P. Green, “Using Experiments to Estimate the Effects of Education on Voter Turnout,” American Journal of Political Science, 54 (2010), 174–189.

Sorlie, Paul D., and Norman J. Johnson, “Validity of Education Information on the Death Certificate,” Epidemiology, 7 (1996), 437–439 (Lippincott Williams & Wilkins).

Wheeler, Christopher H., “Human Capital Externalities and Adult Mortality in the U.S.,” SSRN Scholarly Paper, 2008 (Rochester, NY, Social Science Research Network).

Tables and Figures

Figure 1: The relationship between area human capital and adjusted mortality per 100,000



Note: This graph is a binned scatter plot showing all-cause mortality across ventiles of area education in each year using mortality data that includes those with missing individual education. Each point includes approximately 5% of the population in that year, plotted at the mean percent college graduates and mean mortality rate (adjusted for age-sex using direct adjustment) across areas within each bin. The coefficients (and standard errors in parentheses) of the corresponding OLS regressions are -4.4*** (0.97) in 1990, -6.8*** (0.73) in 2000, and -9.7*** (0.97) in 2010. Fitted lines extend to the full range of the underlying data in each year. Areas are defined as consistent public use microdata areas (CONSPUMAs). *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Table 1: Regression results of all-cause mortality rates per 100,000 on area human capital

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Area characteristics								
% college graduates		-10.30***	-6.41***	-0.72	-8.46***	-6.00***	-8.22***	-0.51
		(1.38)	(1.31)	(1.67)	(1.22)	(1.48)	(1.80)	(2.31)
% college graduates * year=2000				-4.25***		-1.47**		-2.09***
				(1.06)		(0.65)		(0.69)
% college graduates * year=2010				-7.32***		-3.35***		-4.22***
				(1.28)		(0.75)		(0.84)
Log density (pop/sq mi)	-7.42	-4.18	-5.89	-6.12	-3.35	-3.55	-1401.06**	-1029.62*
	(6.39)	(7.08)	(6.89)	(6.72)	(5.44)	(5.41)	(585.39)	(561.99)
Log population	-1.93	-1.04	-1.12	-0.91	-2.71	-2.29	1377.60**	1007.31*
	(4.77)	(4.87)	(4.76)	(4.78)	(4.18)	(4.17)	(584.55)	(560.48)
% Black	2.92***	2.45***	2.66***	2.71***	1.76***	1.83***	5.25**	6.35***
	(0.61)	(0.63)	(0.61)	(0.60)	(0.59)	(0.59)	(2.26)	(2.21)
% Hispanic	1.16*	0.21	1.37**	1.15*	-0.67	-0.75	-2.21	-2.98*
	(0.64)	(0.63)	(0.63)	(0.62)	(0.50)	(0.51)	(1.62)	(1.54)
Cell characteristics (ref. group: no high school)								
High school graduate	-265.23***		-265.65***	-266.58***	-263.47***	-263.81***	-261.23***	-261.43***
	(40.31)		(40.32)	(40.36)	(40.54)	(40.56)	(40.47)	(40.48)
Some college	-448.30***		-446.65***	-448.31***	-446.51***	-447.03***	-443.81***	-444.09***
	(29.07)		(29.06)	(29.12)	(29.26)	(29.30)	(29.17)	(29.18)
College graduate	-543.94***		-539.04***	-540.12***	-537.61***	-537.98***	-534.69***	-534.93***
	(35.32)		(35.23)	(35.26)	(35.42)	(35.44)	(35.33)	(35.34)
Post-graduate education	-729.68***		-724.10***	-724.89***	-722.09***	-722.38***	-720.41***	-720.49***
	(36.64)		(36.47)	(36.51)	(36.69)	(36.72)	(36.53)	(36.54)
% deaths with missing education	-7.10***	-6.22***	-7.32***	-7.02***	-4.41***	-4.27***	-3.94**	-3.84**
	(2.25)	(2.32)	(2.27)	(2.31)	(1.61)	(1.61)	(1.74)	(1.75)
State-year FE	No	No	No	No	Yes	Yes	Yes	Yes
Area FE	No	No	No	No	No	No	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966
Area-year-age-sex-race-educ cells	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707
Areas	486	486	486	486	486	486	486	486
R-squared	0.860	0.855	0.860	0.860	0.862	0.862	0.862	0.862
Dependent var. mean	1,196	1,196	1,196	1,196	1,196	1,196	1,196	1,196
% change from 10pp increase in % college grads		*-8.6*	*-5.4*		*-7.1*		*-6.9*	*-0.4*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, year. Areas are defined as consistent public use microdata areas (CONSPUMAs).

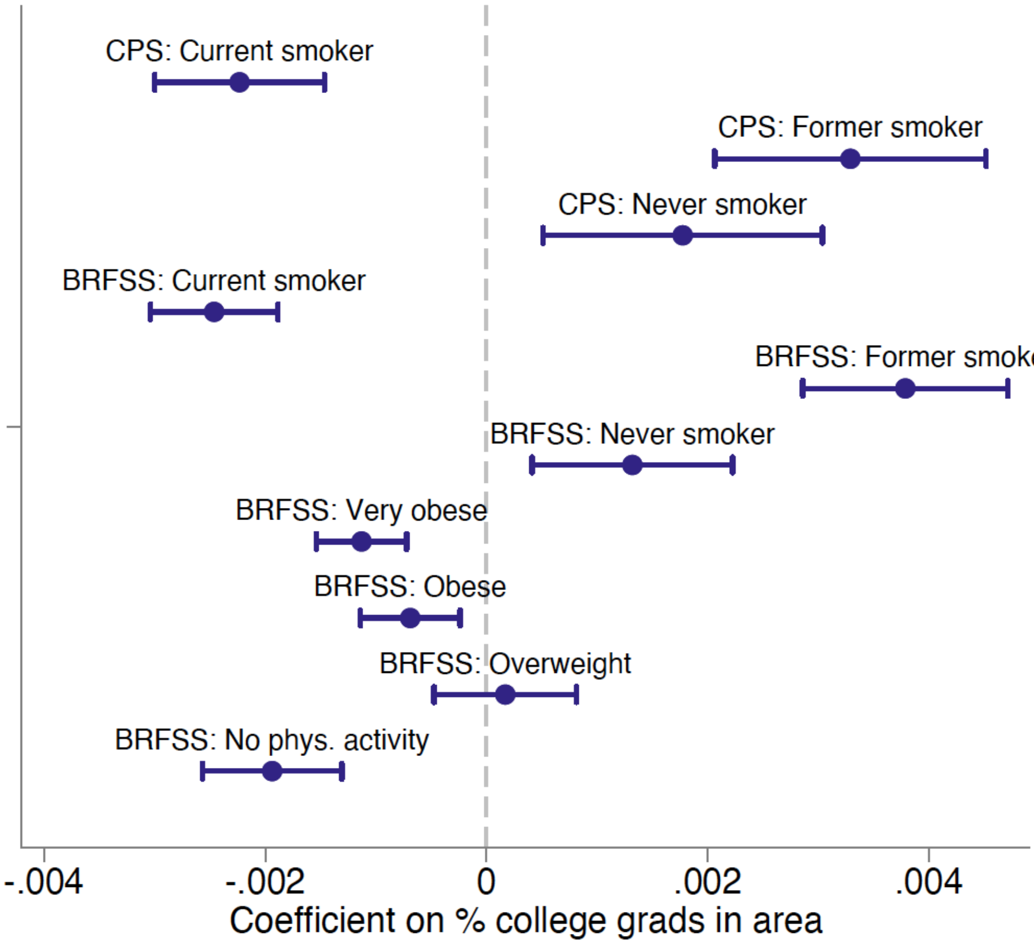
Table 2: Spatial sorting, baseline health status, and selective migration

	(1)	(2)	(3)	(4)	(5)
Dependent var.	**Migrated to different county in next 2 years?**	**% college grads in destination county - minus - % college grads in county of origin**	**Migrated to different SMSA/county between interviews?**	**Stayed in SMSA or moved to SMSA from non-SMSA?**	**Migrated to different SMSA/county between interviews?**
	HRS 1992-2008	HRS 1992-2008	NLSY Young Women 1968-2001	NLSY Young Women 1968-2001	NLSY Young Men 1969-1971, 1976
Model	Probit, dy/dx (SE)	OLS, coef. (SE)	Probit, dy/dx (SE)	Probit, dy/dx (SE)	Probit, dy/dx (SE)
Baseline health status: predicted mortality until next interview	0.051** (0.022)	0.092 (5.23)	0.162** (0.064)	0.568 (0.385)	-0.114 (0.172)
Observations	71,717	3,101	50,722	3,010	4,527
R-squared	0.0031	0.017	0.0341	0.08	0.0318
Dependent var. mean	0.043	-0.42	0.094	0.573	0.092

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the individual-level.

Note: All regressions use sampling weights and control for individual education, 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, National Longitudinal Survey of Youth (NLSY) wave or Health and Retirement Survey (HRS) survey wave, Black, and Hispanic. In the NLSY Young Men regressions, we use 1-year age categories instead of 5-year due to similar ages in the sample. Columns 1-2 additionally control for area characteristics: log population, log density, and industry shares. Baseline health in the HRS regressions was measured as the probability of death between the current and next interview and as probability of dying by 2011 in the NLSY regressions; it is estimated in a separate probit regression of mortality on measures of health status in the current interview and demographics.

Figure 2: Regression results of health-related behaviors on area human capital



Note: This figure plots the coefficient on area human capital estimated separately for each smoking-related and obesity-related behavior, all of which are defined as binary variables. Former smoker is defined conditional on ever smoking. All probit regressions pool data from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS) or the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015. All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

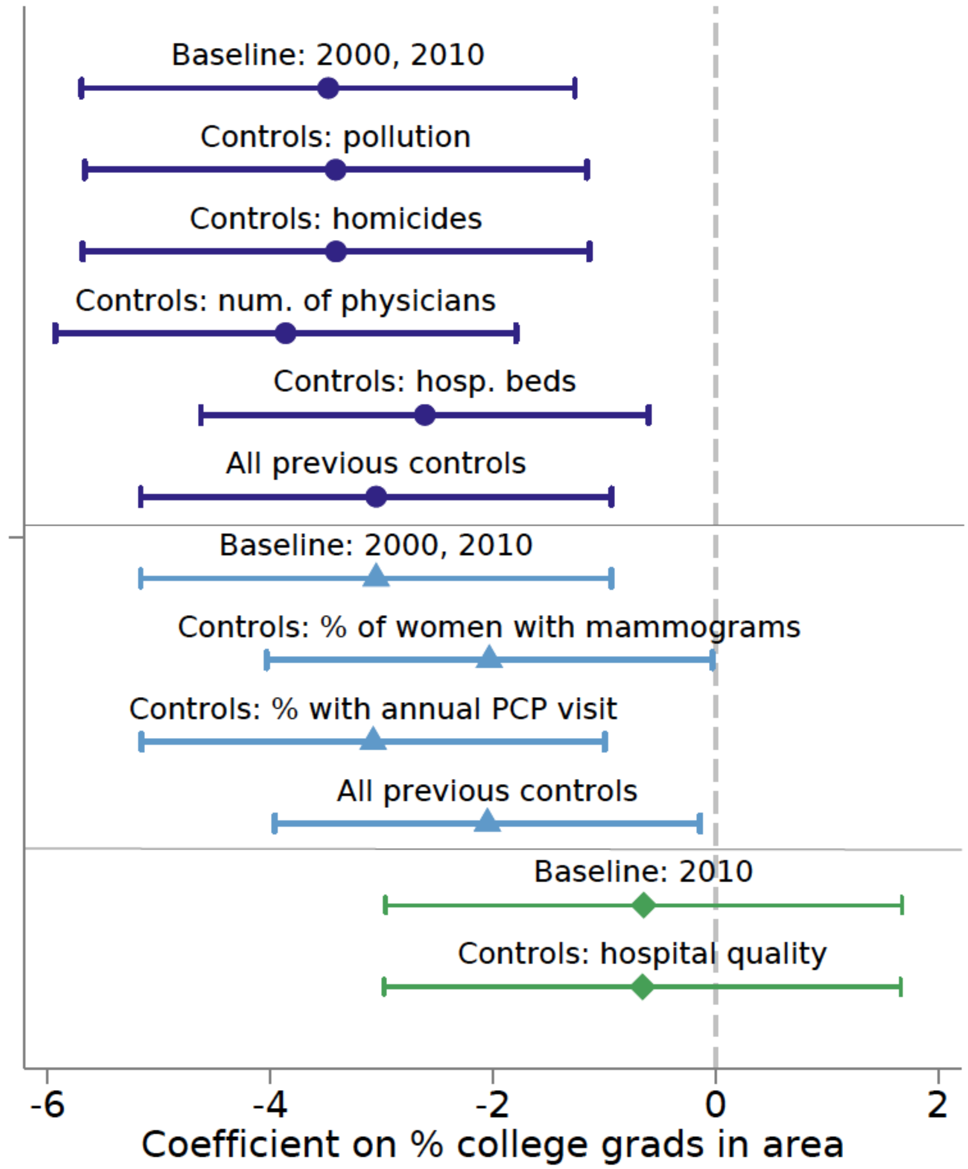
Table 3: Regression results of all-cause mortality per 100,000 on area human capital and health-related behaviors

	(1)	(2)	(3)	(4)	(5)	(6)
	Cells with non-missing data on smoking behavior		**Cells with non-missing data on smoking behavior and obesity-related behaviors**			
	1990-2010		2000-2010			
Area characteristics						
% college graduates	-8.97***	-5.58***	-8.53***	-4.63***	-7.51***	-3.48***
	(1.20)	(1.09)	(1.22)	(1.14)	(1.22)	(1.13)
% current smoker		5.24**		7.23***		6.76**
		(2.19)		(2.76)		(2.70)
% former smoker		-229.65		-165.63		-205.95
		(142.57)		(184.21)		(176.17)
% overweight, obese, very obese					3.88***	4.18***
					(1.08)	(1.09)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	433,629,834	433,629,834	357,070,962	357,070,962	357,070,962	357,070,962
Cells	308,969	308,969	257,906	257,906	257,906	257,906
Areas	485	485	485	485	485	485
R-squared	0.871	0.871	0.879	0.879	0.879	0.879
Dependent var. mean	1,192	1,192	1,206	1,206	1,206	1,206
% change from 10pp increase in % college grads	*-7.5*	*-4.7*	*-7.1*	*-3.8*	-6.2	*-2.9*

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and year. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. The percentage of current or former smokers was calculated using the 1995-1996, 1998-1999 CPS, 1999-2001 BRFSS, and 2009-2011 BRFSS. The percent of individuals that were overweight, obese, or very obese, and those with no physical activity were calculated using the 1999-2001 and 2009-2011 BRFSS. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Figure 3: Regression results of all-cause mortality per 100,000 on area human capital and health-related amenities



Note: All regressions are estimated separately at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. The main control in the baseline regressions is the percentage of individuals currently smoking, the percentage of individuals formerly smoking (conditional on ever smoking), and the percentage of individuals who are overweight, obese, or very obese. All regressions further control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and year. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Table 4: Regression results of workplace smoking policies on area human capital

	(1) **Any official smoking policy**	(2) **Smoking ban in all areas of the office**
Area characteristics		
% college graduates	0.00033	0.00181**
	(0.00048)	(0.00071)
Individual characteristics (ref. group: no high school)		
High school graduate	0.03908***	0.02555***
	(0.00382)	(0.00378)
Some college	0.06120***	0.04699***
	(0.00399)	(0.00449)
College graduate	0.08425***	0.08420***
	(0.00383)	(0.00533)
Post-graduate education	0.10847***	0.11909***
	(0.00450)	(0.00629)
Individual chars	Yes	Yes
State-year FE	Yes	Yes
Area chars	Yes	Yes
Area-level industry shares	Yes	Yes
Weighted obs.	200,075,093	180,130,632
Cells	192,286	173,071
Areas	297	297
R-squared	0.055	0.058
Dependent var. mean	0.908	0.852
% change from 10pp increase in % college grads	*0.4*	*2.1*

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the area level.

Note: All probit regressions pool data from the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015, use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and year. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Table 5: Regression results of beliefs about smoking on area human capital

	(1)	(2)	(3)	(4)	(5)	(6)
	Strongly agree or agree with the following statement					
	"Smoking by pregnant women is harmful for baby"	**"Someone else's smoke is harmful"**	**"Most lung cancer deaths are caused by smoking"**	**"Smoking should be banned in bars"**	**"Smoking should be banned in restaurants"**	**"Smoking should be banned in work areas"**
Area characteristics						
% college grads	0.0103*	0.00328	0.0112***	0.00381***	0.00313*	0.00285***
	(0.00537)	(0.00336)	(0.00412)	(0.00124)	(0.00162)	(0.00107)
Individual characteristics (ref. group: no high school)						
High school graduate	0.218***	0.182***	0.122***	-0.03241***	-0.00892	0.01466***
	(0.0455)	(0.0294)	(0.0413)	(0.00518)	(0.00581)	(0.00439)
Some college	0.500***	0.335***	0.234***	-0.01654**	0.03696***	0.06028***
	(0.0595)	(0.0345)	(0.0471)	(0.00664)	(0.00791)	(0.00654)
College graduate	0.597***	0.541***	0.387***	0.04482***	0.09166***	0.11960***
	(0.0721)	(0.0377)	(0.0591)	(0.00699)	(0.00991)	(0.00800)
Post-graduate	0.771***	0.695***	0.552***	0.08092***	0.13279***	0.15032***
	(0.0782)	(0.0422)	(0.0595)	(0.00728)	(0.01146)	(0.00965)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	No	No	No	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes
Dependent var. mean	0.922	0.836	0.740	0.451	0.607	0.775
% change from 10pp increase in % college graduates	*11.2%*	*3.9%*	*15.1%*	*8.5%*	*5.2%*	*3.7%*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: Outcomes were defined relative to no opinion, disagreeing, or strongly disagreeing. All probit regressions in columns 1-3 pool data from National Health Interview Survey (NHIS) from 1987, 1992, and 2000. All probit regressions in columns 4-6 pool data from the Tobacco Use Supplement in the Current Population Survey (CPS). All regressions use sampling weights and include controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, and year. We exclude individuals with missing education. % college graduates was measured at the county level using data from the immediately preceding census for the given year for NHIS regressions and from the immediately preceding census or 3-year ACS for the CPS regressions. We also include controls for area log density and log population, % Black, % Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix A: Data sources

Mortality data

Data type	Source	Years	Key variables
Mortality	Multiple Cause Mortality Files (MCMF)	1990, 2000, 2010	All-cause and cause-specific mortality, age, sex, race, education, county of residence
Population counts	U.S. Decennial Census (5% sample), American Community Survey (ACS)	1990 (Census), 2000 (Census), 2009-2011 (ACS)	

Data on non-fatal health outcome

Data type	Source	Years	Key variables
Disease prevalence	Health and Retirement Survey	1992-2008	Prevalence of cancer, lung disease, diabetes, heart disease, stroke
Self-reported health	Behavioral Risk Factor Surveillance System	1999-2001, 2009-2011	Good/very good/excellent self-reported health, number of days in poor mental or physical health

Data on health-related behaviors

Data type	Source	Years	Key variables
Smoking behavior	Behavioral Risk Factor Surveillance System	1999-2001, 2009-2011	Current former, never smoker
Smoking behavior	Current Population Survey	1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015	Current former, never smoker

Obesity and physical activity	Behavioral Risk Factor Surveillance System	1999-2001, 2009-2011	BMI, any physical activity in last 30 days

Area characteristics data

Data type	**Source**	**Years**	**Key variables**
Human capital	U.S. Decennial Census (5% sample) and ACS	1990 (Census), 2000 (Census), 2009-2011 (ACS)	% with at least a college degree
Demographics	U.S. Decennial Census (5% sample) and ACS	1990 (Census), 2000 (Census), 2009-2011 (ACS)	% Black, % Hispanic, industry shares
Demographics	Area Resource Files	1990, 2000, 2010	Population, land area
Healthcare	Area Resource Files	1990, 2000, 2010	Number of MDs per 1,000, number of hospital beds per 1,000
Healthcare	Dartmouth Health Atlas	2003-2015	% Medicare enrollees with annual ambulatory visit, % Medicare enrollees with mammogram
Homicides	Uniform Crime Reports	1989-1991, 1999-2001, 2009-2011	Homicide rate
Pollution	van Donkelaar (2019)	1999-2001, 2009-2011	PM-2.5
Pollution	Environmental Protection Agency (EPA)	1989-1991	PM-10
Hospital quality	Centers for Medicare and Medicaid Services (CMS) Hospital Compare Database	2003-2008	Process-of-care indicators for pneumonia, congestive heart failure, and acute myocardial infarction

Migration data

Cross-county migration	Health and Retirement Survey	1992-2008	Probability of changing county of residence in the next two years; difference in area human capital across counties of residence
Cross-county or cross-MSA migration	National Longitudinal Survey of Young Women and Men	1969-1971, 1976 (Young Men); 1968-2001 (Young Women)	Probability of changing county or SMSA of residence since last interview; moving to non-SMSA

Smoking regulations

Workplace smoking policies	Current Population Survey	1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015	Any official smoking policy in place for indoor workers, smoking ban in public or work areas at the workplace for individuals working in indoor workplace

Beliefs about smoking

Beliefs about harms of smoking	National Health Interview Survey	1987, 1992, 2000	Agreement with whether smoking is harmful for pregnant women, whether most lung cancer deaths are caused by smoking, whether someone else's smoke is harmful
Beliefs about smoking regulation	Current Population Survey	1995-1996, 1998-1999, 2001-2002, 2006-2007, 2014-2015	Whether smoking should be banned in bars, restaurants, workplaces

Appendix B: Definitions of mortality due to smoking-related, obesity-related, medically amenable, and external causes

Smoking-related

Malignant Neoplasms: of the Lip, Oral Cavity, Pharynx, Esophagus, Stomach, Pancreas, Larynx, Trachea, Lung, Bronchus, Cervix Uteri, Kidney and Renal Pelvis, Urinary Bladder, and Acute Myeloid Leukemia; Cardiovascular Diseases: Ischemic Heart Disease, Other Heart Disease, Cerebrovascular Disease, Atherosclerosis, Aortic Aneurysm, Other Arterial Disease; Respiratory Diseases: Pneumonia, Influenza, Bronchitis, Emphysema, Chronic Airway Obstruction.

Source: CDC's National Center for Chronic Disease Prevention and Health Promotion (2014).

Obesity-related

Coronary Heart Disease, Other Cardiovascular Diseases; Cancers of the Colon, Breast, Esophagus, Uterus, Ovaries, Kidney, and Pancreas; Diabetes, and Kidney Disease.

Source: Flegal et al. (2007).

Medically amenable

Intestinal Infections, Tuberculosis, Other Infections (Diphtheria, Tetanus, Septicaemia, Poliomyelitis), Whooping Cough, Measles; Malignant Neoplasms of: Colon and Rectum, Skin, Breast, Cervix Uteri, Uterus, Testis; Hodgkin's Disease, Leukaemia, Diseases of the Thyroid, Diabetes, Epilepsy, Chronic Rheumatic Heart Disease, Hypertensive Disease, Ischaemic Heart Disease (50% of all such deaths), Cerebrovascular Disease, All Respiratory Diseases, Peptic Ulcer, Appendicitis, Abdominal Hernia, Cholelithiasis and Cholecystitis, Nephritis and Nephrosis, Benign Prostatic Hyperplasia, Misadventures to Patients during Surgical and Medical Care, Maternal Death, Congenital Cardiovascular Anomalies, Perinatal Deaths (excl. stillbirths).

Source: Nolte & McKee (2008).

External causes

Accidents, Intentional Self-Harm, Assault, Events of Undetermined Intent, Legal Intervention, Operations of War and Their Sequelae, Complications of Medical and Surgical Care.

Source: ICD-10-CM Codes V01-Y9.

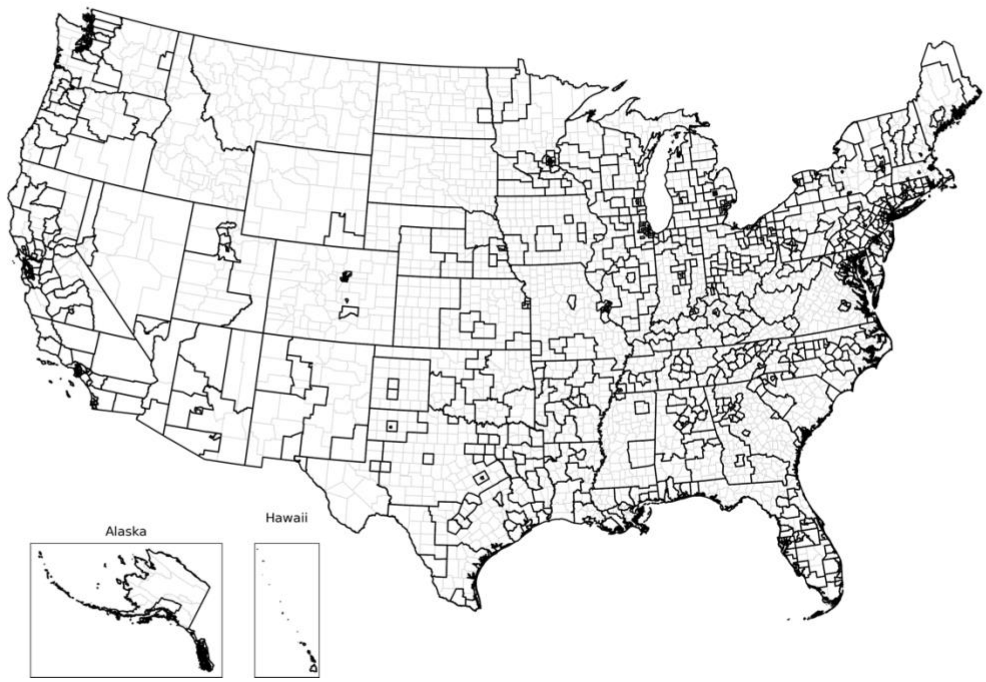
Drug poisoning (overdose)

Deaths from unintentional overdose of a drug, suicide, or drug poisoning of undetermined intent

Source: ICD-10-CM Codes X40–X44, X60–X64, or Y10–Y14.

Appendix C: Additional Tables and Figures

Appendix Figure C1: Map of CONSPUMAs



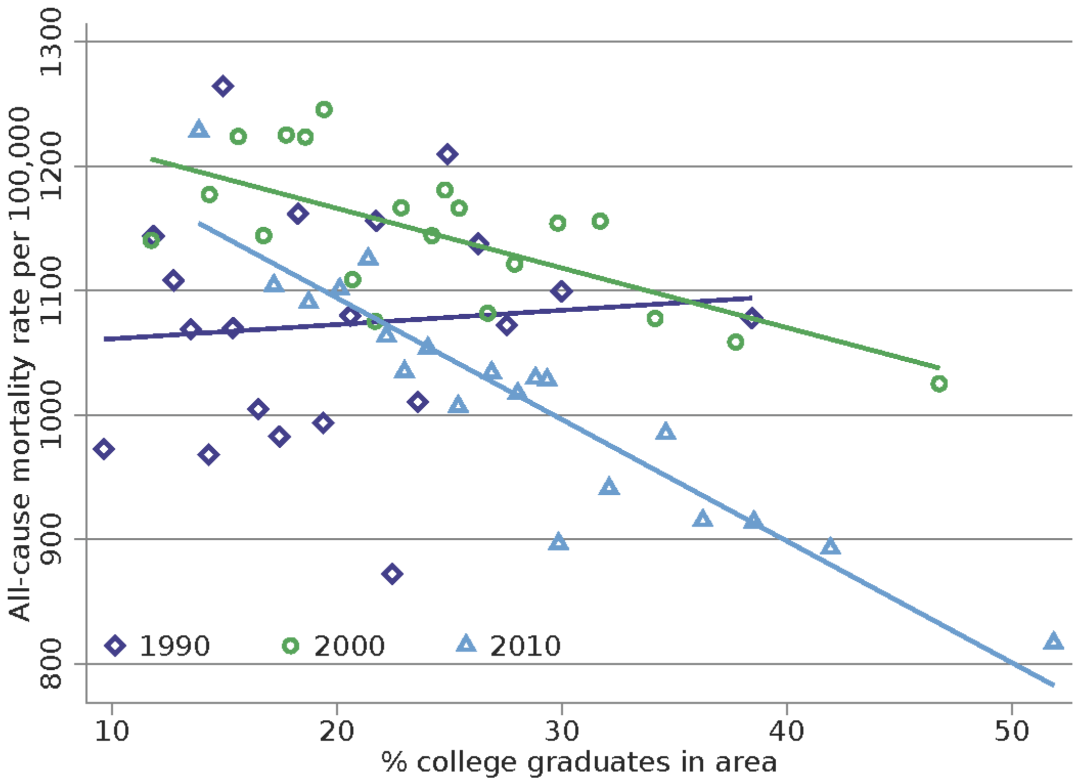
Note: All county-level data was aggregated to areas representing consistent public use microdata areas (CONSPUMAs). Counties are shown with gray borders and CONSPUMAs are shown with black borders.

Appendix Table C1: Descriptive statistics on mortality and area characteristics

	Mean	SD
Cell characteristics		
Age 25-64	80.1%	---
Age 65+	19.9%	---
Female	51.8%	---
White non-Hispanic	74.8%	
Black non-Hispanic	10.4%	
Hispanic	10.1%	
Other non-Hispanic	4.8%	
No high school	15.6%	---
High school graduate	36.9%	---
Some college	22.6%	---
College graduate	15.8%	---
Graduate education	9.0%	---
Missing education on death certificate	3.1%	4.3%
Mortality rates by cause (per 100,000)		
All cause	1,196	2,650
Heart disease	349	966
Cancer	289	518
Medically amenable causes	488	1,203
Smoking-related causes	672	1,653
Obesity-related causes	494	1,196
External causes	69	121
Area characteristics		
% college graduates	24.9%	8.8%
% Black	11.6%	11.4%
% Hispanic	13.2%	14.7%
Density (persons per square mile)	1,804	6,007
Population	1,878,063	2,051,136
Industry share: manufacturing	8.1%	3.6%
Number of observations		
Area-year-age-sex-race-education cells	369,707	---
Areas	486	---
Population	495,778,966	---
Deaths	5,934,489	---

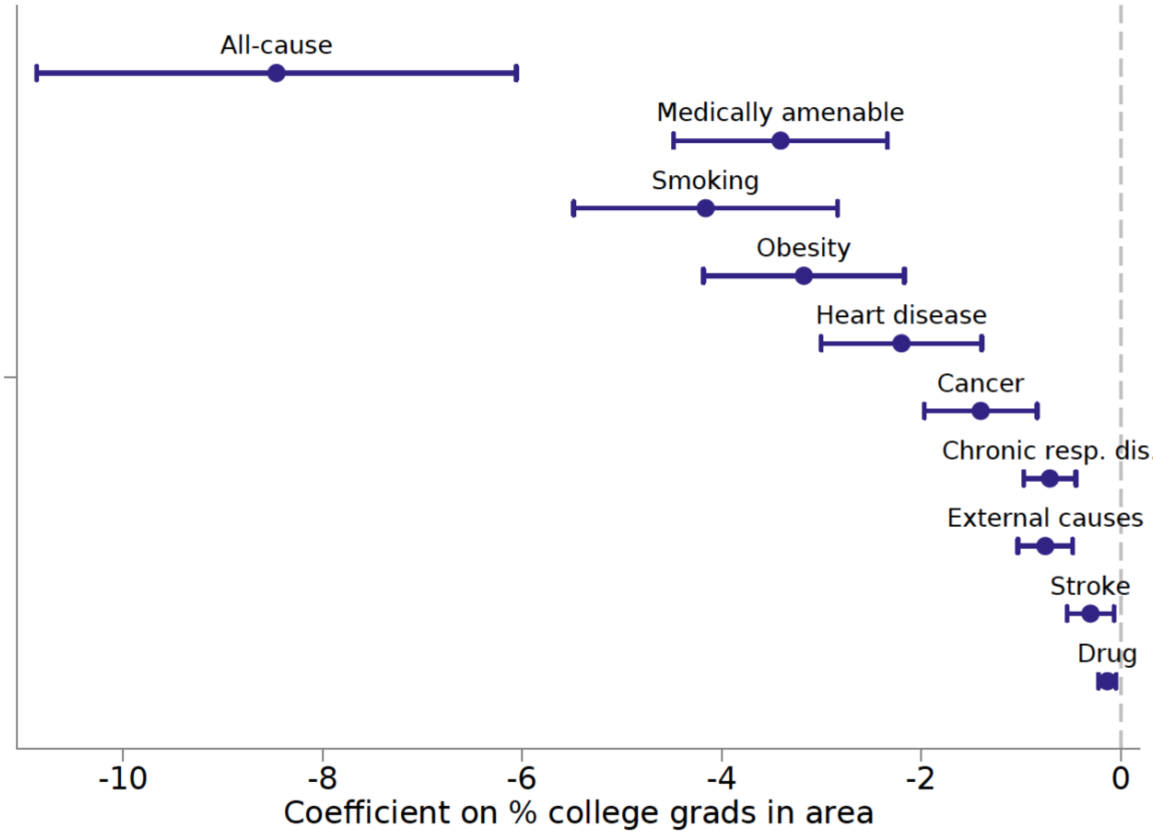
Note: Death data by county-year-age-sex-education comes from the 1990, 2000, and 2010 Multiple Cause Mortality Files. Counties were aggregated to areas representing consistent public use microdata areas (CONSPUMAs). Mortality rates were calculated using population sizes from the 1990 and 2000 Census 5% samples, and the 2009-2011 ACS 5-year file for 2010. We exclude county-year-age-race-sex cells where 25% or more of reported deaths lacked education data. Statistics are weighted by cell size.

Appendix Figure C2: The relationship between area human capital and adjusted mortality per 100,000



Note: This graph is a binned scatter plot showing all-cause mortality across ventiles of area education in each year for all data in our sample (excluding deaths with missing individual education). Each point includes approximately 5% of the population in that year, plotted at the mean percent college graduates and mean mortality rate (adjusted for age-sex using direct adjustment) across areas within each bin. The coefficients (and standard errors in parentheses) of the corresponding OLS regressions are -1.2 (2.2) in 1990, -4.8*** (0.98) in 2000, and -9.8*** (0.96) in 2010. Fitted lines extend to the full range of the underlying data in each year. Areas are defined as consistent public use microdata areas (CONSPUMAs). *** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Appendix Figure C3: Regression results of cause-specific mortality rates per 100,000 on area human capital



Note: This figure plots the coefficient on area human capital estimated separately for each cause of death. OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Smoking-related, medically amenable, and obesity-related causes of death include all deaths to causes associated with that risk factor and are not mutually exclusive categories (see Appendix B for details). Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Appendix Table C2: Regression results of cause-specific mortality rates per 100,000 on area human capital

	(1) All-cause	(2) Medically amenable	(3) Smoking	(4) Obesity	(5) Heart disease	(6) Cancer	(7) Chronic resp. dis.	(8) External causes	(9) Stroke	(10) Drug poisoning
Area characteristics										
% college graduates	-8.46***	-3.41***	-4.16***	-3.18***	-2.20***	-1.40***	-0.71***	-0.75***	-0.30**	-0.13***
	(1.22)	(0.55)	(0.67)	(0.51)	(0.41)	(0.29)	(0.13)	(0.14)	(0.12)	(0.04)
Cell characteristics (ref. group: no high school)										
High school graduate	-263.47***	-122.73***	-191.17***	-109.50***	-98.86***	-19.99**	-29.11***	-25.87***	-17.81***	-5.56***
	(40.54)	(18.03)	(26.28)	(17.82)	(13.75)	(8.95)	(3.07)	(1.90)	(2.87)	(0.43)
Some college	-446.51***	-188.44***	-298.10***	-179.06***	-152.48***	-64.15***	-42.07***	-50.50***	-22.18***	-9.75***
	(29.26)	(13.21)	(19.45)	(12.72)	(10.01)	(6.30)	(2.66)	(1.75)	(2.14)	(0.48)
College graduate	-537.61***	-220.75***	-343.36***	-204.21***	-168.76***	-82.50***	-52.75***	-65.79***	-23.09***	-15.53***
	(35.42)	(15.72)	(22.88)	(15.01)	(11.55)	(7.84)	(3.16)	(2.16)	(2.52)	(0.60)
Post-graduate education	-722.09***	-300.69***	-468.79***	-283.60***	-229.10***	-131.24***	-68.78***	-68.04***	-32.36***	-16.11***
	(36.69)	(16.37)	(23.72)	(15.34)	(11.74)	(8.24)	(3.44)	(2.21)	(2.58)	(0.60)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966
Cells	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707
Areas	486	486	486	486	486	486	486	486	486	486
R-squared	0.862	0.824	0.828	0.811	0.795	0.710	0.563	0.252	0.648	0.111
Dependent var. mean	1,196	488	672	494	349	289	62	69	77	9
% change from 10pp increase in % college grads	*-7.1*	*-7.0*	*-6.2*	*-6.4*	*-6.3*	*-4.9*	*-11.4*	*-10.9*	*-3.9*	*-14.8*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

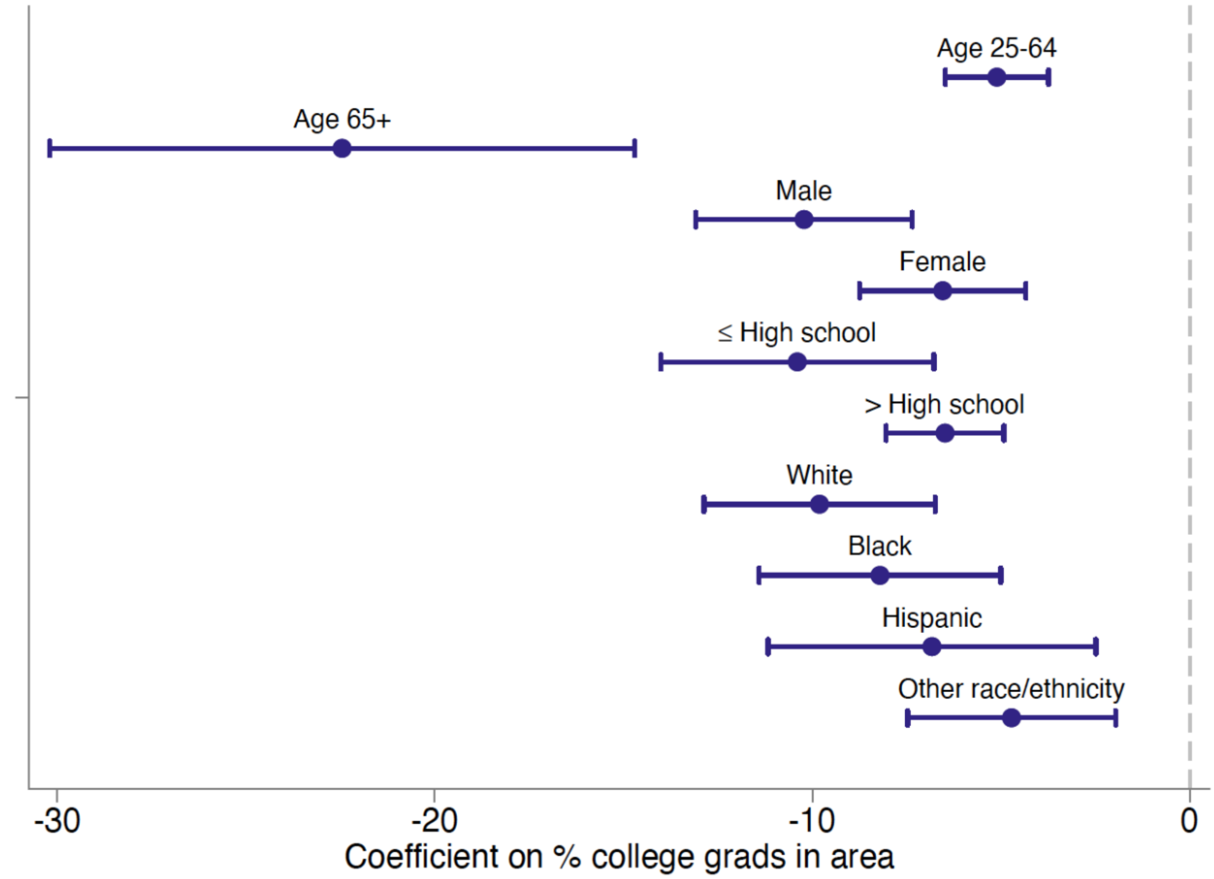
Appendix Table C3: Regression results of cause-specific mortality rates per 100,000 on area human capital by year

	(1) **All-cause**	(2) **Medically amenable**	(3) **Smoking**	(4) **Obesity**	(5) **Heart disease**	(6) **Cancer**	(7) **Chronic resp. dis.**	(8) **External causes**	(9) **Stroke**	(10) **Drug poisoning**
Area characteristics										
% college graduates	-6.00***	-2.27***	-4.03***	-3.15***	-2.76***	-0.86**	-0.21	-0.56***	-0.25	-0.05
	(1.48)	(0.66)	(0.85)	(0.63)	(0.51)	(0.36)	(0.16)	(0.17)	(0.15)	(0.05)
% college graduates *	-1.47**	-0.81***	0.04	-0.25	0.30	-0.47***	-0.30***	-0.10	0.02	0.00
year=2000	(0.65)	(0.31)	(0.42)	(0.33)	(0.28)	(0.16)	(0.07)	(0.10)	(0.09)	(0.03)
% college graduates *	-3.35***	-1.47***	-0.24	0.10	0.78**	-0.65***	-0.67***	-0.28**	-0.10	-0.14***
year=2010	(0.75)	(0.34)	(0.47)	(0.36)	(0.31)	(0.18)	(0.09)	(0.11)	(0.10)	(0.04)
Cell characteristics (ref. group: no high school)										
High school graduate	-263.81***	-122.88***	-191.19***	-109.49***	-98.78***	-20.06**	-29.17***	-25.90***	-17.82***	-5.57***
	(40.56)	(18.04)	(26.29)	(17.83)	(13.76)	(8.95)	(3.07)	(1.90)	(2.87)	(0.43)
Some college	-447.03***	-188.67***	-298.14***	-179.04***	-152.35***	-64.25***	-42.17***	-50.54***	-22.20***	-9.77***
	(29.30)	(13.23)	(19.47)	(12.74)	(10.02)	(6.31)	(2.66)	(1.76)	(2.14)	(0.48)
College graduate	-537.98***	-220.91***	-343.39***	-204.20***	-168.67***	-82.57***	-52.83***	-65.82***	-23.10***	-15.54***
	(35.44)	(15.73)	(22.89)	(15.02)	(11.56)	(7.85)	(3.17)	(2.16)	(2.52)	(0.60)
Post-graduate education	-722.38***	-300.82***	-468.80***	-283.60***	-229.03***	-131.30***	-68.84***	-68.06***	-32.37***	-16.13***
	(36.72)	(16.38)	(23.73)	(15.35)	(11.75)	(8.25)	(3.44)	(2.21)	(2.58)	(0.60)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966	495,778,966
Cells	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707	369,707
Areas	486	486	486	486	486	486	486	486	486	486
R-squared	0.862	0.824	0.828	0.811	0.795	0.710	0.563	0.252	0.648	0.111
Dependent var. mean	1,196	488	672	494	349	289	62	69	77	9

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Figure C4: Regression results of all-cause mortality rates per 100,000 on area human capital by demographic subgroups



Note: This figure plots the coefficient on area human capital estimated separately for each subgroup. OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Appendix Table C4: Regression results of mortality rates per 100,000 on area human capital by demographic subgroup

	(1) **Age 25-64**	(2) **Age 65+**	(3) **Male**	(4) **Female**	(5) **<= High school**	(6) **>= Some college**	(7) **White**	(8) **Black**	(9) **Hispanic**	(10) **Other race/ethnicity**
Area characteristics										
% college graduates	-5.11***	-22.46***	-10.22***	-6.54***	-10.40***	-6.48***	-9.81***	-8.21***	-6.83***	-4.72***
	(0.70)	(3.94)	(1.46)	(1.12)	(1.84)	(0.79)	(1.56)	(1.63)	(2.21)	(1.40)
Cell characteristics (ref. group: no high school)										
High school graduate	-210.80***	-228.68**	-332.57***	-192.10***	-221.45***	--	-416.12***	-108.01**	10.65	-76.25
	(16.20)	(105.39)	(41.54)	(39.93)	(40.88)	--	(49.36)	(46.61)	(28.15)	(46.38)
Some college	-347.65***	-613.23***	-615.91***	-286.73***	--	242.29***	-598.28***	-393.96***	-86.43***	-173.40***
	(14.62)	(57.13)	(30.96)	(28.65)	--	(9.19)	(36.02)	(31.54)	(15.25)	(44.77)
College graduate	-413.81***	-898.80***	-725.04***	-355.09***	--	148.01***	-699.62***	-470.64***	-109.05***	-159.81***
	(15.91)	(89.43)	(37.66)	(33.90)	--	(7.28)	(41.76)	(41.39)	(31.97)	(50.23)
Post-graduate education	-483.75***	-1613.11***	-942.33***	-495.44***	--	--	-891.86***	-726.68***	-230.46***	-203.69***
	(17.01)	(83.40)	(39.29)	(34.61)	--	--	(40.73)	(35.70)	(35.88)	(46.05)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	396,880,881	98,898,085	238,768,366	257,010,600	260,481,118	235,297,848	370,677,266	51,326,357	50,075,991	23,699,352
Cells	240,034	129,673	188,263	181,444	161,373	208,334	172,361	86,386	57,403	53,557
Areas	486	485	485	486	486	486	485	477	485	480
R-squared	0.626	0.793	0.842	0.882	0.882	0.811	0.887	0.716	0.675	0.582
Dependent var. mean	364	4,536	1,232	1,164	1,661	682	1,333	1,197	528	470
% change from 10pp increase in % college grads	*-14.0*	*-5.0*	*-8.3*	*-5.6*	*-6.3*	*-9.5*	*-7.4*	*-6.9*	*-12.9*	*-10.0*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

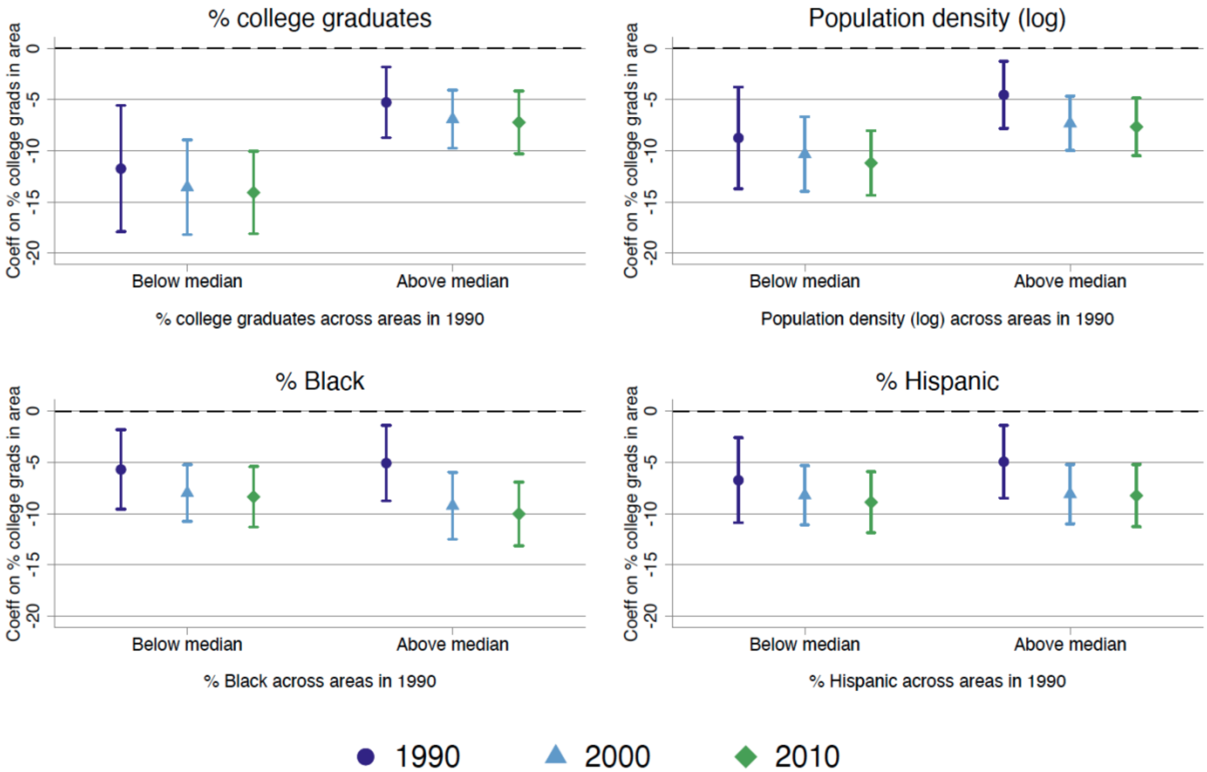
Appendix Table C5: Regression results of mortality per 100,000 on area human capital by demographic subgroup, by year

	(1) Age 25-64	(2) Age 65+	(3) Male	(4) Female	(5) <= High school	(6) >= Some college	(7) White	(8) Black	(9) Hispanic	(10) Other race/ethnicity
Area characteristics										
% college graduates	-3.09***	-17.05***	-8.16***	-3.54**	-5.93**	-3.51***	-6.75***	-6.44***	-11.19**	-5.08***
	(0.77)	(5.07)	(1.65)	(1.42)	(2.31)	(0.87)	(1.83)	(2.36)	(5.14)	(1.86)
% college graduates *	-1.42***	-0.10	-1.10	-1.94***	-1.75*	-3.10***	-1.49**	-1.39	4.07	1.88*
year=2000	(0.42)	(2.29)	(0.87)	(0.65)	(0.99)	(0.52)	(0.72)	(1.46)	(3.53)	(1.13)
% college graduates *	-2.65***	-8.63***	-2.87***	-3.99***	-6.66***	-3.29***	-4.50***	-2.00	4.46	-0.20
year=2010	(0.48)	(2.60)	(0.91)	(0.79)	(1.24)	(0.56)	(0.83)	(1.65)	(3.83)	(1.38)
Cell characteristics (ref. group: no high school)										
High school graduate	-211.00***	-229.91**	-332.84***	-192.53***	-221.62***	0.00	-416.87***	-108.11**	10.72	-76.08
	(16.22)	(105.47)	(41.56)	(39.96)	(40.89)	(.)	(49.40)	(46.62)	(28.14)	(46.43)
Some college	-348.01***	-614.63***	-616.28***	-287.44***	0.00	242.33***	-599.30***	-394.07***	-86.28***	-173.38***
	(14.64)	(57.21)	(30.98)	(28.69)	(.)	(9.19)	(36.07)	(31.56)	(15.23)	(44.81)
College graduate	-414.07***	-899.31***	-725.37***	-355.51***	0.00	147.99***	-700.31***	-470.66***	-108.95***	-159.74***
	(15.93)	(89.46)	(37.69)	(33.92)	(.)	(7.28)	(41.80)	(41.39)	(31.96)	(50.25)
Post-graduate	-483.98***	-1612.54***	-942.75***	-495.52***	0.00	0.00	-892.41***	-726.78***	-230.26***	-203.44***
education	(17.03)	(83.37)	(39.33)	(34.61)	(.)	(.)	(40.76)	(35.71)	(35.85)	(46.07)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	396,880,881	98,898,085	238,768,366	257,010,600	260,481,118	235,297,848	370,677,266	51,326,357	50,075,991	23,699,352
Cells	240,034	129,673	188,263	181,444	161,373	208,334	172,361	86,386	57,403	53,557
Areas	486	485	485	486	486	486	485	477	485	480
R-squared	0.626	0.793	0.842	0.882	0.882	0.811	0.887	0.716	0.675	0.582
Dependent var. mean	364	4,536	1,232	1,164	1,661	682	1,333	1,197	528	470

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are clustered at the area level.

Note: OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Figure C5: Regression results of all-cause mortality rates per 100,000 on area human capital by area characteristics in 1990



Note: This figure plots the coefficient on area human capital interacted by whether the area-year-age-sex-race-education cell is in an area above/below median percent college graduates, population density (log), percent Black, or percent Hispanic across areas in 1990, weighted by population. OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and state-year fixed effects. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Appendix Table C6: Regression results of non-fatal health outcomes on area human capital

	(1) **Cancer per 100,000**	(2) **Lung disease per 100,000**	(3) **Diabetes per 100,000**	(4) **Heart disease per 100,000**	(5) **Stroke per 100,000**	(6) **Good, very good, or excellent self-reported health**	(7) **Poor health, number of days in last 30 days**
	HRS 1992-2008	HRS 1992-2008	HRS 1992-2008	HRS 1992-2008	HRS 1992-2008	BRFSS 1999-2001 2009-2011	BRFSS 1999-2001 2009-2011
Area characteristics							
% college grads	6.5	-18.7**	-7.4	-37.8**	9.1	0.00215***	-0.03897***
	(11.5)	(8.2)	(10.5)	(16.0)	(9.1)	(0.00025)	(0.00517)
Cell characteristics (ref. group: no high school)							
High school graduate	-135.9	-883.2***	-1,051.5***	-1,080.2***	-461***	0.10657***	-2.16753***
	(132.6)	(140.6)	(181.1)	(220.1)	(102.2)	(0.00221)	(0.11493)
Some college	56.9	-980.1***	-1,185.6***	-959.4***	-625.9***	0.14587***	-2.54005***
	(135.5)	(158.6)	(213.0)	(282.1)	(130.4)	(0.00261)	(0.12252)
College graduate	-101.7	-1,676.3***	-1,843.4***	-2,372.7***	-879.8***	0.23313***	-3.91896***
	(162.7)	(152.3)	(195.7)	(256.7)	(118.0)	(0.00357)	(0.12271)
Observations	113,890	115,694	108,075	100,174	119,554	1,553,211	1,459,505
R-squared	0.0307	0.0228	0.0321	0.0579	0.0346	0.116	0.048
Dependent var. mean	2,800	2,200	3,300	6,200	1,700	0.834	3.774
% increase from 10pp increase in % college graduates	*2.3%*	*-8.5%*	*-2.2%*	*-6.1%*	*5.3%*	*2.6%*	*-10.3%*

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the area level.

Note: Cancer, lung disease, diabetes, heart disease, stroke, and hospitalizations in the Health and Retirement Survey (HRS) were defined as conditions or hospitalizations reported since the prior wave. Area characteristics in the HRS are measured at the time of HRS entry. OLS regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, Behavioral Risk Factor Surveillance System (BRFSS) state-year fixed effects or HRS wave fixed effects. We exclude individuals with missing education. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Table C7: Regression results of smoking on area human capital

	(1)	(2)	(3)	(4)	(5)	(6)
	CPS: Current smoker	**CPS: Former smoker**	**CPS: Never smoker**	**BRFSS: Current smoker**	**BRFSS: Former smoker**	**BRFSS: Never smoker**
Area characteristics						
% college graduates	-0.00223***	0.00329***	0.00177***	-0.00246***	0.00379***	0.00132***
	(0.00039)	(0.00063)	(0.00064)	(0.00029)	(0.00047)	(0.00046)
Individual characteristics (ref. group: no high school)						
High school graduate	-0.02866***	0.06475***	0.00256	-0.05600***	0.06444***	0.04979***
	(0.00461)	(0.00510)	(0.00518)	(0.00508)	(0.00520)	(0.00668)
Some college	-0.06789***	0.13527***	0.02177***	-0.10064***	0.12258***	0.08155***
	(0.00627)	(0.00543)	(0.00719)	(0.00573)	(0.00554)	(0.00757)
College graduate	-0.16886***	0.23568***	0.15031***	-0.23354***	0.24969***	0.24235***
	(0.01004)	(0.00678)	(0.01135)	(0.00802)	(0.00651)	(0.00952)
Post-graduate education	-0.22027***	0.30450***	0.19999***	--	--	--
	(0.01196)	(0.00943)	(0.01264)	--	--	--
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	613,858	238,830	613,858	1,551,524	738,024	1,551,524
Cells	613,858	238,830	613,858	1,551,524	738,024	1,551,524
Areas	297	297	297	484	484	484
R-squared	0.086	0.123	0.081	0.088	0.121	0.063
Dependent var. mean	0.165	0.561	0.624	0.192	0.585	0.538
% change from 10pp increase in % college grads	*-13.5*	*5.9*	*2.8*	*-12.8*	*6.5*	*2.5*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. All probit regressions pool data from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS) or the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015. All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Table C8: Regression results of obesity on area human capital

	(1) BRFSS: Very obese	(2) BRFSS: Obese	(3) BRFSS: Overweight	(4) BRFSS: No physical activity
Area characteristics				
% college graduates	-0.00113***	-0.00069***	0.00017	-0.00194***
	(0.00021)	(0.00023)	(0.00033)	(0.00032)
Individual characteristics (ref. group: no high school)				
High school graduate	-0.01824***	-0.01485***	0.01653***	-0.07400***
	(0.00172)	(0.00266)	(0.00357)	(0.00292)
Some college	-0.02380***	-0.01687***	0.01938***	-0.14805***
	(0.00170)	(0.00274)	(0.00347)	(0.00261)
College graduate	-0.06571***	-0.05732***	0.02175***	-0.24288***
	(0.00186)	(0.00320)	(0.00382)	(0.00378)
Individual chars	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes
Weighted obs.	1,490,722	1,490,722	1,490,722	1,452,041
Cells	1,490,722	1,490,722	1,490,722	1,452,041
Areas	484	484	484	484
R-squared	0.049	0.023	0.024	0.063
Dependent var. mean	0.091	0.168	0.378	0.258
% change from 10pp increase in % college grads	*-12.4*	*-4.1*	*0.5*	*-7.5*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: All probit regressions pool data from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS). All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Table C9: Regression results of smoking on area human capital

	(1) **CPS: Current smoker**	(2) **CPS: Former smoker**	(3) **CPS: Never smoker**	(4) **BRFSS: Current smoker**	(5) **BRFSS: Former smoker**	(6) **BRFSS: Never smoker**
Area characteristics						
% college graduates	-0.00204***	0.00282***	0.00157**	-0.00189***	0.00327***	0.00067
	(0.00046)	(0.00078)	(0.00072)	(0.00030)	(0.00052)	(0.00045)
% college graduates *	-0.00007	0.00054	-0.00013	--	--	--
year=2000	(0.00025)	(0.00048)	(0.00033)	--	--	--
% college graduates *	-0.00043*	0.00052	0.00072	-0.00085***	0.00077***	0.00094***
year=2010	(0.00026)	(0.00051)	(0.00047)	(0.00018)	(0.00029)	(0.00024)
Individual characteristics (ref. group: no high school)						
High school graduate	-0.02868***	0.06477***	0.00261	-0.05599***	0.06441***	0.04979***
	(0.00461)	(0.00510)	(0.00518)	(0.00508)	(0.00520)	(0.00668)
Some college	-0.06794***	0.13531***	0.02188***	-0.10064***	0.12257***	0.08156***
	(0.00628)	(0.00543)	(0.00720)	(0.00573)	(0.00554)	(0.00757)
College graduate	-0.16889***	0.23574***	0.15037***	-0.23351***	0.24964***	0.24232***
	(0.01004)	(0.00678)	(0.01136)	(0.00801)	(0.00651)	(0.00951)
Post-graduate education	-0.22031***	0.30456***	0.20004***	--	--	--
	(0.01197)	(0.00944)	(0.01265)	--	--	--
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	613,858	238,830	613,858	1,551,524	738,024	1,551,524
Cells	613,858	238,830	613,858	1,551,524	738,024	1,551,524
Areas	297	297	297	484	484	484
R-squared	0.086	0.123	0.081	0.088	0.121	0.063
Dependent var. mean	0.165	0.561	0.624	0.192	0.585	0.538

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. All probit regressions pool data from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS) or the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015. All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Table C10: Regression results of obesity on area human capital

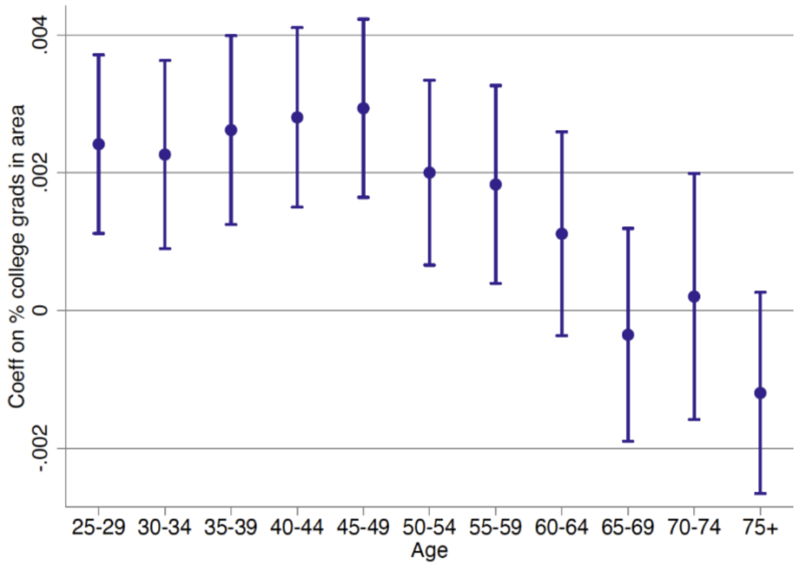
	(1) **BRFSS: Very obese**	(2) **BRFSS: Obese**	(3) **BRFSS: Overweight**	(4) **BRFSS: No physical activity**
Area characteristics				
% college graduates	-0.00087***	-0.00060**	-0.00058**	-0.00165***
	(0.00026)	(0.00030)	(0.00029)	(0.00033)
% college graduates * year=2000	--	--	--	--
	--	--	--	--
% college graduates * year=2010	-0.00034*	-0.00012	0.00108***	-0.00038
	(0.00020)	(0.00020)	(0.00025)	(0.00024)
Individual characteristics (ref. group: no high school)				
High school graduate	-0.01825***	-0.01485***	0.01654***	-0.07400***
	(0.00172)	(0.00266)	(0.00357)	(0.00292)
Some college	-0.02380***	-0.01687***	0.01940***	-0.14805***
	(0.00170)	(0.00274)	(0.00347)	(0.00261)
College graduate	-0.06569***	-0.05732***	0.02171***	-0.24287***
	(0.00185)	(0.00320)	(0.00382)	(0.00378)
Individual chars	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes
Weighted obs.	1,490,722	1,490,722	1,490,722	1,452,041
Cells	1,490,722	1,490,722	1,490,722	1,452,041
Areas	484	484	484	484
R-squared	0.049	0.023	0.024	0.063
Dependent var. mean	0.091	0.168	0.378	0.258

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

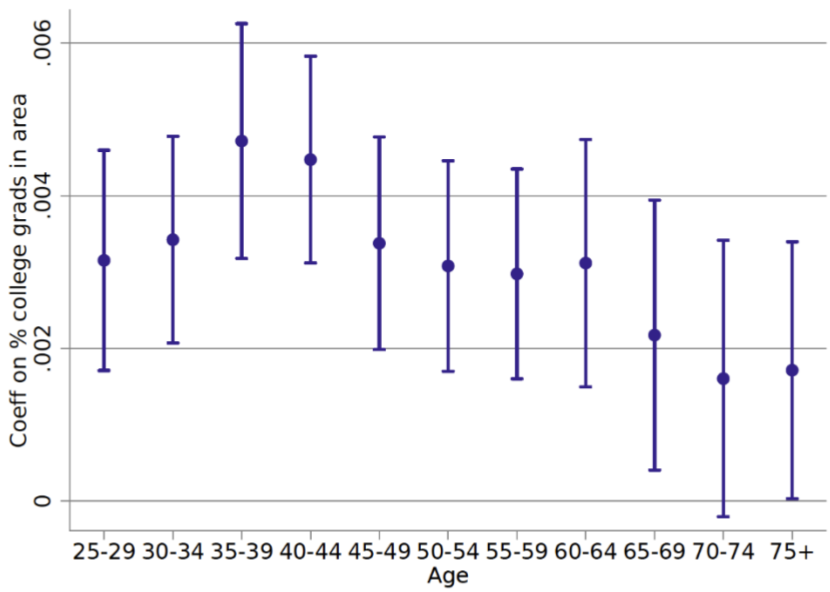
Note: All probit regressions pool data from the 1999-2001 and 2009-2011 Behavioral Risk Factor Surveillance System (BRFSS). All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Figure C6: Regression results of smoking on area human capital interacted with age

(a) Never smoker



(b) Former smoker



Note: This figure plots the coefficient on area human capital interacted by age and estimated separately for each smoking-related behavior, all of which are defined as binary variables. Former smoker is defined conditional on ever smoking. All probit regressions pool data from the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015. All regressions use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and state-year fixed effects. We exclude individuals with missing education. Area-level percent college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares, defined similarly as percent college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Appendix Table C11: Regression results of mortality per 100,000 by cause of death on area human capital and health-related behaviors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cancer		**Chronic lower respiratory disease**		**Stroke**		**Drug poisoning**	
Area characteristics								
% college graduates	-1.46***	-0.44	-0.69***	-0.10	-0.32**	-0.04	-0.15**	-0.04
	(0.28)	(0.29)	(0.15)	(0.14)	(0.13)	(0.16)	(0.06)	(0.07)
% current smoker		2.50***		0.88***		0.26		0.18
		(0.62)		(0.31)		(0.37)		(0.11)
% former smoker		32.23		-26.51		-17.30		-3.40
		(38.39)		(20.79)		(22.38)		(6.52)
% overweight, obese, very obese		0.73***		0.30*		0.25*		0.08
		(0.22)		(0.15)		(0.15)		(0.06)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	357,070,962	357,070,962	357,070,962	357,070,962	357,070,962	357,070,962	357,070,962	357,070,962
Cells	257,906	257,906	257,906	257,906	257,906	257,906	257,906	257,906
Areas	485	485	485	485	485	485	485	485
R-squared	0.749	0.749	0.614	0.614	0.672	0.672	0.110	0.111
Dependent var. mean	288	288	66	66	75	75	12	12
% change from 10pp increase in % college grads	-5.1	*-1.5*	-10.5	-1.5	-4.3	-0.6	-12.3	-3.0

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. OLS regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 2000 and 2010. All regressions control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and year. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. The percentage of current or former smokers was calculated using the 1995-1996, 1998-1999 CPS, 1999-2001 BRFSS, and 2009-2011 BRFSS. The percent of individuals that were overweight, obese, or very obese, and those with no physical activity were calculated using the 1999-2001 and 2009-2011 BRFSS. Areas are defined as consistent public use microdata areas (CONSPUMAs).

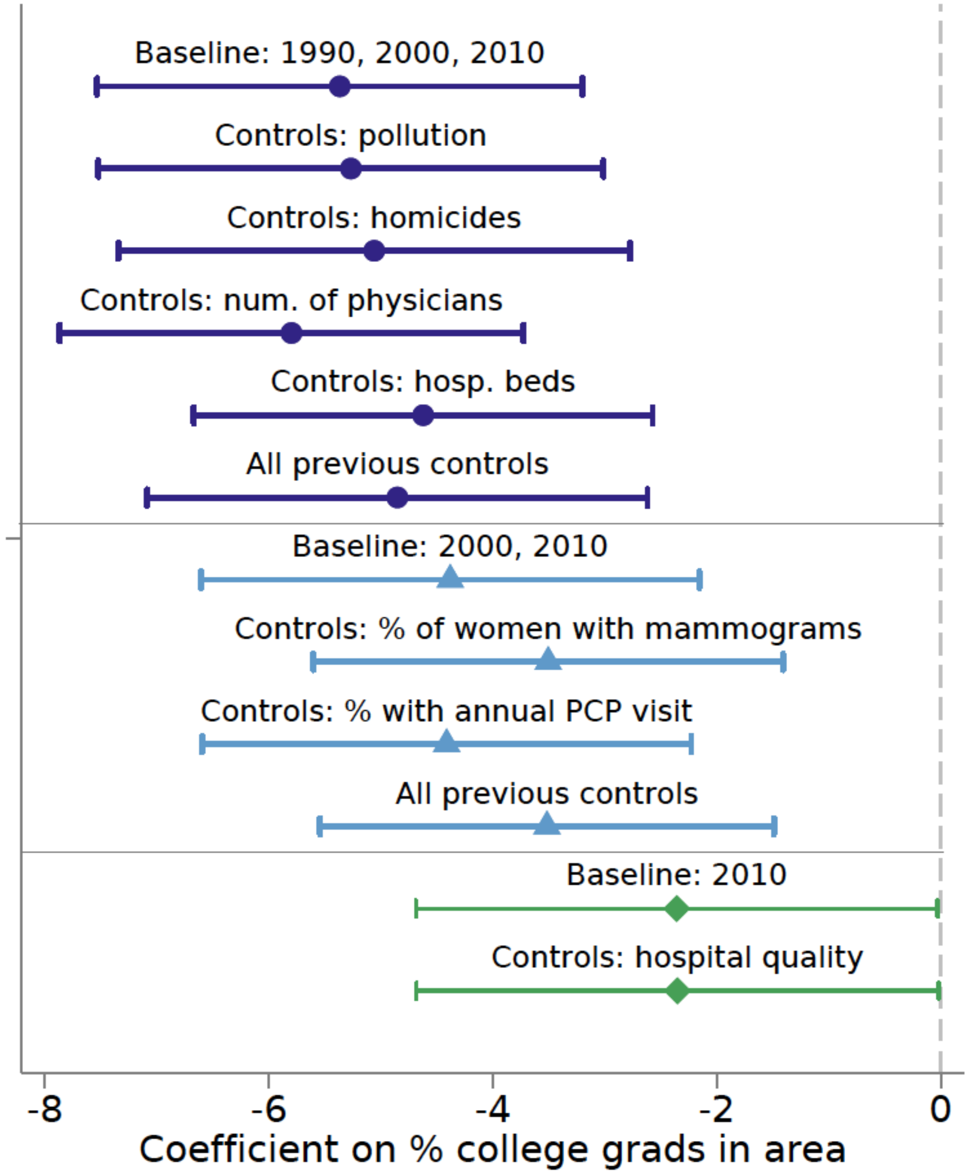
Appendix Table C12: Regression results of all-cause mortality per 100,000 on area human capital and health-related amenities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Area characteristics												
% college graduates	-3.48***	-3.41***	-3.41***	-3.86***	-2.61**	-3.05***	-2.03**	-3.08***	-2.05**	-0.65	-0.66	-0.66
	(1.13)	(1.15)	(1.16)	(1.05)	(1.02)	(1.08)	(1.02)	(1.06)	(0.97)	(1.18)	(1.18)	(1.18)
Air pollution (PM-2.5)		3.66				4.05	2.93	3.93	1.34	-1.92	-2.03	-2.03
		(4.27)				(3.96)	(4.08)	(4.01)	(4.05)	(5.27)	(5.33)	(5.33)
Homicide rate per 100,000			0.48			-0.55	-1.73	-0.52	-1.70	-1.61	-1.58	-1.58
			(1.67)			(1.64)	(1.62)	(1.63)	(1.46)	(3.31)	(3.29)	(3.29)
Physicians per 1,000				13.12***		8.77**	9.18**	8.89**	10.71***	10.52**	10.55**	10.55**
				(4.12)		(4.10)	(4.06)	(4.03)	(3.94)	(4.13)	(4.14)	(4.14)
Hospital beds per 1,000					12.10***	9.66***	10.40***	9.68***	10.84***	10.49**	10.34**	10.34**
					(3.14)	(3.30)	(3.40)	(3.29)	(3.31)	(4.10)	(4.17)	(4.17)
% with mammogram							-5.81***		-7.39***	-8.27***	-8.36***	-8.36***
							(1.25)		(1.27)	(1.16)	(1.21)	(1.21)
% with annual PCP visit								0.32	3.52***	2.22**	2.21**	2.21**
								(1.01)	(0.99)	(1.06)	(1.06)	(1.06)
Hospital quality (z-score)											4.21	4.21
											(11.82)	(11.82)
% currently smoking	6.76**	7.06***	6.71**	5.63**	5.34**	5.26**	5.59**	5.27**	5.77**	6.78***	6.84***	6.84***
	(2.70)	(2.69)	(2.75)	(2.61)	(2.47)	(2.44)	(2.41)	(2.44)	(2.34)	(2.44)	(2.44)	(2.44)
% formerly smoking	-205.93	-187.47	-205.84	-247.03	-244.69	-244.05	-189.96	-241.08	-142.74	-273.32*	-272.27	-272.27
	(176.18)	(172.34)	(177.15)	(168.95)	(166.43)	(159.08)	(153.37)	(161.27)	(148.80)	(165.76)	(166.00)	(166.00)
% overweight, obese, very obese	4.18***	4.01***	4.19***	4.67***	4.44***	4.52***	4.85***	4.51***	4.88***	7.13***	7.17***	7.17***
	(1.09)	(1.13)	(1.09)	(1.08)	(1.09)	(1.10)	(1.08)	(1.11)	(1.07)	(1.47)	(1.47)	(1.47)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area-level industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cells	257,598	257,598	257,598	257,598	257,598	257,598	257,598	257,598	257,598	145,531	145,531	145,531
Areas	485	485	485	485	485	485	485	485	485	485	485	485
R-squared	0.879	0.879	0.879	0.879	0.879	0.879	0.879	0.879	0.879	0.874	0.874	0.874
Dependent var. mean	1,206	1,206	1,206	1,206	1,206	1,206	1,206	1,206	1,206	1,175	1,175	1,175
% change from 10pp increase in % college grads	-2.9	-2.8	-2.8	-3.2	-2.2	-2.5	-1.7	-2.6	-1.7	-0.5	-0.6	-0.6

* p<0.1, ** p<0.05, *** p<0.01. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. All regressions are estimated at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. All regressions further control for cell-level 5-year age (25-29, 30-34, …, 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and year. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Areas are defined as consistent public use microdata areas (CONSPUMAs).

Appendix Figure C7: Regression results of all-cause mortality per 100,000 on area human capital and health-related amenities



Note: All regressions are estimated separately at the area-year-age-sex-race-education cell level, weighted by cell population, and pooled across 1990, 2000, and 2010. The main control in the baseline regression is the percent of individuals currently smoking and the percent of individuals formerly smoking (conditional on ever smoking). All regressions further control for cell-level 5-year age (25-29, 30-34, . . . , 85+) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic) interactions, individual education, percent of death certificates without education information, and year. We also include controls for area log density and log population, percent Black, percent Hispanic, and industry shares. Confidence intervals are clustered at the area level. Areas are defined as consistent public use microdata areas (CONSPUMAs). Confidence intervals are clustered at the area level.

Appendix Table C12: Regression results of smoking behavior on area human capital, workplace smoking policies, and smoking beliefs

	(1) Current smoker	(2) Current smoker	(3) Current smoker	(4) Former smoker	(5) Former smoker	(6) Former smoker	(7) Never smoker	(8) Never smoker	(9) Never smoker
Area characteristics									
% college graduates	-0.00218***	-0.00191***	-0.00186***	0.00407***	0.00344***	0.00339***	0.00132*	0.00103	0.00098
	(0.00050)	(0.00049)	(0.00049)	(0.00092)	(0.00089)	(0.00088)	(0.00068)	(0.00066)	(0.00066)
Smoking should be banned in bars, restaurants, and work		-0.19644***	-0.19467***		0.26890***	0.26686***		0.20864***	0.20697***
		(0.00695)	(0.00689)		(0.00855)	(0.00850)		(0.00497)	(0.00495)
Smoking ban in all areas of the office			-0.02910***			0.04168***			0.03257***
			(0.00327)			(0.00600)			(0.00438)
Individual chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area chars	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Area industry shares	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weighted obs.	123,182,251	123,182,251	123,182,251	123,182,251	123,182,251	123,182,251	123,182,251	123,182,251	123,182,251
Cells	120,508	120,508	120,508	120,627	120,627	120,627	120,627	120,627	120,627
Areas	297	297	297	297	297	297	297	297	297
R-squared	0.082	0.143	0.144	0.091	0.129	0.129	0.080	0.113	0.114
Dependent var. mean	0.168	0.168	0.168	0.549	0.549	0.549	0.629	0.629	0.629
% change from 10pp increase in % college grads	*-13.0*	*-11.4*	*-11.1*	*7.4*	*6.3*	*6.2*	*2.1*	*1.6*	*1.6*

*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered at the area level.

Note: Former smoker is defined conditional on ever smoking. All probit regressions pool data from the Tobacco Use Supplement in the Current Population Survey (CPS) from waves 1995-1996, 1998-1999, 2001-2002, 2003, 2006-2007, 2010-2011, and 2014-2015, use sampling weights and include individual-level controls for 5-year age (25-29, 30-34, …, 85+, missing) by sex by race (white non-Hispanic, Black non-Hispanic, other non-Hispanic, Hispanic, missing race/ethnicity) interactions, individual education, and year. We exclude individuals with missing education. Area-level % college graduates in each year was measured using data from the immediately preceding census or 3-year ACS. We also include controls for area log density and log population, % Black, % Hispanic, and industry shares, defined similarly as % college graduates. Areas are defined as consistent public use microdata areas (CONSPUMAs).