Published in final edited form as:
Soc Sci Med. 2012 June ; 74(12): 1900–1910. doi:10.1016/j.socscimed.2012.02.029.

Using quantile regression to examine the effects of inequality across the mortality distribution in the U.S. counties

Tse-Chuan Yang[a,*], **Vivian Yi-Ju Chen**[b], **Carla Shoff**[c], and **Stephen A. Matthews**[d]

[a]Population Research Institute, Social Science Research Institute, The Pennsylvania State University, United States

[b]Department of Statistics, Tamkang University, Taipei, Taiwan

[c]Department of Agricultural Economics and Rural Sociology, Population Research Institute, The Pennsylvania State University, United States

[d]Department of Sociology, Department of Anthropology, Population Research Institute, The Pennsylvania State University, United States

Abstract

The U.S. has experienced a resurgence of income inequality in the past decades. The evidence regarding the mortality implications of this phenomenon has been mixed. This study employs a rarely used method in mortality research, quantile regression (QR), to provide insight into the ongoing debate of whether income inequality is a determinant of mortality and to investigate the varying relationship between inequality and mortality throughout the mortality distribution. Analyzing a U.S. dataset where the five-year (1998–2002) average mortality rates were combined with other county-level covariates, we found that the association between inequality and mortality was not constant throughout the mortality distribution and the impact of inequality on mortality steadily increased until the 80th percentile. When accounting for all potential confounders, inequality was significantly and positively related to mortality; however, this inequality–mortality relationship did not hold across the mortality distribution. A series of Wald tests confirmed this varying inequality–mortality relationship, especially between the lower and upper tails. The large variation in the estimated coefficients of the Gini index suggested that inequality had the greatest influence on those counties with a mortality rate of roughly 9.95 deaths per 1000 population (80th percentile) compared to any other counties. Furthermore, our results suggest that the traditional analytic methods that focus on mean or median value of the dependent variable can be, at most, applied to a narrow 20 percent of observations. This study demonstrates the value of QR. Our findings provide some insight as to why the existing evidence for the inequality–mortality relationship is mixed and suggest that analytical issues may play a role in clarifying whether inequality is a robust determinant of population health.

Keywords

Quantile regression; Income inequality; Mortality; County; USA

*Corresponding author. Tel.: +1 814 865 5553. tuy111@psu.edu (T.-C. Yang).

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.socscimed.2012.02.029.

Introduction

Since the 1970s, several developed countries have witnessed a resurgence of income inequality (Alderson & Nielsen, 2002). Income inequality (hereafter inequality) has been the most striking in the U.S., where inequality increased by nearly 20 percent from 1970 to 2008 (DeNavas-Walt, Proctor, & Smith, 2009; Moller, Nielsen, & Alderson, 2009). It has been suggested that this phenomenon is a function of economic development, educational expansion, racial/ethnic composition, and urbanization, as well as political and institutional factors (Moller et al., 2009). Researchers have also paid attention to the consequences and implications of inequality. For example, Gilbert (2008) systematically discussed the changes in social class structure in the U.S. and emphasized potential social problems that emerge during the "age of inequality."

In the health literature, two plausible pathways have been proposed to theorize how and why inequality affects population health. First, higher inequality is more likely to increase the sense of relative deprivation and relative disadvantage (Marmot, 2004; Wilkinson, 2006). For instance, the poor in developed countries may have electricity, water, as well as some household durables, but their poverty status is determined through comparison with other residents in their country. While the absolute level of living standards is crucial to human health, it does not inform our understanding of why inequality is negatively related to human health in developed countries (Wilkinson, 2006). It is suggested that higher inequality within a society will cause a stronger sense of relative deprivation, resulting in stress, frustration, depression, anxiety, hostility, and other mental pressures (Marmot, 2004). All of these psychosocial characteristics can be translated into high-risk behaviors such as smoking, excessive alcohol consumption, and antisocial behavior (Kawachi, Levine, Miller, Lasch, & Amick, 1994; Wilkinson, 1997). The negative health consequences of relative inequality are not only observed at the bottom of the social hierarchy, but also across it including at the top (Lynch & Kaplan, 1997).

The second pathway between inequality and health is related to the underinvestment in physical, cultural, and civic resources (Daly, Duncan, Kaplan, & Lynch, 1998; Kawachi, Kennedy, Lochner, & Prothrow-Stith, 1997; Lynch & Kaplan, 1997). Inequality may lead to limited public services, poor infrastructure, and weaker social conditions, as stated in the "fundamental causes" hypothesis (Link & Phelan, 1995). An area with less inequality is expected to be one where people enjoy more equal opportunities and access to education, housing, medical care, and a range of social services. In a study by Kaplan, Pamuk, Lynch, Cohen, and Balfour (1996), the authors found that among the 50 U.S. states, a more equal income distribution indicated higher educational expenditure per capita, more library books per capita, fewer people without health insurance coverage, a lower rate of violent crimes, and a smaller proportion of people in jail. Indeed, investment in social programs that provide access to public services is a form of wealth redistribution, which may mitigate relative deprivation. Furthermore, underinvestment in society undermines self-esteem, mutual respect, trust, and confidence. In addition, friendships, social trust, reciprocity, and social capital are difficult to establish in an environment of high inequality (Kawachi et al., 1997; Marmot, 2004; Wilkinson, 2006).

The theoretical pathways identified above suggest that inequality is a relative concept with an independent association with health. Since the 1990s, many county- and state-level analyses have been conducted in the U.S. (Lynch et al., 2004; Wilkinson & Pickett, 2006); however, the findings are far from conclusive. Deaton (2003) suggested that "it is *not* true that income inequality itself is a major determinant of population health… and the correlation across the states and cities of the United States is almost certainly the result of something that is correlated with income inequality, but that is not income inequality itself

(p.151).” This debate is ongoing and warrants a detailed investigation of “something that is correlated with inequality (p.151),” such as racial composition and absolute living standards (Deaton, 2003).

As we explain in this paper, the inconsistent findings may be, in part, the result of methodological shortcomings related to levels/ units of analysis and the number of observations (e.g., an *N* of 50 for state-level studies). In turn, these shortcomings have limited both the range and diversity of variables included in the models and the sophistication of the analytical techniques used. To address these issues, we use quantile regression (QR) to analyze data on 3072 counties in the contiguous U.S. (forty-eight states plus the District of Columbia) based on data from around the year 2000 (see later discussion).

Our paper proceeds as follows. First, we provide a review of prior ecological studies of the associations between inequality and mortality in the U.S. In this review, we both expand and build upon our discussion of methodological constraints. Our next section introduces quantile regression (QR) as a method that utilizes information across the entire distribution of the outcome variable (i.e., mortality). QR is a well-known statistical approach (Koenker & Bassett, 1978), but one that is rarely used in mortality research. We follow this with a discussion of our data and methods. As noted, our analysis is based on U.S. county-level data. We measure health using all-cause standardized mortality rates, a widely used health indicator in the field of inequality-health research that facilitates a comparison of our findings and those of earlier studies. Our county-level covariates include a more comprehensive set of predictors than most other research on inequality and health including inequality, racial/ethnic composition, rurality, socioeconomic status (SES), and measures of social capital. These data and the use of QR allow us to address three specific substantive questions regarding the relationship between inequality and mortality. (1) Is inequality significantly associated with mortality after controlling for the socioeconomic confounders found in the literature? (2) If yes, is this association constant throughout the distribution of mortality? Or does inequality have a greater influence in counties with higher mortality rates compared to those with lower mortality rates? And (3) if not, how does the relationship between inequality and mortality vary by the levels of mortality? We present our results and close with a discussion of our findings, policy implications, the limitations of our study, and the value of using QR in social science research.

Inequality and mortality in the U.S

The association between inequality and mortality in the U.S. has drawn much attention in recent decades. After reviewing nearly 100 articles, Lynch et al. (2004) concluded that little support was found for the idea that inequality is a “major and generalizable” concept accounting for the population health differences within or between developed countries. However, the most consistent effect of inequality on population health was identified in the U.S. (Lynch et al., 2004), where most of these studies used mortality to measure population health. Wilkinson and Pickett (2006) further extended the scope of the review to over 150 published papers. The authors noted that over 80 percent of the analyses where nation states were the analytic unit confirmed the association between inequality and mortality; however, in studies using small areas (e.g., neighborhood and county or parish) as the analytic unit, less than fifty percent reported a significant relationship. Similarly, Wilkinson and Pickett (2006, 2009) not only found that the studies in the U.S. offered more consistent findings than other countries, but also documented a distinction by ecological level. Specifically, they noted that state-level research provided more coherent evidence of a positive inequality–mortality association compared to county- or tract-level studies. One plausible explanation for this pattern is that the ecological studies that use large areas better capture social

stratification (Wilkinson & Pickett, 2006, 2009). We note that these systematic reviews did not suggest that nations or states are the appropriate analytic units, but the mixed findings by unit choice indicate more research is necessary—primarily on larger samples and using a range of statistical techniques.

State-level analyses of inequality and mortality

Several state-level U.S. studies published in the late 1990s investigated the association between wealth (and income) distribution and mortality. Kaplan et al. (1996) defined the degree of income inequality as the percentage of total income received by the households in the bottom 50 percent of the wealth distribution. Based on the results of the analysis, a more even distribution of wealth was associated with lower mortality rates. In addition, the authors further linked their inequality measure to multiple outcomes including age-specific mortality, low birth weight rates, homicides, violent crimes, work disability, expenditures on medical care, police protection, smoking, and sedentary activity. These outcomes are all negatively associated with their measure of income inequality (Kaplan et al., 1996).

Kennedy, Kawachi, and Prothrow-Stith (1996) measured income inequality with the Gini coefficient and the Robin Hood index. Although the Gini coefficient is the most popular measure of income inequality, the study found little association with different cause-specific mortality rates. However, the Robin Hood Index told a different story. This index measures the proportion of aggregate income that needs to be transferred from households above the mean to those households below the mean in order to reach an even distribution. Higher values of the Robin Hood Index indicate stronger inequality in a society. Age-adjusted mortality, infant mortality, coronary heart disease, cancer, and homicides were all positively associated with the Robin Hood Index. Even after adjustment for absolute living standards, the significant associations remained (Kennedy et al., 1996). The insignificant relationships between the Gini coefficient and many health outcomes in this study led to an investigation of whether the measure of inequality matters. Linking six indicators of inequality to state-level mortality, it was suggested that the choice of indicator did not alter the positive association between inequality and mortality (Kawachi & Kennedy, 1997). Since then, many state-level analyses have reported similar findings (Kawachi et al., 1997; Shi, Starfield, Kennedy, & Kawachi, 1999; Wilkinson, Kawachi, & Kennedy, 1998; Wolfson et al., 1999).

However, some state-level studies have argued that inequality was associated with racial composition, education, poverty, and rurality and that these covariates were mostly ignored in the earlier studies. Moreover, several researchers found that inequality was not related to mortality after accounting for these potential confounders (Deaton & Lubotsky, 2003; Laporte, 2002; Mellor & Milyo, 2001; Muller, 2002). Multivariate analytic techniques were used to consider the effects of confounders and inequality on mortality instead of simply calculating bivariate correlations of inequality with mortality, which is the approach that was commonly used in research in the late 1990s and has since been criticized for the lack of statistical robustness (Deaton, 2003). These later studies also illustrate how different methodological approaches applied to the study of inequality and mortality can generate new findings, which has contributed to the mixed results observed at the state-level.

County-level analyses of inequality and mortality

State-level analyses are subject to the small sample problem (maximum $n = 50$); hence, complex multivariate analyses are unsuitable. County-level research does not face such a constraint. For instance, using Poisson regression and mortality data between 1985 and 1994, Massing et al. (2004) found that cardiovascular disease mortality rates were closely related to income inequality among county population aged 35 to 74. Similarly, Shi et al. (2005) concluded that counties with high income inequality experienced more than a 10

percent increase in all-cause, heart disease, and cancer mortality compared to their counterparts with low inequality. Nevertheless, findings at the county-level are also mixed and seemingly dependent on more local social factors (e.g., population size) than those at the state-level. A study in Texas found that the impact of inequality on mortality was significant only among the counties with a population over 150,000 (Franzini, Ribble, & Spears, 2001). A study conducting stratified regression analyses on all 100 North Carolina counties found that all-cause mortality was significantly related to inequality, even after controlling for per capita income; however, this significant relationship depended upon rurality, a dimension not incorporated into state-level analyses (Brodish, Massing, & Tyroler, 2000).

Varying relationship of inequality with mortality

Most of the previous research has concentrated on the issue of how different levels of inequality are associated with mortality (McLaughlin & Stokes, 2002; McLaughlin, Stokes, & Nonoyama, 2001), and relatively few studies have explicitly investigated the question of whether the effect of inequality varies by different levels of mortality. Notwithstanding, some findings from the literature inform and guide our study. James and Cossman (2006) explored the effects of inequality on three different types (and levels) of standardized mortality rates across different regions in the U.S. While inequality was not a significant predictor in the final model, the results implied that the same one unit change in the Gini index would lead to a larger increase in mortality for those counties with higher mortality rates. The results also showed that this trend holds true across regions. Another study found that the same level of the Gini index was differently associated with mortality, with higher average mortality rates associated with a stronger impact of inequality in general (Cossman, Cossman, Cosby, & Reavis, 2008). It should be noted that these studies were not designed to directly examine the varying relationship of inequality with mortality, but their findings provide auxiliary evidence that the effect of inequality on mortality increases with mortality rates and does not need to be constant across the mortality distribution. Again, the findings on the varying relationship between mortality and inequality are circumstantial and our study is among the first to explicitly test this issue.

The U.S. studies discussed above are important for at least three reasons. First, inequality is intertwined with many aspects of our society, including socioeconomic conditions, race/ethnicity, and factors such as rurality. Thus, these potential confounders need to be considered if we are to better understand the association between inequality and mortality. As the major source of endogeneity is likely an omitted variable simultaneously related to both inequality and mortality, including a range of independent covariates may minimize endogeneity. Second, there is a lack of consistent evidence supporting the argument that population health is better where there is low inequality and this is especially the case in county-level studies (Lynch et al., 2004; Wilkinson & Pickett, 2006). Third, while recent county-level analyses imply that the effect of inequality on mortality differs by the level of mortality (Cossman et al., 2008; James & Cossman, 2006), the findings are far from conclusive. The analytic methods used in previous studies cannot portray the varying association. Building upon past literature, the goal of this study is to answer the three interrelated research questions stated in the introduction and to explore the potentially varying relationship between inequality and mortality.

Analytic approach

Two analytical approaches dominated the aforementioned ecological studies. One is to calculate the bivariate correlations between inequality and mortality rates and the other is to compute the effects of inequality on the “mean” function of the conditional distribution of mortality by minimizing the sum of squared errors (also known as ordinary least squares [OLS] regression). The former has been criticized for ignoring potential confounders and the

latter is unable to answer our three research questions. That is, while statistically the mean value represents the central tendency of a distribution, little information is known about the entire distribution. The conventional OLS approach only estimates models for conditional mean functions of mortality and fails to consider other locations throughout the mortality distribution.

By contrast, QR is a non-parametric approach that was developed to estimate a full range of conditional quantile functions (also referred to as percentile functions in some literature) by minimizing asymmetrically weighted absolute errors (Koenker & Bassett, 1978). QR thus complements and ameliorates the OLS regression approach. To compute the covariance and correlation matrices of the parameter estimates, both bootstrap and asymptotic approaches are often used in QR. There are three advantages of using QR. QR (1) precisely portrays the stochastic associations between random variables; (2) provides robust estimates when the dependent variable is not normally distributed, and (3) minimizes the influence of data outliers (Koenker & Bassett, 1978). These methodological merits allow the associations of the independent covariates with the response variable to vary by the location, scale, and shape of the response distribution. QR has drawn the most attention from economists (Fitzenberger, Koenker, & Machado, 2002) and from statistical theorists (Buhai, 2005). Other social scientists have been unusually slow to explore the advantages QR may bring to mortality research. The literature on inequality and mortality leads us to believe that QR has the potential to be advantageous in this area.

Given a set of explanatory variables, we use QR to estimate mortality rates conditional on the selected quantile functions (e.g., 5th, 25th, 75th, and 95th). More specifically, QR allows us to evaluate the associations of the predictors with mortality on different segments of the conditional mortality distribution. By estimating the model at a series of segments (defined by the researchers), we can depict a more complete picture of how the associations vary throughout the analyzed distribution. The QR approach is also statistically more efficient and less affected by outliers than the conventional approaches in which researchers have used several dummy variables to capture the potential varying association between inequality and mortality. In the next section, we provide a more formal overview of the QR method and an explanation of how to interpret the estimated coefficients.

Fundamentals of QR

The classic QR model, first introduced by Koenker and Bassett (1978), could be regarded as an extension of the OLS model. Specifically, OLS only estimates how the predictor variables are related to the mean value of the dependent variable; QR, however, allows researchers to model the predictors against different locations/measurements of the dependent variable. Statistically, consider a random variable Y (mortality in this study) with a probability distribution function that can be expressed as (Buhai, 2005, p.3; Koenker & Hallock, 2001, p.145):

$$F(y)=\text{Prob}(Y \leq y) \quad (1)$$

the τ th quantile of Y can be defined as:

$$Q(\tau)=\inf\{y:F(y) \geq \tau\} \quad (2)$$

where $0 < \tau < 1$. Several notable locations are the first quartile, $Q(0.25)$, the median, $Q(0.5)$, the third quartile $Q(0.75)$, and the first and last deciles Q(0.1) and Q(0.9), respectively. Researchers can specify any values of τ to implement regressions. In our empirical mortality data, the τ th quantile, $\xi(\tau)$, could be expressed as the solution of the optimization problem (without any covariates):

$$\min_{\xi \in R} \sum_{i=1}^{n} \rho_\tau (y_i - \xi) \quad (3)$$

where $\rho_\tau(z) = z(\tau - I(z < 0))$. $I(.)$ represents the usual indicator function. When the explanatory variables (x) are included, the τ quantile, $\xi(\tau)$, can be rewritten as $x_i' \beta$ and the linear conditional quantile function becomes $Q_Y(\tau|X=x)=x_i' \beta_\tau$. The estimated coefficients of the explanatory variables can then be obtained by solving the function:

$$\widehat{\beta}(\tau)=\underset{\xi \in R}{\operatorname{argmin}} \sum_{i=1}^{n} \rho_\tau (y_i - x_i' \beta) \quad (4)$$

The interpretation of the estimates is similar to, but slightly different from that of OLS (Buhai, 2005; Koenker & Hallock, 2001). In OLS, the coefficient of a specific predictor, X, represents the expected change in the dependent variable that is associated with a unit change in X. However, the coefficient of X in the τ th quantile can be interpreted as the marginal change (relative to the value of the τ th quantile of the dependent variable) due to a one unit change in X. As τ can be specified as any value between 0 and 1, the estimated coefficients can be many, but we report those at commonly used quantiles (such as 0.05, 0.5, and 0.95).

In addition, we implement the Wald tests for equality of coefficients across quantiles in order to understand whether the differences in estimates are statistically significant. Specifically, for a pair of coefficients, say β_j^p and β_j^q, corresponding to the jth covariate at quantiles p and q, the following Wald statistics are used:

$$W^* = \frac{\widehat{\beta}_j^p - \widehat{\beta}_j^q}{Var\left(\widehat{\beta}_j^p - \widehat{\beta}_j^q\right)} \quad (5)$$

and $Var(\widehat{\beta}_j^p - \widehat{\beta}_j^q)=Var(\widehat{\beta}_j^p)+Var(\widehat{\beta}_j^q) - 2Cov(\widehat{\beta}_j^p, \widehat{\beta}_j^q)$ where $Var(\widehat{\beta}_j^p)$, $Var(\widehat{\beta}_j^q)$, and $Cov(\widehat{\beta}_j^p - \widehat{\beta}_j^q)$ are the estimated variances and covariance for $\widehat{\beta}_j^p$ and $\widehat{\beta}_j^q$ obtained via bootstrapping. Equation (5) can be extended to test the equality of multiple coefficients across quantiles (Hao & Naiman, 2007, pp.49–50). Note that bootstrapping method is used for the QR analysis because it has been argued to provide more suitable estimations for standard errors than the asymptotic method (Hao & Naiman, 2007, pp:43–76). Detailed discussions of the QR methodology and its applications in the social sciences can be found elsewhere (Fitzenberger et al., 2002; Hao & Naiman, 2007).

Data and measures

This is a county-level analysis of all counties in the continental U.S. and the District of Columbia. There are three reasons why the county could be an appropriate unit of analysis for mortality research. First, the administrative hierarchy capable of carrying out effective policy solutions is made up of federal, state, and county governments; however, since the 1980s, there has been a decentralization of governing responsibilities and public service activities from higher to lower levels of government (Lobao & Kraybill, 2005). County is the smallest analytic unit with useful policy implications (Allen, 2001) and policy decisions rarely extend to units below the county-level. Second, data from the U.S. decennial census and other governmental agencies such as the National Center for Health Statistics are more

readily available at the county-level. Third, using counties as the unit of analysis allows all geographic areas in the U.S. to be included, even the most remote rural areas and the largest cities, which helps the examination of the mortality rates across diverse places (McLaughlin, Stokes, Smith, & Nonoyama, 2007). Furthermore, county boundaries remain relatively stable over time, which makes it easier to replicate this study in the future. Finally, counties, for a variety of reasons, represent more appropriate units of analysis for our research questions than smaller units (e.g., ZIP codes or census tracts) as they better capture people's daily activities, their immediate functional ties (Matthews, 2011), and thus, an array of social and economic ties. According to a recent report, in the U.S., the average trip distance to work, work related business, shopping, personal/family errands, school, religious and recreational activities is approximately 10–20 miles (USDT, 2009), which is much larger than the average ZIP codes or census tracts, but well below the average county size. While the distributions of sizes are highly skewed, the average area of counties, ZIP codes, and tracts in the contiguous U.S. is 925 square miles (roughly equal to a 30 × 30 mile cell), 92 square miles (a 10 × 10 mile cell), and 46 square miles (a 7 × 7 mile cell), respectively.

The county-level mortality rate serves as the dependent variable in this study. Drawing upon past literature, we categorize our independent variables into five groups: inequality, racial/ethnic composition, rurality, socioeconomic status (SES), and social capital. Our analytic strategy is to implement five nested QR models where the five sets of variables are added sequentially enabling us to examine the changes in the association between inequality and mortality and determine whether the inequality–mortality associations holds even after controlling for all the other factors. Due to space limitations, only the estimates of the full model will be presented.

Mortality

Two *Compressed Mortality Files* (CMF), 1989–1998 and 1999–2003, from the National Center for Health Statistics (NCHS) need to be used in order to calculate five-year (1998–2002) mortality rates (NCHS, 2003, 2006) standardized by the 2000 U.S. age–sex population structure. The mortality rates are not standardized by race. We control for racial/ethnic covariates as a separate category in the analysis (Deaton & Lubotsky, 2003). As our standardized all-cause mortality rates are close to a normal distribution, we use the Gaussian assumption in our QR analysis. While it would be possible to analyze the number of deaths with Poisson regression, the discrepancy in the age–sex structure across counties would not be controlled for as well as it is with standardization.

Inequality

Several theoretical and methodological differences exist between the most common measures of inequality–the Gini index, the coefficient of variation, and the Theil's index (Allison, 1978), but their adverse impact on health and mortality are unanimously acknowledged (Kawachi & Kennedy, 1997). Due to its popularity in government reports and elsewhere, we used the Gini index to measure the distribution of household income (*U.S. Census, 2000*). The Gini coefficient is defined as a ratio with values between 0 and 1. The numerator is the area between the Lorenz Curve of the household income distribution and the uniform distribution line and the denominator is the area under the uniform distribution line. That is, a smaller Gini coefficient indicates a more even distribution of household income. The Gini index is calculated by multiplying the Gini coefficient by 100, and hence, values can range from 0 (perfect income equity) to 100 (extreme income inequality).

Racial/ethnic composition

Three racial/ethnic groups are included in the analysis: the county-level percentage Hispanic, percentage black, and percentage other races. To avoid collinearity, we do not

consider the percentage of non-Hispanic white population in the QR models. These variables are extracted from the *2000 Census* data. Controlling racial/ethnic composition is necessary and facilitates comparison of our findings with previous research.

Rurality

There is no agreement on how to measure rurality. We follow recent research that focused on multiple aspects of rurality (Brodish et al., 2000; McLaughlin et al., 2001, 2007). Here, rurality is measured by six variables, all derived from the *2000 Census*. Principal components analysis indicated that these variables summarized three dimensions of rurality: industrial structure, an ecological dimension (population), and exogenous economic integration. We calculated the factor scores with the regression method and used them as indicators of rurality in the analysis.

The first dimension, industrial structure, is comprised of one variable: percentage of the population ages 16 and over employed in farming, forestry, and fishing (factor loading is 0.93). The ecological dimension includes three variables related to the total population of a county: population density (the total population divided by land area, (0.93)), road density (the length of major roads per squared kilometer, (0.80)), and the percentage of workers commuting by public transportation (0.95). The third dimension of rurality is exogenous economic integration (EEI), which measures the economic influence of neighboring areas. Two variables are identified to capture this: percentage of workers traveling over an hour to work (0.87), and percentage of workers who work outside their county of residence (0.82). A more integrated county is expected to have a higher score of EEI.

SES

Following Sampson, Raudenbush, and Earls (1997), we use principal components analysis to capture the SES characteristics of a county. Specifically, we extracted eight variables from the *2000 U.S. Census* and from these derived two factors, social affluence and concentrated disadvantage. Social affluence is comprised of the following four variables: log of per capita income (factor loading is 0.88), percentage of the population age 25 years or older with a bachelor's degree or higher (0.93), percentage of the population employed in professional, administrative, and managerial positions (0.78), and the percentage of families with incomes over 75,000 dollars (0.92). Almost 80 percent of the variance among these variables is explained. Concentrated disadvantage includes the remaining four variables: poverty rate (0.89), the percentage of persons receiving public assistance (0.85), the unemployment rate (0.87), and the percentage of female-headed families with children (0.78). Principal component analysis indicates that one factor could explain 72 percent of the common variance.

Social capital

Following the underinvestment explanation for why inequality is associated with mortality, the social environment (e.g., mutual trust, safety, and reciprocity) must be considered if one wants to explore the net effect of inequality on mortality. Kawachi et al. (1997) suggested that social capital was related to both inequality and mortality and deserved further investigation. Rupasingha, Goetz, and Freshwater (2006) pulled together numerous indicators of social capital (e.g., Census response rate and the density of non-profit organizations) and created a social capital index for all U.S. counties. We followed previous research (Yang & McManus, 2010; Yang, Teng, & Haran, 2009) to include this social capital index and we also consider county-level safety and residential stability. Safety is a factor score drawn on a variety of crimes (e.g., embezzlement and robbery) that were calculated from the 1998–2002 *Uniform Crime Reports*. This covariate is used to reflect the absence of mutual trust and the sense of safety (and thus, weaker social capital). Social

capital was reported to be higher among homeowners (Glaeser, Laibson, & Sacerdote, 2000), indicating that a stable environment is good for the growth of social capital. In our dataset, two *2000 U.S. Census* variables were obtained to create a residential stability index: the percentage of the population living at the same address in 1995 and the percentage of owner-occupied housing units. These variables were standardized and averaged to create our residential stability index. Note that this study is focused on the inequality–mortality relationship. We refer readers to recent papers where the impacts of social capital on mortality are discussed in greater detail (Song, Son, & Lin, 2010; Yang, Jensen, & Haran, 2011).

Results and discussion

Descriptive analysis results

The descriptive statistics of the variables used in this study are shown in Table 1. The average age–sex standardized mortality rate was 8.90 per 1000 population in U.S. counties. While the mortality variation was not large (standard deviation was only 1.38), the range was wide. The Gini index at the county-level was 43 and the maximal inequality score was 61. However, the small value of the standard deviation (3.75) suggests that most counties had a Gini coefficient close to the mean. In 2000, on average, the counties' population was 8.72 percent Black and 6.16 percent Hispanic. Then again, as documented elsewhere (Jensen & Yang, 2009) and as indicated by the large standard deviations in Table 1, the geographical distribution of minorities, especially those of Hispanic-origin, varies greatly across the U.S. Since other socioeconomic and social capital variables were derived from principal components analysis, they have a mean close to 0 and a variance of 1. We examined multicollinearity with the variance inflation factors (VIF). The VIF is an index that measures the severity of multicollinearity among the independent variables. The square root of the VIF for a certain covariate can be interpreted as the extent to which the standard error is inflated due to the correlations with other independent variables. Note that the VIF is only concerned with the correlations among explanatory covariates; therefore, it is easily applied to multiple regression analyses (Mansfield & Helms, 1982). The rule of thumb is that a VIF greater than 10 would be problematic (Kutner, Nachtsheim, & Neter, 2004). The largest VIF (4.29) in Table 1 was well below this cut-off point, suggesting that multicollinearity would not increase the standard errors of our model estimates.

QR modeling results for the inequality–mortality relationship

While we implemented different models to explore how the association between inequality and mortality varies with the inclusion of nested sets of explanatory variables, our discussion focuses on the final model where all explanatory covariates were included. To better explain our findings, we provided the percentile-mortality distribution in Table 2. For example, the mortality rate at the 20th percentile was 7.81, indicating that 20 percent of the counties have a mortality rate lower than 7.81.

Table 3 displays the inequality–mortality relationship by 19 different quantiles. There are several important observations to note. First, the magnitude of the association between inequality and mortality varied greatly across quantiles. Specifically, the estimated coefficient was smaller than 0.01 for those counties below the 10th percentile of mortality. However, the inequality–mortality relationship more than doubled from the 10th to the 15th percentile and increased steadily until the 80th percentile where the estimated coefficient was 0.05. After which, the association between the Gini index and mortality decreased to roughly the same level at the 50th percentile. In contrast to the weakest inequality–mortality relationship (0.003) at the 10th percentile, the strongest association was 17 times larger (0.05 at the 80th percentile). This variability of the magnitude of the inequality–mortality

relationship has not been reported in the previous literature and suggests that close attention needs to be paid to the entire mortality distribution.

Second, inequality was not a significant explanatory variable for those counties with a mortality rate lower than 7.36 deaths per 1000 population (the 10th percentile). This is an important finding. It confirmed our hypothesis that the inequality–mortality relationship was not constant across the mortality distribution at the U.S. county-level (Cossman et al., 2008; McLaughlin & Stokes, 2002). After considering all other confounders, the QR results suggest that inequality is only relevant to mortality for counties with a mortality rate over a certain threshold. In addition, this varying inequality–mortality association across the mortality distribution challenges the findings based on more traditional analytic approaches. For example, OLS regression aims to model the conditional mean value of the dependent variable, but it is well-known that mean values are easily affected by extreme cases, which may seriously bias the estimates and undermine the conclusions.

We used the Wald tests via the bootstrapping approach in STATA 11.0 to examine the equality of coefficients across the quantiles of mortality to further confirm the pattern in the inequality–mortality relationship found above. We looked at the results from four tests. The first three tests compared the current estimated coefficient (say, at the 5th percentile) against three other estimates: at the median, at the corresponding position in the opposite tail (i.e., the 95th percentile), and at the adjacent higher percentile (i.e., the 10th percentile). The fourth test examined whether more than two out of the four quantile coefficients above were jointly the same. As the test for the corresponding position was included, we only presented the tests that focus on one half of the distribution (Hao & Naiman, 2007). The *p*-values of these tests were shown in Table 4.

According to the Wald tests, the estimated coefficients below the 20th percentile were significantly lower than the estimate at the median, providing evidence for an increasing trend from the lower to the higher percentiles. Moreover, the tests for the corresponding positions in the higher tail suggested that the inequality–mortality relationship was only comparable between the 40th and 60th percentile and it was significantly different between the lower and higher tails. The Wald test provides evidence of the appropriateness of QR in our study as the results from the traditional analytic methods (focusing on mean or median) can be, at most, applied to a narrow middle range of observations.

The variation of the inequality–mortality relationship across the mortality distribution was identified by the Wald test. It is not surprising that the only significant Wald test fell between the 10th and the 15th percentile. As discussed earlier, the inequality–mortality relationship was not significant in the lowest 10 percentiles of the mortality distribution. The non-significant result between the 5th and the 10th percentile (p-value = 0.54) suggests that the estimated inequality–mortality relationship was comparable (stable) in the low tail, but the apparent difference in the Gini index coefficients was statistically significant above the 15th percentile. One plausible explanation for the non-significant result at the lower tails may be endogeneity between the dependent and independent variables that is not captured by our independent variables. Also it should be noted that the estimations at the lower quantiles may be less precise due to larger standard errors, which may contribute to the non-significant inequality–mortality relationship. Finally, the last column of Table 4 provided evidence to reject the null hypothesis that more than two quantile coefficients for the Gini index were jointly the same, especially the lower 10 percentiles. We note that the Wald test via bootstrapping is more conservative than other common approaches, e.g., studentized range test (Hao & Naiman, 2007). Therefore, we used a larger p-value threshold in Table 4 than is typically used in these other approaches.

QR modeling results for other covariates

Having identified the varying association between inequality and mortality across the mortality distribution, we now focus on the associations of the other covariates in our model with mortality. In Table 5 we listed QR parameter estimates at seven selected percentiles and summarized the key findings below (a graphical representation of the QR estimates is included as an online Supplementary file).

First, following Koenker and Hallock (2001), the intercept may be interpreted as the estimated conditional quantile function of the mortality distribution of a county when the other variables are zero. It should be noted that most of our independent variables (with the exception of inequality and the racial/ethnic compositions) were centered by the mean values at the county-level, which is the preferable treatment of the independent variables (Koenker & Hallock, 2001). The primary reason why the Gini index and the racial/ethnic composition variables were not centered or standardized is to allow for the comparisons of our results with previous studies where racial groups were found to account for the effect of the Gini index on mortality (Deaton, 2003; Deaton & Lubotsky, 2003). Although the intercept seemed to have an increasing trend (Table 5), particularly at the higher end of the distribution, the Wald tests suggest that the difference in the intercept across quantiles was not statistically significant (results not shown).

Second, with respect to the racial/ethnic composition of counties, the relationship for the percentage Black was relatively stable across the distribution. The Wald tests did not find any significant discrepancy across quantiles, suggesting a constant relationship between the percentage Black and mortality. As has been suggested in the literature, counties with a higher percentage Black population tend to have higher mortality rates than those counties with a lower percentage of Black residents. This association was confirmed and found to be unrelated to the levels of mortality. While some have argued that the percentage Black was highly correlated with inequality and that by controlling for this factor it would make inequality an insignificant predictor of mortality (Deaton & Lubotsky, 2003; Mellor & Milyo, 2001), our analysis did not support this conclusion. In addition, although some researchers find that Hispanics and African Americans have comparable socioeconomic statuses, published research suggests that being Hispanic is negatively associated with mortality (the Hispanic Paradox) (Abraido-Lanza, Dohrenwend, Ng-Mak, & Turner, 1999). Our results not only confirmed this finding, but also demonstrate that the association of the percentage Hispanic on mortality was stronger at the lower quantile. Specifically, the tests for equality across quantiles indicated that the association of the percentage Hispanic with mortality almost doubled for those counties with a mortality rate around 6.95 deaths per 1000 population (the 5th percentile) in contrast to those counties with mortality rates around 8.85 deaths per 1000 population (the 50th percentile).

Third, exogenous economic integration (EEI) was positively associated with mortality, but this relationship was only significant for those counties above the threshold of 7.57 deaths per 1000 population (the 15th percentile). The association of EEI with mortality increased between the 15th percentile to the median and this upward trend was statistically significant based on the equality tests. In other words, being more economically integrated with neighboring counties was associated with higher mortality. Industrial structure was negatively associated with mortality, with the effect being strengthened until the 30th percentile and declining through the 90th percentile (10.56 deaths per 1000 population), above which it was not significant. Importantly, our findings match those of several recent studies by rural sociologists (McLaughlin et al., 2001, 2007).

With respect to the SES variables, social affluence and concentrated disadvantage had the expected relationships with mortality. Their point estimates at different percentiles were

significant across the entire mortality distribution with relatively little variation. All of the Wald tests were not significant, which indicates that the impact of the SES variables on mortality do not vary by the levels of mortality. The QR results not only suggest that a superior socioeconomic status in a county was related to lower mortality even after controlling for other factors, but also supported the argument that social conditions are the fundamental determinants of health (Link & Phelan, 1995).

As expected, residential stability, safety, and the social capital index were found to be significant predictors of county mortality rates. The social capital index demonstrates a declining profile (based on the absolute values of the estimated coefficients). That is, the counties in the lower 20th percentiles of the mortality distribution benefitted from social capital more than those counties in the upper half of the mortality distribution. The Wald tests for the social capital index, again, confirmed this finding (test results not shown). However, the associations of residential stability and safety with mortality were constant across the mortality distribution. That is, the Wald tests did not find any apparent differences in the estimated coefficients for both variables, but pointed out the universal importance of building a stable and safe neighborhood.

Finally, we observed that the magnitude of the impact of inequality on mortality was relatively minor in contrast to other the covariates. Specifically, Table 5 showed that a 10-point increase in the Gini coefficient can raise the county mortality rate by about 0.3–0.5 (deaths per 1000 population) in most quantiles. However, the same level of increase in rurality, SES, or social capital variables (e.g., safety and social affluence) may lead to a change in mortality by 1.5–5 deaths per 1000 population across quantiles. As discussed in Introduction, the two pathways from inequality to population health–relative deprivation and underinvestment–are “indirect” and subject to other social factors such as poverty and social support. The finding here seems to echo the theoretical frameworks and the results of recent county-level studies that used advanced analytic methods (Sparks & Sparks, 2010; Yang et al., 2011).

Several modeling issues should be noted. First, due to the space constraint, we did not include the Wald tests for the equality of coefficients across the mortality distribution for the explanatory variables (available upon request). Note that, as with inequality, the Wald tests for the other covariates confirmed the findings shown in Table 5. Second, we examined whether there are any interaction effects among the independent variables and whether there are any non-linear associations of the explanatory covariates with mortality. As the results did not alter our findings and conclusions, we only focus on the main effects for model parsimony. Third, we also used the asymptotic approach to estimate the QR coefficients and found that the results are comparable to those when using the bootstrap approach. These different model specifications and estimations strengthen the validity and reliability of our findings.

Conclusion

The QR methodology provides an empirical basis for answering the three research questions proposed in this study. First, in general, our results indicated that inequality was significantly and positively related to mortality, even after controlling for racial/ethnic composition, rurality, SES, and social capital. The QR analyses further demonstrated that this adverse impact of inequality on mortality was not significant until the 15th percentiles. The tests for equality further bolstered this finding and may suggest that inequality is irrelevant to mortality when county-level mortality is low, specifically lower than 7.4 deaths per 1000 population. These findings from QR have not been reported previously and may provide new motivation for future research to explore whether inequality is a determinant of

mortality (Lynch et al., 2004; Wilkinson & Pickett, 2006). The global one-model-fits-all approach (i.e., OLS regression) conventionally used in the field ignores information from the entire mortality distribution and as we have shown is unable to provide a complete picture of the inequality–mortality relationship.

Second, we found that the association between inequality and mortality was not constant throughout the mortality distribution. When accounting for all potential confounders, the impact of inequality on mortality steadily increased until the 80th percentile. A series of Wald tests confirmed that the inequality–mortality relationship differed significantly between the lower and upper tails. The large variation in the estimated coefficients of the Gini index suggested that inequality had the greatest influence on those counties with a mortality rate of roughly 9.95 deaths per 1000 population (80th percentile, see Table 2) compared to any other counties. While some earlier findings implied that the impact of inequality on mortality increases with mortality (Cossman et al., 2008; James & Cossman, 2006), these studies did not specifically report how the association varied across the mortality distribution. Our QR results in Tables 3 and 4 test and reveal the varying inequality–mortality relationship across the mortality distribution in U.S. counties. This addresses our third research question. That is, QR, a rarely used method in mortality research, suggests the variability in the inequality–mortality relationship.

Several policy implications can be drawn from our findings. Our QR results suggest that the associations of the explanatory variables with mortality are not universal. This has implications for how health and public policy makers may think about inequality and its association with other social factors contributes to mortality differentials across counties. While our results suggest that a one unit decrease in the Gini index would lead to larger improvement in mortality among the counties at the high percentiles than their counterparts below the 15th percentiles, it is not a trivial task to reduce inequality within a county. However, as noted in the previous section, other social variables play a larger role than inequality and it would be more realistic to take both social affluence and disadvantage into account, along with inequality. For example, improving other social conditions via subsidization and promoting public services might be considered as inequality is associated with social disadvantage (Pearson correlation coefficient = 0.61) and affluence (−0.22). Policies aimed at reducing poverty and unemployment may directly affect inequality, and they also may help minimize mortality. Similarly, the significant association of social capital and mortality suggested that developing a safe, stable, and cohesive community would likely see improvements in both mortality health disparities and social inequality. As the health inequality literature has suggested (Kawachi et al., 1997; Link & Phelan, 1995; Wilkinson et al., 1998), social conditions, including inequality, are the fundamental determinants of population health and the policy implications drawn from this study should simultaneously consider these factors.

This study has several limitations. First, an ecological approach implies that our findings should not be generalized to the individual-level (Piantadosi, Byar, & Green, 1988) and changing the spatial scale and unit of analysis may also lead to different conclusions (Openshaw, 1984). The former is known as the ecological fallacy and the latter refers to the modifiable areal unit problem. These are two issues shared by any study using aggregate data. Second, while our inequality measure is one of the most commonly used, it emphasizes the household income distribution within a county, which may be confounded by other social subsidies, especially in areas with a greater proportion of disadvantaged populations. More importantly, income inequality is just one dimension of inequality and it may not be the only component of relative deprivation. For instance, land inequality and political inequality may also have implications for human health and well-being (Deaton, 2003). It should also be noted that while we endeavored to minimize endogeneity by considering

many potential confounders in the analysis, the possibility of endogeneity is not fully eliminated and future research may use the instrumental variable to address this issue. Third, our study is cross-sectional and focuses only on the U.S., a country where the association of inequality within the country and population health has been extensively studied. While it has been argued that cross-sectional analysis may be more appropriate than the analyses using the changes in mortality and inequality (Kawachi & Blakely, 2001), more efforts should be made to investigate whether the varying inequality–mortality relationship is subject to temporal lags. Similar research could also be extended to other countries, specifically those with detailed sub-national health, demographic, and socioeconomic data.

Beyond the limitations specific to this study, some economists have argued that several issues in the literature may falsely lead to the conclusion that inequality is a determinant of population health (Mellor & Milyo, 2001; 2002a). These criticisms concentrate on whether the use of a standardized mortality rate is appropriate, whether the factors highly correlated with inequality, such as income and education, should be included in an ecological analysis, and also whether the temporal dimension should be considered in the inequality–mortality research. While Kawachi and Blakely (2001) have addressed these issues in detail, the debate is ongoing (Kawachi & Blakely, 2002; Mellor & Milyo, 2002b). The take-home message from the debate is for researchers to pay attention to data, measurement, and analytical issues in the continued investigation of the relationship between population health consequences and economic disparities (House, 2001). Although the focus of our study was not to directly address the previously raised issues, we have joined the debate by introducing a rarely used tool, quantile regression, to investigate the inequality–mortality relationship and we believe that we have offered new findings not been previously revealed using more conventional data and analytical methods.

Our research was motivated by a desire to better understand the relationship between inequality and mortality. The most significant contribution of our work is to reveal that the association between the social condition variables and mortality varies across counties depending upon where they are located in the mortality distribution. The significantly varying pattern across quantiles underscores the fact that previous studies overlooked the inherent heterogeneity in the inequality–mortality association. Methodologically we introduce quantile regression to health-mortality research and our empirical analysis provides a useful demonstration of how a QR approach can help mortality researchers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was partially supported by internal funds from the Social Science Research Institute at Penn State. Additional support has been provided by the Geographic Information Analysis Core at Penn State's Population Research Institute, which receives core funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R24-HD41025).

References

Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Turner JB. The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. American Journal of Public Health. 1999; 89(10):1543–1548. [PubMed: 10511837]

Alderson AS, Nielsen F. Globalization and the great U-turn: income inequality trends in 16 OECD countries. American Journal of Sociology. 2002; 107(5):1244–1299.

Allen DW. Social class, race, and toxic releases in American counties, 1995. The Social Science Journal. 2001; 38(1):13–25.

Allison PD. Measures of inequality. American Sociological Review. 1978; 43(6):865–880.

Brodish PH, Massing M, Tyroler HA. Income inequality and all-cause mortality in the 100 counties of North Carolina. Southern Medical Journal. 2000; 93(4):386–391. [PubMed: 10798507]

Buhai S. Quantile regression: overview and selected applications. Ad-Astra-The Young Romanian Scientists' Journal. 2005; 5(1):1–17.

Cossman RE, Cossman JS, Cosby AG, Reavis RM. Reconsidering the ruraleurban continuum in rural health research: a test of stable relationships using mortality as a health measure. Population Research and Policy Review. 2008; 27(4):459–476.

Daly MC, Duncan GJ, Kaplan GA, Lynch JW. Macro to micro links in the relation between income inequality and mortality. Milbank Quarterly. 1998; 76(3):315–339. [PubMed: 9738166]

Deaton A. Health, inequality, and economic development. Journal of Economic Literature. 2003; 41:113–158.

Deaton A, Lubotsky D. Mortality, inequality and race in American cities and states. Social Science & Medicine. 2003; 56(6):1139–1153. [PubMed: 12600354]

DeNavas-Walt, C.; Proctor, B.; Smith, J. U.S. Census Bureau. Income, poverty, and health insurance coverage in the United States: 2008. Washington, DC: U.S. Government Printing Office; 2009.

Fitzenberger, B.; Koenker, R.; Machado, JAF. Economic applications of quantile regression. Physica-Verlag; 2002.

Franzini L, Ribble J, Spears W. The effects of income inequality and income level on mortality vary by population size in Texas counties. Journal of Health and Social Behavior. 2001; 42(4):373–387. [PubMed: 11831138]

Gilbert, D. The American class structure in an age of growing inequality. Pine Forge Pr; 2008.

Glaeser, EL.; Laibson, D.; Sacerdote, B. The economic approach to social capital. Mass., USA: National Bureau of Economic Research Cambridge; 2000.

Hao, L.; Naiman, DQ. Quantile regression. Sage Publications, Inc; 2007.

House JS. Relating social inequalities in health and income. Journal of Health Politics, Policy and Law. 2001; 26(3):523.

James WL, Cossman JS. Does regional variation affect ecological mortality research? An examination of mortality, income inequality and health infrastructure in the Mississippi Delta. Population Research and Policy Review. 2006; 25(2):175–195.

Jensen, L.; Yang, TC. Taken by surprise: new immigrants in the rural United States. In: Jentsch, B.; Simard, M., editors. International migration and rural areas – Global perspectives. Hampshire, UK: Ashgate Publishing, Ltd; 2009. p. 17-41.

Kaplan GA, Pamuk ER, Lynch JW, Cohen RD, Balfour JL. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. British Medical Journal. 1996; 312(7037):999–1003. [PubMed: 8616393]

Kawachi I, Blakely TA. When economists and epidemiologists disagree. Journal of Health Politics, Policy and Law. 2001; 26(3):533.

Kawachi I, Blakely TA. Response: substance and sophistry in the analysis of social inequalities in health. Journal of Health Politics, Policy and Law. 2002; 27(2):297–302.

Kawachi I, Kennedy BP. The relationship of income inequality to mortality: does the choice of indicator matter? Social Science & Medicine. 1997; 45(7):1121–1127. [PubMed: 9257403]

Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality, and mortality. American Journal of Public Health. 1997; 87(9):1491–1498. [PubMed: 9314802]

Kawachi, I.; Levine, S.; Miller, M.; Lasch, K.; Amick, B. Joint Program in Society & Health, Health Institute. Income inequality and life expectancy: Theory, research, and policy. New England Medical Center and Harvard School of Public Health; 1994.

Kennedy BP, Kawachi I, Prothrow-Stith D. Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States. British Medical Journal. 1996; 312(7037):1004–1007. [PubMed: 8616345]

Koenker R, Bassett G. Regression quantiles. Econometrica: Journal of the Econometric Society. 1978; 46(1):33–50.

Koenker R, Hallock KF. Quantile regression: an introduction. The Journal of Economic Perspectives. 2001; 15(4):143–156.

Kutner, MH.; Nachtsheim, C.; Neter, J. Applied linear regression models. McGraw-Hill; 2004.

Laporte A. A note on the use of a single inequality index in testing the effect of income distribution on mortality. Social Science & Medicine. 2002; 55(9):1561–1570. [PubMed: 12297242]

Link BG, Phelan J. Social conditions as fundamental causes of disease. Journal of Health and Social Behavior. 1995; 35(Extra issue):80–94. [PubMed: 7560851]

Lobao L, Kraybill DS. The emerging roles of county governments in metropolitan and nonmetropolitan areas: findings from a national survey. Economic Development Quarterly. 2005; 19(3):245.

Lynch J, Kaplan GA. Understanding how inequality in the distribution of income affects health. Journal of Health Psychology. 1997; 2(3):297–314. [PubMed: 22013024]

Lynch J, Smith GD, Harper S, Hillemeier M, Ross N, Kaplan GA, et al. Is income inequality a determinant of population health? Part 1. A systematic review. Milbank Quarterly. 2004; 82(1):5–99. [PubMed: 15016244]

Mansfield ER, Helms BP. Detecting multicollinearity. American Statistician. 1982; 36(3):158–160.

Marmot, M. Status syndrome. London: Bloomsbury; 2004.

Massing MW, Rosamond WD, Wing SB, Suchindran CM, Kaplan BH, Tyroler HA. Income, income inequality, and cardiovascular disease mortality: relations among county populations of the United States, 1985 to 1994. Southern Medical Journal. 2004; 97(5):475–484. [PubMed: 15180024]

Matthews, SA. Spatial polygamy and the heterogeneity of place: studying people and place via egocentric methods. In: Burton, LM.; Kemp, SP.; Leung, M.; Matthews, SA.; Takeuchi, DT., editors. Communities, neighborhoods, and health: Expanding the boundaries of place. Springer; 2011. p. 35-55.

McLaughlin DK, Stokes CS. Income inequality and mortality in US counties: does minority racial concentration matter? American Journal of Public Health. 2002; 92(1):99–104. [PubMed: 11772770]

McLaughlin DK, Stokes CS, Nonoyama A. Residence and income inequality: effects on mortality among US counties. Rural Sociology. 2001; 66(4):579–598.

McLaughlin, DK.; Stokes, CS.; Smith, PJ.; Nonoyama, A. Differential mortality across the United States: the influence of place-based inequality. In: Lobao, LM.; Hooks, G.; Tickamyer, AR., editors. The sociology of spatial inequality. Albany: State University of New York Press; 2007.

Mellor JM, Milyo J. Reexamining the evidence of an ecological association between income inequality and health. Journal of Health Politics, Policy and Law. 2001; 26(3):487–522.

Mellor JM, Milyo J. Income inequality and individual health: evidence from the Current Population Survey. Journal of Human Resources. 2002a; 37(3):510–539.

Mellor JM, Milyo J. On the use of age-adjusted mortality rates in studies of income inequality and population health. Journal of Health Politics, Policy and Law. 2002b; 27(2):293.

Moller S, Nielsen F, Alderson AS. Changing patterns of income inequality in US counties, 1970–2000. American Journal of Sociology. 2009; 114(4):1037–1101.

Muller A. Education, income inequality, and mortality: a multiple regression analysis. British Medical Journal. 2002; 324(7328):23–25. [PubMed: 11777800]

NCHS. Compressed mortality file, 1989–1998 (machine readable data file and documentation, CD-ROM series 20, No F). Hyattsville, Maryland: National Center for Health Statistics; 2003.

NCHS. Compressed mortality file, 1999–2007 (machine readable data file and documentation, CD-ROM series 20, No 2M). Hyattsville, Maryland: National Center for Health Statistics; 2006.

Openshaw, S. The modifiable areal unit problem. Norwich, UK: Geo Books; 1984.

Piantadosi S, Byar DP, Green SB. The ecological fallacy. American Journal of Epidemiology. 1988; 127(5):893. [PubMed: 3282433]

Rupasingha A, Goetz SJ, Freshwater D. The production of social capital in US counties. Journal of Socio-Economics. 2006; 35(1):83–101.

Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science. 1997; 277(5328):918–924. [PubMed: 9252316]

Shi L, Macinko J, Starfield B, Politzer R, Wulu J, Xu J. Primary care, social inequalities, and all-cause, heart disease, and cancer mortality in US counties, 1990. American Journal of Public Health. 2005; 95(4):674–680. [PubMed: 15798129]

Shi L, Starfield B, Kennedy B, Kawachi I. Income inequality, primary care, and health indicators. Journal of Family Practice. 1999; 48(4):275–284. [PubMed: 10229252]

Song, L.; Son, J.; Lin, N. Social capital and health. In: Cockerham, W., editor. The new blackwell companion to medical sociology. Malden, MA: Blackwell; 2010. p. 184-210.

Sparks PJ, Sparks CS. An application of spatially autoregressive models to the study of US county mortality rates. Population, Space and Place. 2010; 16(6):465–481.

USDT. Summary of travel trends: 2009 national household travel survey. U.S. Department of Transportation Federal Highway Administration; 2009.

Wilkinson RG. Income, inequality and social cohesion. American Journal of Public Health. 1997; 87(9):1504–1506. [PubMed: 9314804]

Wilkinson RG. The impact of inequality. Social Research: An International Quarterly. 2006; 73(2): 711–732.

Wilkinson RG, Kawachi I, Kennedy BP. Mortality, the social environment, crime and violence. Sociology of Health & Illness. 1998; 20(5):578–597.

Wilkinson RG, Pickett KE. Income inequality and population health: a review and explanation of the evidence. Social Science & Medicine. 2006; 62(7):1768–1784. [PubMed: 16226363]

Wilkinson, RG.; Pickett, KE. The spirit level: Why equality is better for everyone. Penguin; 2009.

Wolfson M, Kaplan G, Lynch J, Ross N, Backlund E, Gravelle H, et al. Relation between income inequality and mortality: empirical demonstration diminishing returns to aggregate level studies two pathways, but how much do they diverge? British Medical Journal. 1999; 319(7215):953–955. [PubMed: 10514157]

Yang TC, Jensen L, Haran M. Social capital and human mortality: explaining the rural paradox with county-level mortality data. Rural Sociology. 2011; 76(3):347–374.

Yang TC, McManus B. Infant mortality and social environment in Georgia: an application of hotspot detection and prioritization. Environmental and Ecological Statistics. 2010; 17(4):473–502.

Yang TC, Teng HW, Haran M. The impacts of social capital on infant mortality in the US: a spatial investigation. Applied Spatial Analysis and Policy. 2009; 2(3):211–227.

Table 1

Descriptive statistics of the variables in this study ($N = 3072$).

Variable	VIF[a]	Mean	Median	Maximum	Minimum	Std dev
Dependent variable						
Mortality	N.A.	8.90	8.85	19.78	0.00	1.38
Independent variables						
Inequality						
Gini	1.92	43.42	43.21	60.50	33.33	3.75
Racial/ethnic compositions						
Percent Black	2.06	8.72	1.64	86.08	0.00	14.50
Percent Hispanic	1.51	6.16	1.75	98.10	0.00	12.12
Percent other races	1.74	3.49	1.86	93.58	0.00	6.76
Rurality						
Industrial structure	1.36	0.00	−0.20	8.65	−2.89	1.00
Ecological dimension	1.31	0.00	−0.17	28.70	−0.61	1.00
EEI	1.76	0.00	−0.12	4.54	−1.92	1.00
SES						
Social affluence	2.80	0.00	−0.19	6.01	−2.43	1.00
Concentrated disadvantage	4.29	0.00	−0.18	9.06	−2.54	1.00
Social capital						
Social capital index	1.93	0.00	−0.12	7.66	−4.06	1.29
Safety	1.38	0.00	−0.20	12.12	−1.37	1.00
Stability	2.29	0.00	0.05	1.70	−4.15	0.59

[a]VIF: Variance inflation factor.

Table 2

Percentiles and mortality rates ($N = 3072$).

Percentile	Mortality (per 1000 population)
5	6.95
10	7.36
15	7.57
20	7.81
25	8.01
30	8.18
35	8.34
40	8.50
45	8.69
50	8.85
55	9.01
60	9.18
65	9.36
70	9.55
75	9.74
80	9.95
85	10.24
90	10.56
95	11.05

Table 3

The estimated coefficients, bootstrap standard errors, and *p*-values for inequality by quantiles.

Quantiles (τ)	**0.05**	**0.10**	**0.15**	**0.20**	**0.25**	**0.30**	**0.35**	**0.40**	**0.45**	**0.50**	**0.55**	**0.60**	**0.65**	**0.70**	**0.75**	**0.80**	**0.85**	**0.90**	**0.95**
Coefficient	0.009	0.003	0.021	0.026	0.030	0.031	0.032	0.037	0.037	0.040	0.042	0.042	0.049	0.049	0.051	0.051	0.048	0.045	0.041
S.E.	0.011	0.010	0.009	0.008	0.008	0.008	0.006	0.006	0.006	0.007	0.007	0.007	0.007	0.007	0.008	0.008	0.009	0.011	0.013
p-value	0.412	0.748	0.020	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002

Table 4

The Wald tests for equality of inequality coefficients across quantiles of mortality.

Quantile (τ)	*p*-value				
	Coefficient	Different from coefficient at median?	Different from coefficient at (1 − τ)th quantile?	Different from coefficient at (τ + 0.05)th quantile?	4 Coefficients jointly different?
0.05	0.009	**0.011**	**0.065**	0.537	**0.007**
0.10	0.003	**0.001**	**0.005**	**0.013**	**0.005**
0.15	0.021*	**0.041**	**0.020**	0.433	0.112
0.20	0.026***	**0.069**	**0.007**	0.303	**0.059**
0.25	0.030***	0.194	**0.021**	0.951	**0.098**
0.30	0.031***	0.154	**0.023**	0.755	0.109
0.35	0.032***	0.144	**0.010**	0.188	**0.075**
0.40	0.037***	0.470	0.368	0.819	0.796
0.45	0.037***	0.457	0.272	0.546	–
0.50	0.040***	–	–	0.501	–

Note:

* $p \leq 0.05$;

*** $p \leq 0.001$;

$p \leq 0.10$ in bold for significant differences from coefficients.

Table 5

Quantile regression estimates for all covariates in the analysis ($N = 3072$).

Variables	τ = 0.05	τ = 0.15	τ = 0.25	τ = 0.50	τ = 0.75	τ = 0.85	τ = 0.95
Intercept	7.361*** (0.514)	7.303*** (0.405)	7.148*** (0.361)	7.171*** (0.313)	7.140*** (0.327)	7.588*** (0.388)	8.313*** (0.598)
Inequality							
Gini	0.009 (0.011)	0.021* (0.009)	0.030*** (0.008)	0.040*** (0.007)	0.051*** (0.008)	0.048*** (0.009)	0.041** (0.013)
Racial/ethnic compositions							
Percent Black	0.010** (0.003)	0.008*** (0.002)	0.009*** (0.002)	0.012*** (0.002)	0.012*** (0.002)	0.010*** (0.003)	0.012*** (0.003)
Percent Hispanic	−0.036*** (0.005)	−0.025*** (0.004)	−0.021*** (0.003)	−0.018*** (0.002)	−0.018*** (0.002)	−0.019*** (0.003)	−0.013** (0.005)
Percent other races	−0.012 (0.008)	−0.012 (0.008)	−0.004 (0.007)	0.004 (0.005)	0.009 (0.008)	0.012 (0.008)	0.038* (0.018)
Rurality							
Industrial structure	−0.521*** (0.074)	−0.431*** (0.038)	−0.374*** (0.043)	−0.309*** (0.029)	−0.240*** (0.028)	−0.158*** (0.037)	0.012 (0.058)
Ecological dimension	0.071* (0.036)	0.064 (0.047)	0.050 (0.038)	0.045 (0.032)	0.013 (0.042)	−0.017 (0.054)	−0.002 (0.055)
EEI	0.017 (0.049)	0.094* (0.045)	0.147*** (0.035)	0.200*** (0.021)	0.181*** (0.025)	0.179*** (0.027)	0.275*** (0.041)
SES							
Social affluence	−0.477*** (0.077)	−0.403*** (0.052)	−0.412*** (0.035)	−0.428*** (0.030)	−0.433*** (0.028)	−0.385*** (0.040)	−0.384*** (0.051)
Concentrated disadvantage	0.351*** (0.081)	0.370*** (0.058)	0.366*** (0.049)	0.337*** (0.048)	0.300*** (0.050)	0.372*** (0.064)	0.452*** (0.085)
Social capital							
Social capital index	−0.224*** (0.044)	−0.198*** (0.037)	−0.161*** (0.031)	−0.124*** (0.024)	−0.147*** (0.024)	−0.119*** (0.026)	−0.063 (0.034)
Safety	0.133** (0.042)	0.131*** (0.031)	0.158*** (0.026)	0.145*** (0.025)	0.122*** (0.026)	0.141*** (0.029)	0.116** (0.040)
Stability	−0.142 (0.090)	−0.130* (0.063)	−0.165*** (0.047)	−0.172*** (0.042)	−0.192*** (0.047)	−0.206** (0.066)	−0.306** (0.133)

Note:

* $p \leq 0.05$;

** $p \leq 0.01$;

*** $p \leq 0.001$;

Standard errors are reported in the parentheses.