Composition and decomposition in US gender-specific self-reported health disparities, 1984–2007

Hui Zheng [a,*], Kenneth C. Land [b]

[a] Department of Sociology, The Ohio State University, Columbus, OH 43210, USA
[b] Center for Population Health and Aging, Duke Population Research Institute and Department of Sociology, Duke University, Durham, NC 27708, USA

ARTICLE INFO

ABSTRACT

Variance function regression models and demographic decomposition methods are applied to identify two dimensions of changes in health disparities (SES-demographic effects vs. compositional effects, between-group disparities vs. within-group disparities) in the US from 1984 to 2007. Using National Health Interview Survey data on self-reported health, we find that disparities in men's health increased, while those of women decreased, for the whole period. Widening men's health disparities are largely driven by increases in the effects of SES-demographic statuses on within-group disparities. These increases are moderated by increasing levels of men's college attainment. But decreasing middle and upper income attainment and a decreasing employment rate further increase men's health disparities. For women, the effects of SES-demographic statuses on health disparities also increased over time. This, however, was outweighed by increases in women's college attainment, middle and upper income attainment, and employment rate. The result is overall declining self-reported health disparities for women.

1. Introduction

The growth of health disparities in the US across the past two decades has been one of the most extensively researched topics in demography, medical sociology, epidemiology and public health (William and Collins, 1995; Hummer et al., 1998). Disparities in health may occur across categories of socioeconomic status (SES, including income, education and occupation) and such sociodemographic characteristics as work status, marital status, race/ethnicity and gender. Researchers have examined health disparities by studying several different health outcomes, including mortality (e.g., Papps et al., 1993), cause-specific mortality (e.g., Steenland et al., 2004; Jemal et al., 2008), morbidity (e.g., Schoeni et al., 2005), life expectancy (e.g., Crimmins and Saito, 2001; Meara et al., 2008), and self-rated health (e.g., Goesling, 2007).

Previous research, however, tended to focus on changes in gradients of SES and sociodemographic statuses (hereafter termed SES-demographic statuses) on health outcomes, that is, on the extent to which a health outcome varies as a function of each of a set of specific socioeconomic and demographic statuses. But health disparities can grow or decline due to two distinct mechanisms: the *gradient* or set of differential relationships of SES-demographic statuses to health outcomes can change and the *composition of the population* with respect to SES-demographic statuses itself can change. These two factors are confounded with each other. For example, growing health disparities can be caused not only by increasing gradients of the effects of SES-demographic statuses on health (i.e., a SES-demographic effect), but also can result from a decreasing

---

* Corresponding author.
  E-mail address: zheng.64@sociology.osu.edu (H. Zheng).

probability for members of the population to be located at middle statuses (i.e., a compositional effect). Separating SES-demographic effects from compositional effects then are essential in studying health disparities.

Moreover, health disparities are composed of between-group disparities and within-group disparities. *Between-group health disparities* represent variations in health across groups with different SES-demographic characteristics. *Within-group health disparities* describe health heterogeneity within each group with the same characteristics. Previous research has focused on between-group health disparities based on regression-based analysis, but within-group disparities in outcomes of social processes can far exceed between-group disparities (Western et al., 2008).

To further explore and analyze the impacts of SES-demographic effects vs. compositional effects, between-group disparities vs. within-group disparities on health disparities, this study applies a variance function regression model and standard decomposition methods to the analysis of self-reported health outcomes in the US National Health Interview Survey. Extant research suggests that temporal trends of health disparities in the past two decades have differed for men and women. Specifically, health disparities have consistently widened for men, while they have slightly increased, stagnated, contracted or significantly increased for women, depending on the time periods and health outcomes examined (Feldman et al., 1989; Preston and Elo, 1995; Liu and Hummer, 2008). Therefore, we investigate the sources of the divergent gender-specific self-rated health disparities trends.

We commence in the next section with a review of prior research on health disparities and trends therein over recent decades in the United States. We then review explanations of these trends. This is followed by a description of the data analyzed, the statistical models applied to the data, and the findings that result therefrom. We conclude with a discussion of the findings, limitations of the study, and directions for additional research.

2. Trends in health disparities

2.1. Widening health disparities

Health disparities have increasingly become a major dimension of general social inequality in the United States. Socioeconomic statuses (SESs: along the dimensions of education, occupation, and income – a Weberian construct) are major social determinants of the distribution of health and disease in American society. Research has consistently documented an inverse relationship between SES and risks of diseases and death in the recent decades (e.g., William and Collins, 1995; Link and Phelan, 1995). The mechanisms accounting for this relationship include greater exposure to stress and hardship and more limited access to valuable resources (e.g., food, housing, health care, and medical knowledge), which can help prevent and cure disease, among lower SES individuals (George, 2005; Adler and Newman, 2002). In addition, health is not a singular condition and different components of SES (e.g., education, income and occupation) may have different impacts on different health outcomes. For example, some studies suggest that education exerts a stronger effect on the onset of disease and sickness, while income plays a more important role on the progression afterwards (e.g., Zimmer and House, 2003).

In the US context, race is another important predictor of health and disease; particularly, whites usually have better health than blacks. But research has shown divergent findings on the extent to which SES differences can account for race disparities in health. On the one hand, research suggests SES differences account for a large proportion of racial disparities in health (e.g., Rogers, 1992), although it does not eliminate the disparities (e.g., Krieger and Fee, 1994). On the other hand, some research finds even larger racial disparities in health at higher levels of SES – that is, when SES increases, blacks do not achieve as much improvement in health as whites have (Farmer and Ferraro, 2005). In spite of these divergent findings, most studies coherently demonstrate that race is a potent predictor of variations in health.

Many recent studies in demography, medical sociology and epidemiology have observed rising socioeconomic differentials in health in the US in the past several decades, that is, socioeconomic status has increasingly exerted a stronger effect on health and diseases in the past several decades (e.g., Feldman et al., 1989; Papps et al., 1993; Preston and Elo, 1995; Hummer et al., 1998; Meara et al., 2008). For example, Papas et al. (1993) found a sharper increase in socioeconomic disparity in mortality differentials since 1960 and this result holds across gender, race and marital status. Mortality differentials across socioeconomic statuses grew between 1990 and 2000 (Meara et al., 2008; Jemal et al., 2008). Besides mortality differentials, research has found growing educational disparities in old-age disability (Schoeni et al., 2005), disability-free life expectancy (Crimmins and Saito, 2001), and total life expectancy (Meara et al., 2008) over recent decades. Similarly, a trend towards widening gaps in self-rated health by educational levels for middle-aged and older adults has continued since the early 1980s (Goesling, 2007; Liu and Hummer, 2008) or for all ages among younger cohorts (Lynch, 2003).

Besides a growing health disparity by SES, disparities also have widened across marital status, race and ethnicity. With regard to marital status, studies found that in the past several decades male mortality differences between the married and unmarried statuses have widened (Smith, 1996; Hummer et al., 1998), and women's self-rated health has improved more for the married than for the widowed, divorced and separated (Liu and Umberson, 2008). And the racial gap in mortality and life expectancy continues widening, which is largely explained by a slower rate of decline among blacks than whites for heart disease, while HIV infection, diabetes, pneumonia and homicide are major causes of decreasing life expectancy for blacks (Kochanek et al., 1994).

2.2. Trends in gender-specific health disparities

Although enlarging health disparities have become a widespread phenomenon in the US that holds across gender (Papps et al., 1993), some research actually suggests different trends of health disparities by sex. For example, Feldman et al. (1989) compared the mortality rate in the 1960 Matched Records Study data and in 1971–1984 data from the first National Health and Nutrition Examination Survey (NHANES I) Epidemiologic Followup Study (NHEFS). They found that educational differentials in men's death rates widened between 1960 and 1984, because death rates declined more rapidly among the more educated than the less educated. In contrast, women's death rates declined at similar rates across levels of educational attainment, which produces similar magnitudes of mortality disparities between 1960 and 1984 for women. The National Center for Health Statistics (1994) reported that the racial gap in life expectancy widened much more for men (from 6.9 years to 8.3 years) than for women (from 5.6 years to 5.8 years) between 1980 and 1991. Moreover, Preston and Elo (1995) found that educational disparities in adult mortality widened for men but contracted for working-age women. In contrast, most recent research suggests that educational disparities in self-rated health and life expectancy have widened among women in the two most recent decades (Liu and Hummer, 2008; Meara et al., 2008), which may be due to the increasing importance of education not only for labor market and income outcomes (Hamil-Luker, 2005), but also for a higher probability of marriage, a higher standard of living and insurance against poverty among women than ever before (DiPrete and Buchmann, 2006). In sum, these studies generally have found that health disparities have widened for men, while only slightly increasing, stagnating, contracting, or significantly increasing for women – but the findings for women also depend on the time periods and health outcomes examined. Nonetheless, there is substantial evidence of major gender differences in trends in health disparities. Accordingly, in the analyses reported below, we examine *gender-specific self-rated health disparities trends*, with the recognition that findings may not generalize to health disparity trends associated with other health outcomes.

3. Explaining trends in health disparities

3.1. Changing gradients and changing population composition

Much the prior research on health disparities has focused on health differences across individual-level characteristics (e.g., income, education, occupation, work status, race, ethnicity, gender, marital status) and how these differences change over time. This line of research is more concerned with *changing gradients* of the relationship of individual-level characteristics to health. However, changing health disparities can also be attributed to *changing population composition*, which has been neglected in previous research.

Enlarging health disparities can be caused, for example, by increasing gradients of the relationship of SES-demographic characteristics to health (e.g., a SES-demographic effect) or a decreasing probability for lower SES groups to enter middle or higher statuses (i.e., a compositional effect), and vice versa for declining health disparities. For instance, health disparities can increase due to the increasing importance of education on earnings (DiPrete and Buchmann, 2006), access to advanced health care and medical technology innovation (Glied and Lleras-Muney, 2008), or others beneficial for health (i.e., a socioeconomic effect), but also can be reduced by increasing higher education attainment and labor force participation among lower income families (i.e., a compositional effect).

Therefore, although health disparity has increased among women in recent years due to the increasing importance of education for labor force participation, income and health (Liu and Hummer, 2008) (i.e., a socioeconomic effect), this increase is possibly offset by increasing higher education achievement in lower status families (Buchmann and DiPrete, 2006) and by increasing labor force participation (Juhn and Potter, 2006) among women (i.e., a compositional effect). In contrast, health outcomes for men may be worsened by the double effect of the increasing importance of education for earnings and standard of living (i.e., a socioeconomic effect), and by declining labor force participation (Juhn and Potter, 2006) among men (i.e., a compositional effect). Therefore, in order to attain a complete picture of changing health disparities, it is very important to separate SES-demographic effects from compositional effects.

3.2. Between-group and within-group disparities

Recent research suggests increases in income inequality can occur both between groups and within groups (Western et al., 2008; Western and Bloome, 2009; Lemieux, 2006). This research proceeds by decomposing increases in income inequality into between-group disparities and within-group disparities. The same logic can be applied to health disparities. Between-group health disparities represent variation in health across groups with different characteristics, for example, health difference between less educated persons and better educated persons. Within-group health disparities describe health heterogeneity within each group with the same characteristics – for example, within a group defined by low levels of educational attainment. Previous research has focused on between-group health disparities based on regression-based analysis, but within-group disparities also can be substantial and should be studied.

Within-group health disparities can be estimated by residual variances from regression analyses. These residuals usually are neglected by scholars. Within-group health disparities may be due to variations in health within groups, differing

variances of measurement error, and genetic effects. Research suggests that genetic factors can account for as much as 50% of frailty and mortality differences among individuals (Yashin et al., 1999; Iachine et al., 1998). Added to variations in health within groups, it follows that within-group disparities may contribute more to total health disparities than between-group disparities. However, we are more interested in explaining how *changes* in health disparities are associated with *changes* in between-group disparities and *changes* in within-group disparities. In this case, *changes* in health disparities are not necessarily largely driven by *changes* in within-group disparities. Assuming group-specific variances of measurement error are relatively constant over time (Lemieux, 2006), and noting that the time scale of genetic changes involves many generations, changes in within-group health disparities are then mostly driven by changes in variations in health within groups.

The previous section has described how changes in between-group health disparities can be explained by changing gradients of SES-demographic effects on health and by the changing SES-demographic composition of the population, which also contribute to changes in within-group health disparities. In other words, changes in within-group disparities can be caused by changes in variations in health within each socioeconomic-demographic group (e.g., a SES-demographic effect) and/or changes in population distribution in groups with different within-group health variations (e.g., a compositional effect). For example, an increase in within-group health disparities can be associated with increasing within-group health variations among college graduates (a SES-demographic effect) and/or with a decreasing population proportion of college graduates who have lower within-group health variations than non-college graduates (a compositional effect).

In this study, we focus on how health disparities may change due to the changing gradient of education, income and employment status on health, and the composition of the population with respect to these three statuses.

4. Data

Our analysis is based on annual data from the US National Health Interview Survey (NHIS) for the 24-year period 1984–2007. NHIS is a repeated cross-sectional sample survey of civilian non-institutionalized US population ages 18 and over conducted by National Center for Health Statistics. NHIS collects health information for each member of a family or household sampled, as reported by one primary respondent. In order to reduce reporting/measurement errors, we limit our analysis to the primary respondent. The sample size for men is about 16,837 each year (in total 16,837 ∗ 24 = 404,098), and for women is about 12,439 each year (in total 12,439 ∗ 24 = 298,546).

The sample frame for the NHIS is redesigned every 10 years. Nonetheless, the fundamental design of the 1995–2007 NHIS is similar to that of the 1985–1994 NHIS. Two changes in the sampling design are notable. First, the number of primary sampling locations has increased from 198 to 358 since 1995. Second, both black and Hispanic populations were oversampled in the 1995–2007 NHIS, while only blacks were oversampled in the 1985–1994 NHIS. These two redesigns potentially increase the variances (health disparities in our paper) among samples. As discussed in the Results section, we study trends in health disparities by adjusting for population compositional changes, which encompasses both real changes in the society and sampling changes in the survey.

The question and response structure for the key outcome variable analyzed here, *self-rated health*, has remained largely unchanged across periodic revisions of the NHIS questionnaires, which facilitates the analysis of trends. It has five response categories: poor, fair, good, very good, and excellent. Self-rated health is a widely used measure of general health status that has been found to be very predictive of mortality and strongly correlated with objective assessments of health, including physician diagnoses (Idler and Benyamini, 1997; Case and Paxson, 2005). Close relationships between self-rated health and objective health indicators also hold across population subgroups (Bosworth et al., 1999; Kennedy et al., 2001). In order to differentiate respondents with very good health from those with average or poor health and get a better (less skewed) distribution of the outcome variable, we recoded the measure into a dichotomous variable coded 1 for people reporting "very good" or "excellent" health and 0 for people reporting "poor", "fair" or "good" health[1].

To facilitate application of variance function regressions and decomposition methods (described in the Research Strategy and Models section), we also recoded all explanatory variables into categorical variables. In the NHIS, income was measured by several income categories. We first calculated the mid-point of each income category and then converted the mid-points to 2007 US dollars. We recoded family income to three income levels: below the 20th percentile, 20th to 50th percentile, and above the 50th percentile.[2] A detailed description of explanatory variables is given in Table 1. In brief, we grouped the explanatory variables for each gender into three income levels, three education categories, two employment statuses, six age groups, two race categories, and two marital statuses.[3] This yields 3 ∗ 3 ∗ 2 ∗ 6 ∗ 2 ∗ 2 = 432 groups for each gender.

---

[1] As robustness analyses, we also conducted analyses in which we: a) dichotomized the responses into fair or poor versus excellent, very good, or good (see, e.g., Lynch, 2003), and b) treated this ordinal measure as continuous (e.g., Schnittker, 2007). Overall findings are similar. This is consistent with Manor et al. (2000) who found that a dichotomous coding of self-rated health responses produced similar results to those obtained with alternative statistical methods that accommodate the ordered nature of self-rated health, e.g., polytomous regression, cumulative odds, continuation ratio and adjacent categories models.

[2] We could not identify whether people are above the 80th or 90th percentile due to questionnaire limitations in the NHIS. For example, in the 2007 NHIS, the highest income category is $75,000 and above.

[3] We also conducted the analyses with the initial age group beginning at 30, so the majority of respondents would have finished their formal educations. But inclusion or exclusion of younger adults does not substantially change the findings. Therefore, the analyses with all adults are presented here.

Table 1
Description of explanatory variables in the NHIS, 1984 to 2007.

Variable	Description
Family income	1 = below 20th percentile, 2 = 20th to 50th percentiles, and 3 = above 50th percentile
Education	1 = without high school degree, 2 = high school graduates, and 3 = college graduates
Employment status	1 = employed, and 2 = unemployed
Age	1 = 18/29, 2 = 30/39, 3 = 40/49, 4 = 50/59, 5 = 60/69, and 6 = 70/85
Race	1 = whites, and 2 = non-whites
Marital status	1 = married, and 2 = unmarried

5. Research strategy and models

We applied variance function regression and decomposition methods to study gender-specific health disparity trends in the US from 1984 to 2007. By using these methods, we can separate SES-demographic effects from compositional effects, between-group disparities from within-group disparities (Western et al., 2008; Western and Bloome, 2009; Lemieux, 2006).

5.1. Variance function regression

We measure health disparity by the variance in probabilities of reporting very good or excellent health. Sociologists often emphasize between-group inequality (or disparity) based on regression-based analyses of inequality. But within-group or residual inequality (or disparity) may far exceed between-group differences (Western et al., 2008). By using variance function regression, we can analyze the extent to which increasing or decreasing health disparity in the US can be attributed to increasing or decreasing between-group differences as compared to increasing or decreasing within-group differences. Variance function regression has two parts, including a regression for conditional mean, y_i, and a regression for logarithm of the residual variances, $\log(\sigma_i^2)$ (Western and Bloome, 2009):

$$\hat{y}_i = x_i'\beta \tag{1}$$

$$\log(\sigma_i^2) = z_i'\lambda, \tag{2}$$

where observations on individual sample members are indexed by i, x_i' is a vector of covariates for mean or expected value $\hat{y}_i$, and z_i' is a vector of covariates (possibly equal to x_i') for the conditional expected value of logarithm of the residual variance $\log(\sigma_i^2)$. The quantity σ_i^2 is the squared residuals, $\hat{e}_i^2$, from the first regression, where $\hat{e}_i = y_i - x_i'\hat{\beta}$. From a substantive viewpoint, the first regression describes how covariates that define the 432 cells described above affect the y_i response variable and account for the deviations of the within-cell sample means from the average or grand mean $\bar{y}$ (which can be termed the between-group inequality), while the second regression explains how covariates affect the within-cell variability of the response variable around the cell means, that is, the unpredictability of y_i within the 432 cells (which can be termed the within-group inequality).

The models are estimated by maximum likelihood, using an iterative two stage method (Western and Bloome, 2009; Aitkin, 1987). In the iteration process, we repeat a weighted linear regression of y_i on x_i, with weights, $1/\hat{\sigma}_i^2$, which is estimated from the gamma regression of the squared OLS residuals using a log link function (i.e., the second regression). This method can generate unbiased, consistent and efficient estimations when sample size is large. Although the outcome variable, self-rated health is coded as a binary variable, we use a linear probability model rather than a binary logistic model for the first regression because the sample size is large enough to generate accurate estimates with the linear probability model.[4]

5.2. Decomposition and standardization

Because we are interested in how health disparities (variances in probabilities of reporting very good or excellent health) changed from 1984 to 2007, we estimated the variance function regressions for each year instead of pooling all 24 years together. The variance function regressions can be denoted:

$$\hat{y}_{ti} = x_{ti}'\beta_t$$

$$\log(\sigma_{ti}^2) = z_{ti}'\lambda_t,$$

where t denotes year, ranged from 1984 to 2007. After estimating all the β_t and λ_t coefficients for each year, we predict the probabilities and residuals of reporting very good or excellent health for each group in each year, that is, the $\hat{y}_{tj}$ and $\hat{\sigma}_{tj}^2$, where

---

[4] Specifically, (1) all of the predicted probabilities of reporting very good or excellent health are between 0 and 1 for all 432 groups for each year (1984, ..., 2007); (2) the linear probability model is much easier to interpret than is the binary logistic model (see Hellevik, 2009 for an extended discussion of this point); (3) Powers and Xie (2008: 37) state that if the observed group-specific probabilities are not close to 0 or 1 and the sample size is sufficiently large (conditions that apply in the present analyses) the linear probability model is attractive due to its ease of use and interpretation; and (4) without modification, the binary logistic model does not permit adjustment for weights $1/\hat{\sigma}_i^2$ in the two-stage estimation method (Western and Bloome, 2009).

j demotes group, ranged from 1 to 432. Each group j has a compositional weight or cell proportions, π_{tj}, giving the fraction of individuals falling into this group. Then the total variances in probabilities of reporting very good or excellent health can be written (Western and Bloome, 2009):

$$V_t = B_t + W_t = \sum_{j=1}^{J} \pi_{tj} \hat{r}_{tj}^2 + \sum_{j=1}^{J} \pi_{tj} \hat{\sigma}_{tj}^2, \tag{3}$$

where B_t is between-group component (weighted sum of squared group deviation or between-group variance $\hat{r}_{tj}^2$), and W_t is within-group component (weighted sum of residual variance or within-group variance $\hat{\sigma}_{tj}^2$). Group deviation $\hat{r}_{tj}$ is the deviation of the group mean from grand mean, that is $\hat{r}_{tj} = \hat{y}_{tj} - \bar{y}_t$.

If we have estimates at two time points, $t = 0, 1$, we then can write changes in total variance in the following way:

$$V_1 - V_0 = B_1 - B_0 + W_1 - W_0$$

$$= \underbrace{\sum_{j=1}^{J}(\pi_{1j} - \pi_{0j})\hat{r}_{1j}^2 + \sum_{j=1}^{J}(\hat{r}_{1j}^2 - \hat{r}_{0j}^2)\pi_{0j}}_{B_1 - B_0} + \underbrace{\sum_{j=1}^{J}(\pi_{1j} - \pi_{0j})\hat{\sigma}_{1j}^2 + \sum_{j=1}^{J}(\hat{\sigma}_{1j}^2 - \hat{\sigma}_{0j}^2)\pi_{0j}}_{W_1 - W_0}$$

$$= \underbrace{\sum_{j=1}^{J}(\pi_{1j} - \pi_{0j})(\hat{r}_{1j}^2 + \hat{\sigma}_{1j}^2)}_{E_C} + \underbrace{\sum_{j=1}^{J}(\hat{r}_{1j}^2 - \hat{r}_{0j}^2)\pi_{0j}}_{E_B} + \underbrace{\sum_{j=1}^{J}(\hat{\sigma}_{1j}^2 - \hat{\sigma}_{0j}^2)\pi_{0j}}_{E_W} = E_C + E_B + E_W \tag{4}$$

The change in the between-group variance $B_1 - B_0$ is associated with a compositional effect $\left(\sum_{j=1}^{J}(\pi_{1j} - \pi_{0j})\hat{r}_{1j}^2\right)$ and a socioeconomic effect $\left(\sum_{j=1}^{J}(\hat{r}_{1j}^2 - \hat{r}_{0j}^2)\pi_{0j}\right)$. The change in the within-group variance $W_1 - W_0$ is associated with a compositional effect $\left(\sum_{j=1}^{J}(\pi_{1j} - \pi_{0j})\hat{\sigma}_{1j}^2\right)$ and a SES-demographic effect $\left(\sum_{j=1}^{J}(\hat{\sigma}_{1j}^2 - \hat{\sigma}_{0j}^2)\pi_{0j}\right)$. In sum, Eq. (4) indicates that changes in health disparity V_t can be decomposed into (1) a compositional effect that changes the distribution of population across groups (i.e., E_C, or changes in π_{tj}), (2) a SES-demographic effect that changes the gradient of SES-demographic status on between-group disparities (i.e., E_B, or changes in $\hat{r}_{tj}^2$), and (3) a SES-demographic effect that changes the gradient of SES-demographic statuses on within-group disparities (i.e., E_w, or changes in $\hat{\sigma}_{tj}^2$).

We also can standardize adjusted variances by fixing π_{tj}, $\hat{r}_{tj}^2$, or $\hat{\sigma}_{tj}^2$ at baseline time point, $t = 0$. Adjusted variances can be interpreted as the variance we could observe if population compositions, between-group variances, or within-group variances remained unchanged at their $t = 0$ values. Additionally, we can calculate an explanatory variable's (e.g., income or education) socioeconomic effect or compositional effect by fixing its regression coefficients or weights at $t = 0$. Due to space limits, we cannot describe these standardization techniques in more details; they, however, can be found in Western et al. (2008), Western and Bloome (2009) and Lemieux (2006).

6. Results

6.1. Disparities in men's self-reported health, 1984–2007

Fig. 1 displays the estimated trend of self-reported health disparities for men from 1984 to 2007. The trajectory over time of predicted total variance from the variance function regression tracks very closely with the trajectory of observed total variance, which suggests that the variance function regression does a good job in capturing the trend of self-reported health disparities and the linear probability model generates very accurate estimations. Overall, self-reported health disparity for men increased from 1984 to 2007. But there were variations in the trend within this overall period – specifically, disparity decreased from 1984 to 1990, rose until around 1995, decreased afterwards, and then rose again after 2000.[5]

Fig. 2 further decomposes the total variance into between-group variance and within-group variance. The left vertical axis is for total variance and within-group variance. The right one is for between-group variance. There are two main findings in Fig. 2. First, within-group variance accounts much more than between-group variance for total variance. In 1984, within-group variance and between-group variance contribute about 81% and 19% to total variance, respectively. In 2007, the corresponding numbers are 84% and 16%. This implies that many unobserved factors are not captured in between-group studies that control only for observed explanatory variables such as those in the present analysis. Life style, health habits, genetic differences, and measurement errors may contribute to within-group variances. Second, across the entire period from 1984 to 2007, within-group variance increases, which drives total variance up, while between-group variance decreases and offsets the effect of increases in within-group variance to some extent.

Panel A of Table 2 further separates SES-demographic and compositional effects on the growth of men's self-reported health disparities for the entire 24-year period from 1984 to 2007. Change in variance is decomposed into between-group variance and within-group variance, both of which are further decomposed into two sub-components: SES-demographic effects and compositional effects. The next two rows sum up the SES-demographic effects and compositional effects. The num-

---

[5] Considering the possible increase in variance caused by a sampling redesign of the NHIS in 1995, the "real" disparity may decrease even more after 1995. So the decrease in health disparity between 1995 and 2000 is not artificial, but may be underestimated.
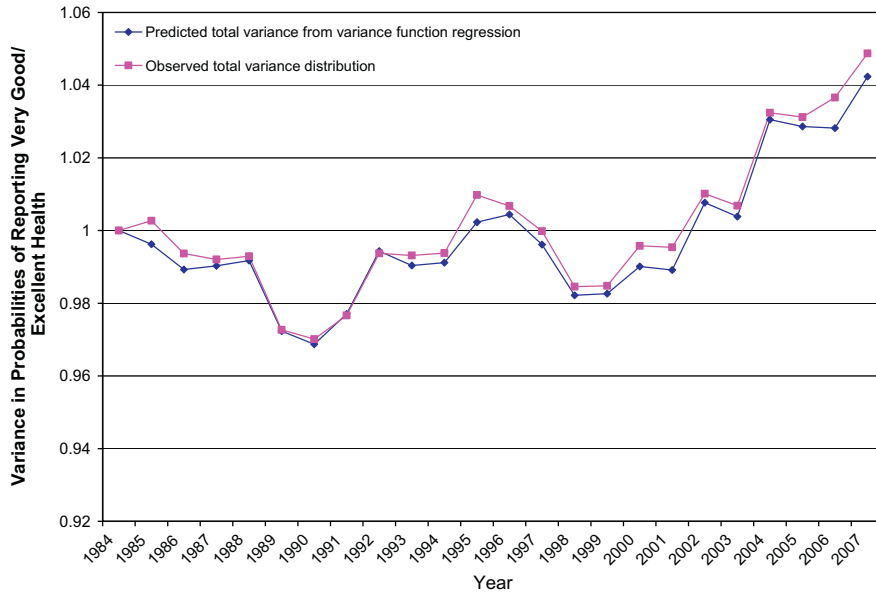
Fig. 1. Comparison between observed variance distribution and predicted variance from variance function regression, US Men, 1984–2007.
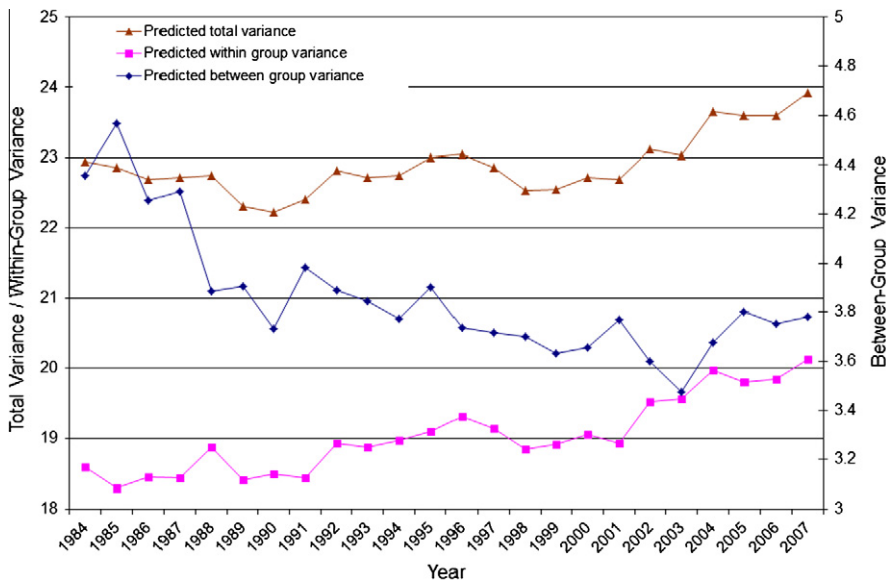

Fig. 2. Decomposition of total variance into between-group variance and within-group variance, US Men, 1984–2007.

bers without parentheses are amounts of change in variance associated with each specific component. The numbers in parentheses are percents of change explained by the corresponding component. For example, the amount of change in variance from 1984 to 2007 associated with between-group SES-demographic effects is −0.059, which contributes about −6.1% (−0.059/0.971) to the total change in variance (i.e., 0.971) from 1984 to 2007. In other words, the decreasing SES-demographic effect on between-group health disparities offsets the overall increasing health disparities to some extent.

There are four main findings from Panel A. First, an increasing within-group variance accounts for a 159% increase in total variance from 1984 to 2007, which is offset by a decreasing between-group variance to some extent. Second, overall, SES-demographic effects on between-group variance decreased from 1984 to 2007, which, however, contribute only about 10% (−0.059/−0.574) to the declines of between-group health disparity. In contrast, compositional effects account for about 90% (−0.515/−0.574) of the declines of between-group health disparity. Third, SES-demographic effects on within-group variance dramatically increased from 1984 to 2007, which contributed 112% (1.739/1.546) to the increase of within-group

Table 2
Decomposition and standardization of the change in variance in probabilities of reporting very good/excellent health, US Men, 1984–2007.

Panel A. Decomposition[a]		
Change in variance	0.971	
Between-group	−0.574	(−59.1)
SES-Demographic effect	−0.059	(−6.1)
Compositional effect	−0.515	(−53.0)
Within-group	1.546	(159.1)
SES-Demographic effect	1.739	(179.1)
Compositional effect	−0.194	(−19.9)
Sum of SES-Demographic effects	1.680	(172.9)
Sum of compositional effects	−0.708	(−72.9)
Panel B. Adjusted change in variance, fixing at 1984		
College effect, β	0.932	(4.1)
College effect, λ	−0.082	(108.4)
College effect, β and λ	−0.121	(112.5)
Middle and upper income effect, β	0.797	(17.9)
Middle and upper income effect, λ	0.609	(37.3)
Middle and upper income effect, β and λ	0.435	(55.2)
Employment status effect, β	1.019	(−4.9)
Employment status effect, λ	0.703	(27.6)
Employment status effect, β and λ	0.751	(22.7)
College attainment	1.148	(−18.2)
Middle and upper income attainment	0.87	(10.4)
Employment status composition	0.697	(28.2)

Note:
[1] Numbers in parentheses are percent of change explained by corresponding factors.
[2] Compositional effects here include compositional changes induced by population composition changes in the reality and sampling design changes in the NHIS surveys. Since sampling redesign may increase variance, the real compositional effects should be more negative. In other words, the compositional effects here may underestimate the real compositional effects.
[a] All the covariates, including income, education, employment status, age, race and marital status, are included in the analysis. Panel A describes the total effects from these six SES-demographic characteristics. Panel B specifies the effects from education, income, and employment status.

health disparities. Compositional effects counterbalance the SES-demographic effects to some extent. Fourth, all the growth of health disparities from 1984 to 2007 results from increases in the gradients of the relationship of SES-demographic statuses on within-group health disparities (the contribution is 179%), while decreases in between-group variance and changes in population compositions offset the former effect to some extent.[6]

Panel B of Table 2 further portrays how the three segments of SES status (i.e., income, education, and employment status) affect total variance by examining the changes in adjusted variances when a particular factor was fixed at the 1984 level. The numbers without parentheses are counterfactual changes in variance estimated by fixing a particular component at the 1984 level. The difference of observed change in variance and adjusted change in variance is the amount of change in observed variance associated with each specific component. The numbers in parentheses are percents of change explained by the corresponding component. For example, by fixing the college education attainment effect on between-group variance β at the 1984 level, the adjusted change in variance is 0.932. Thus, the amount of change in observed variance associated with this effect is 0.039 (=0.971−0.932), which means this effect accounts about 4.1% (0.039/0.971) of the increase in variance from 1984 to 2007. Overall, the changing gradients of SES-demographic effects on within-group health disparities contribute much more to change in total disparities than changing gradients of their effects on between-group health disparities, e.g., 108.4% vs. 4.1% for the college education effect, 37.3% vs. 17.9% for the middle and upper income effect, 27.6% vs. −4.9% for the employment status effect.

This means that increasing gradients of SES-demographic effects on with-group disparities are the main engines of widening health disparities for men. More specifically, education plays a more important role than income. College effects contribute about 112.5% ((0.971 + 0.121)/0.971) to the increase of total disparities, while middle and upper income effects account for about 55.2% ((0.971−0.435)/0.971). Health disparities between employed and unemployed men decrease, while disparities within each employment status increase, which may be due to an increasing underemployment rate within employed status.[7] As for the contributions of population composition to change in total variance, although increasing college attainment slows down the increase in health disparities, decreasing middle and upper income attainments and a decreasing employment rate further increase the health disparities.

---

[6] The "real" compositional effect may be underestimated as explained in notes below Table 2.
[7] In this paper, employed status includes both full-time job and part-time job.
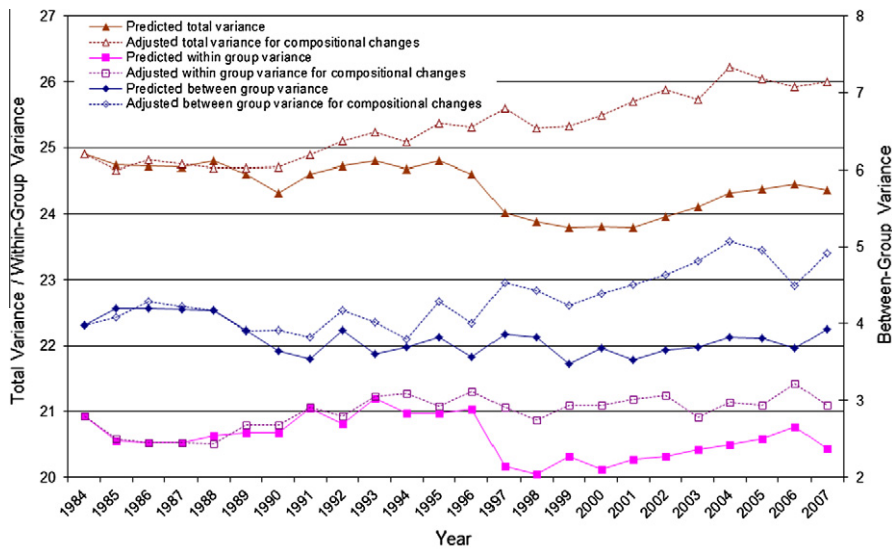
1. Adjusted for compositional changes: population compositions fixed at 1984 level.
2. Taking into account the possible increase in variance caused by sampling redesign after 1995, the sharp decrease in within-group disparity from 1996 to 1997 would be smoother since disparity should be lower in 1995 and 1996 than as shown in above figure.

Fig. 3. Decomposition of total variance into between-group variance and within-group variance, US Women, 1984–2007.

6.2. Disparities in women's self-reported health, 1984–2007

Fig. 3 displays the trend in self-reported health disparities for women from 1984 to 2007. Overall, total health disparity for women decreased from 1984 to 2007, but, as with men, this overall trend contains periods of increases and decreases. Specifically, disparity decreased from 1984 to 1990, rose until around 1995, decreased afterwards, and then rose again after 2000.[8] These period-specific trends are similar to those for men, but men's total health disparity increased overall from 1984 to 2007 while women's shows a slight decline. For women, in addition, both between-group health disparity and within-group health disparity decreased overall from 1984 to 2007. Similarly to men, within-group variance accounts much more than between-group variance as a percentage of total variance, 84% and 16% on average, respectively. The dotted lines represent the trends of variances adjusted for compositional changes which are induced by population composition changes and possibly by sampling design changes in the NHIS surveys as well. If population compositions were fixed at 1984 levels, all the three variances would have increased over time. This suggests compositional changes made a substantial contribution to the declining health disparities.

Panel A of Table 3 separates SES-demographic and compositional effects on the growth of health disparities from 1984 to 2007. There are four main findings from Panel A. First, generally, changes in within-group variance account for more of changes in total variance than changes in between-group variance. Second, different from men, SES-demographic effects on women's overall between-group variance significantly increased from 1984 to 2007, which, however, was outweighed by compositional effects, leading to overall decreased between-group health disparity. Third, similar to men, the SES-demographic effects on women's within-group variance increased from 1984 to 2007, but to a much smaller extent than for men. Compositional effects far outweigh the SES-demographic effects, which results in overall decreased within-group health disparity for women. Fourth, all of the reduction of women health disparities from 1984 to 2007 results from changes in population composition (as indicated by dotted lines in Fig. 3), while changes in gradients of socioeconomic statuses on between-group and within-group variance offset compositional effects to some extent.[9]

Panel B of Table 3 further explains how the three dimensions of SES effects and population compositions affect trends in health disparities for women. As for men, an increasing gradient of the effect of college education on health increases health disparities for women, which accounts for about −14% of the decrease in health disparities. After controlling for other SES-demographic factors, however, the gradients of middle and upper income and employment status effects on women's health variance decrease over time, which thus contribute about 29% and 22% to the decrease in health disparities, respectively. An articulation of the causes of the decreases of the gradients of these two factors over time is beyond the scope of this paper and merits further research. As expected, increasing college attainment, middle and upper income attainments, and employment rate reduce the total variance in probabilities of reporting very good/excellent health. They account for about 52.6%, 96.9%, and 22.2% of the decline of total health disparities, respectively.

---

[8] See footnote 5.
[9] The "real" compositional effect may be underestimated as explained in notes below Table 3.

Table 3
Decomposition and standardization of the change in variance in probabilities of reporting very good/excellent health, US Women, 1984–2007.

Panel A. Decomposition[a]		
Change in variance	−0.535	
Between-group	−0.049	*(9.2)*
SES-Demographic effect	0.942	*(−176.2)*
Compositional effect	−0.991	*(185.3)*
Within-group	−0.486	*(90.8)*
SES-Demographic effect	0.162	*(−30.3)*
Compositional effect	−0.648	*(121.1)*
Sum of SES-Demographic effects	1.104	*(−206.5)*
Sum of compositional effects	−1.639	*(306.5)*
Panel B. Adjusted change in variance, fixing at 1984		
College effect	−0.611	*(−14.2)*
Middle and upper income effect	−0.378	*(29.4)*
Employment status effect	−0.416	*(22.1)*
College attainment	−0.254	*(52.6)*
Middle and upper income attainment	−0.017	*(96.9)*
Employment status composition	−0.416	*(22.2)*

[1] Numbers in parentheses are percent of change explained by corresponding factors.
[2] Compositional effects here include compositional changes induced by population composition changes in the reality and sampling design changes in the NHIS surveys. Since sampling redesign may increase variance, the real compositional effects should be more negative. In other words, the compositional effects here may underestimate the real compositional effects.
   [a] All the covariates, including income, education, employment status, age, race and marital status, are included in the analysis. Panel A describes the total effects from these six SES-demographic characteristics. Panel B specifies the effects from education, income, and employment status.

7. Discussion and conclusions

Health disparities have grown considerably in recent decades, at least for men. The sources of these increases have not been clear, however. Much previous research focuses on widening health disparities across socioeconomic status levels by looking at how the gradients of effects of SES on health changes over time (e.g., Meara et al., 2008; Papps et al., 1993). But population-wide health disparities are affected by changes in population composition as well. Previous research also focused on health disparities across groups with different SES-demographic characteristics (i.e., between-group disparities), but health disparities within each group with the same SES-demographic characteristics (i.e., within-group disparities) can exceed between-group disparities. By applying variance function regression models and demographic decomposition methods to the analysis of health outcomes, we are able to separate SES-demographic effects from compositional effects and between-group disparities from within-group disparities.

Applied to gender-specific self-rated health in the NHIS data from 1984 to 2007, these methods identify sources of the different health disparities trends for men and women. In this study, we focus on the effects of three segments of SES characteristics (i.e., education, income and employment status) and their corresponding population composition on health disparities trends. We find both men's and women's health disparities decreased from 1984 to 1990, rose until around 1995, decreased afterwards, and then rose again after 2000. But, across the whole time period from 1984 to 2007, men's health disparities increased, while women's disparities decreased.

Consistent with previous research (e.g., Preston and Elo, 1995; Feldman et al., 1989), we find the education and income differentials in men's health increase. But widening men's health disparities are largely driven by rising within-group health disparities, which are moderated by the equalizing effects of increasing rates of college attainment to some extent. However, decreasing middle and upper income attainment and a decreasing employment rate (Juhn and Potter, 2006) further increase health disparities among men.

For women, the gradients of SES-demographic statuses on both between-group and within-group health disparities increase over time. For example, the education differential in self-rated health increases, which is consistent with the findings reported in Liu and Hummer (2008). This, however, was outweighed by changing population composition, e.g., increasing college attainment (Buchmann and DiPrete, 2006), middle and upper income attainment, and employment rate (Juhn and Potter, 2006). This leads to overall declining health disparities between 1984 and 2007 for women. But if increases in college attainment and labor participation rates slow down in the future, women's health disparities may also increase like those of men in the past two decades.

This study proposes a new concept, within-group health disparities – disparities in health outcomes within each group defined by the same SES-demographic characteristics. Results suggest within-group health disparities have dramatically increased over the time period studied, especially for men. This poses a question: why did the health disparities among people with the same SES-demographic characteristics increase? Now that this phenomenon of increasing within-group health dis-

parities has been identified, its full study and explanation can become a major topic of future research. Recent studies find within-group income instability has increased (Western et al., 2008) due to deinstitutionalization of the labor market, including declining unionization, minimum wages, long-term employment and increasing numbers of lower-wage, insecure jobs (McCall, 2000; Freeman, 1993; DiNardo et al., 1996), which particularly exists at the bottom of the wage and education distribution (DiNardo et al., 1996; Blau and Kahn, 1996). In addition, globalization (e.g., increased international trade and immigration) contributes to economic insecurity through declining relative demand and increasing demand elasticity for local low-skill workers, who are substituted by their counterparts across borders (Rodrik, 1997). How rising within-group health disparities may interact with increasing within-group income instability merits further studies.

This study has several limitations. First, it focuses on trends in self-rated health disparities. Although self-rated health has been shown to be a valid measurement of overall objective health, it should be emphasized that health is not a singular condition. Thus, further research is needed to examine trends in health disparities associated with other health outcomes, like functional health status (e.g., as defined by limitations in activities of daily living), morbidity (including the presence of specific diseases), and mortality. The conceptual framework developed herein – which emphasizes two dimensions to explore health disparities: SES-demographic effects vs. compositional effects, between-group health disparities vs. within-group health disparities – can fruitfully guide such future studies. Second, this study does not take into account cohort effects, which have been found in some previous studies (e.g., Yang and Lee, 2009; Lynch, 2003; Lauderdale, 2001).[10] Future studies can fill in this gap by applying the conceptual framework developed here to longitudinal cohort data. Third, in these initial analyses the objective of which is to show that the methods yield substantively interesting findings, we have focused on trends in disparities for groups of men and women defined by three SES-demographic statuses (family income, education, and employment status). There are many opportunities to follow these analyses in other studies that examine groups defined by other SES-demographic statuses, e.g., race, ethnicity, marital status, age, and occupation status and/or that focus on specific sub-populations.

Eliminating health disparities is one of the most important goals in public health fields, which however have exclusively focused on the health disparities between different SES groups. In the past two decades, health disparities have dramatically increased within the groups with the same SES characteristics too. How to interpret this phenomenon and how to eliminate within-group health disparities should be an equally important task for public health scholars. The effects of population composition on health disparities have been largely neglected in the existent literature, which actually makes a difference in the divergent self-rated health disparities among men and women. Therefore, changing population composition, e.g., increasing higher education attainment, increasing labor participation, and increasing social mobility should be important approaches to reduce health disparities.

Acknowledgment

References

Adler, Nancy E., Newman, Katherine, 2002. Socioeconomic disparities in health: pathways and policies. Health Affairs 21 (2), 60–76.
Aitkin, Murray, 1987. Modeling variance heterogeneity in normal regression using GLIM. Applied Statistics 36, 332–339.
Blau, Francine D., Kahn, Lawrence M., 1996. International differences in male wage inequality: institutions versus market forces. Journal of Political Economy 104, 791–837.
Bosworth, Hayden B., Siegler, Ilene C., Brummett, Beverly H., et al, 1999. The association between self-rated health and mortality in a well-characterized sample of coronary artery disease patients. Medical Care 37, 1226–1236.
Buchmann, Claudia, DiPrete, Thomas A., 2006. The growing female advantage in college completion: the role of family background and academic achievement. American Sociological Review 71 (4), 515–541.
Case, Anne, Paxson, Christina, 2005. Sex difference in morbidity and mortality. Demography 42, 189–214.
Crimmins, Eileen M., Saito, Yasuhiko, 2001. Trends in healthy life expectancy in the United States, 1970–1990: gender, racial, and educational differences. Social Science and Medicine 52, 1629–1641.
DiNardo, James, Fortin, Nicole M., Lemieux, Thomas, 1996. Labor market institutions and the distribution of wages, 1973–1992. Econometrica 64, 1001–1044.
DiPrete, Thomas A., Buchmann, Claudia, 2006. Gender-specific trends in the value of education and the emerging gender gap in college completion. Demography 43 (1), 1–24.
Farmer, Melissa M., Ferraro, Kenneth F., 2005. Are racial disparities in health conditional on socioeconomic status? Social Science & Medicine 60, 191–204.
Feldman, Jacob J., Makuc, Diane M., Kleinman, Joel C., Cornoni-Huntley, Joan, 1989. National trends in educational differences in mortality. American Journal of Epidemiology 129 (5), 919–933.
Freeman, Richard B., 1993. How much has deunionization contributed to male earnings inequality. In: Sheldon, Danziger, Gottschalk, Peter (Eds.), Uneven Tides: Rising Inequality in America. Russell Sage Foundation, New York, pp. 133–163.
George, Linda K., 2005. Socioeconomic status and health across the life course: progress and prospects. Journal of Gerontology: Social Sciences 60B (special issue), 135–139.
Glied, Sherry, Lleras-Muney, Adriana, 2008. Technological Innovation and Inequality in Health. Demography 45 (3), 741–761.
Goesling, Brian, 2007. The rising significance of education for health? Social Forces 85 (4), 1621–1644.
Hamil-Luker, Jenifer, 2005. Women's wages: cohort differences in returns to education and training over time. Social Science Quarterly 86, 1261–1780.
Hellevik, Ottar, 2009. Linear versus logistic regression when the dependent variable is a dichotomy. Quality and Quantity 43, 59–74.

---

[10] We have explored the incorporation of cohort into our analysis. This substantially increases the number of cells in the analysis (to 1728 for 4 cohorts) and substantially increases the number of cells with no observations as age groups and cohort groups do not overlap in many cells.

Hummer, Robert A., Rogers, Richard G., Eberstein, Isaac W., 1998. Sociodemographic approaches to differentials in adult mortality: a review of analytic approaches. Population and Development Review 24 (3), 553–578.

Iachine, Ivan A., Holm, Niels V., Harris, Jennifer R., et al, 1998. How heritable is individual susceptibility to death? The results of an analysis of survival data on Danish, Swedish and Finnish Twins. Twin Research 1, 196–205.

Idler, Ellen L., Benyamini, Yael, 1997. Self-rated health and mortality: a review of 27 community studies. Journal of Health and Social Behavior 38, 21–37.

Jemal, Ahmedin, Ward, Elizabeth, Anderson, Robert N., et al, 2008. Widening of Socioeconomic Inequalities in US death rates, 1993–2001. PLOS ONE 3 (5), e2181.

Juhn, Chinhui, Potter, Simon, 2006. Changes in labor force participation in the United States. Journal of Economic Perspectives 20 (3), 27–46.

Kennedy, Byron S., Kasl, Stanislav V., Vaccarino, Viola, 2001. Repeated hospitalizations and self-rated health among the elderly: a multivariate failure time analysis. American Journal of Epidemiology 153, 232–241.

Kochanek, Kenneth D., Maurer, Jeffrey D., Rosenberg, Harry M., 1994. Why did black life expectancy decline from 1984 through 1989 in the United States? American Journal of Public Health 84, 938–944.

Krieger, Nancy, Fee, Elizabeth, 1994. Social class: the missing link in US health data. International Journal of Health Service 24 (1), 25–44.

Lauderdale, Diane S., 2001. Education and survival: birth cohort, period, and age effects. Demography 38 (4), 551–561.

Lemieux, Thomas, 2006. Increasing residual wage inequality: composition effects, noisy data, or rising demand for skill? American Economic Review 96 (3), 461–498.

Link, Bruce G., Phelan, Jo, 1995. Social conditions as fundamental causes of disease. Journal of Health and Social Behavior (Extra issue), 80–94.

Liu, Hui, Hummer, Robert A., 2008. Are educational differences in US self-rated health increasing? An examination by gender and race. Social Science & Medicine 67, 1898–1906.

Liu, Hui, Umberson, Debra J., 2008. 'The Times They Are a Changin': marital status and health differentials from 1972 to 2003. Journal of Health and Social Behavior 49 (3), 239–253.

Lynch, Scott M., 2003. Cohort and life-course patterns in the relationship between education and health: a hierarchical approach. Demography 40 (2), 309–331.

Manor, Orly, Matthews, Sharon, Power, Chris, 2000. Dichotomous or categorical response? Analysing self-rated health and lifetime social class. International Journal of Epidemiology 29, 149–157.

McCall, Leslie, 2000. Explaining levels of within-group wage inequality in US labor markets. Demography 37 (4), 415–430.

Meara, Ellen.R., Richards, Seth, Cutler, David M., 2008. The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000. Health Affairs 27 (2), 350–360.

National Center for Health Statistics, 1994. Excess deaths and other mortality measures for the black population: 1979–1981 and 1991. Pub. Health Serv, Hyattsville, MD.

Papps, Gregory, Queen, Susan, Hadden, Wilbur, Fisher, Gail, 1993. The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. New England Journal of Medicine 329, 103–115.

Powers, Daniel A., Xie, Yu, 2008. Statistical Methods for Categorical Data Analysis, second ed. Emerald, Howard House, England.

Preston, Samuel H., Elo, Irma T., 1995. Are educational differentials in adult mortality increasing in the United States? Journal of Aging and Health 7 (4), 476–496.

Rodrik, Dani, 1997. Has Globalization Gone Too Far? Institute for International Economics, Washington, DC.

Rogers, Richard.G., 1992. Living and dying in the USA: sociodemographic determinants of death among blacks and whites. Demography 29 (2), 287–303.

Schnittker, Jason, 2007. Working more and feeling better: women's health, employment, and family life, 1974–2004. American Sociological Review 72, 221–238.

Schoeni, Robert F., Martin, Linda G., Andreski, Patricia, Freedman, Vicki A., 2005. Persistent and growing socioeconomic disparities in disability among the elderly: 1982–2002. American Journal of Public Health 95 (11), 2065–2070.

Smith, Ken R., 1996. All-cause and cause-specific mortality and marital status: findings from the NHIS-NDI Merged File, 1986-91. In: Paper Presented at the Annual Meeting of the Population Association of America, 9–11 May, New Orleans.

Steenland, Kyle., Hu, Sherry, Walker, James, 2004. All-cause and cause-specific by socioeconomic status among employed persons in 27 US states, 1984–1997. American Journal of Public Health 94 (6), 1037–1042.

Western, Bruce, Bloome, Deidre, 2009. Variance function regression for studying inequality. Sociological Methodology 39, 293–325.

Western, Bruce, Bloome, Deirdre, Percheski, Christine, 2008. Inequality among American families with children, 1975–2005. American Sociological Review 73, 903–920.

William, David R., Collins, Chiquita, 1995. US Socioeconomic and racial differences in health: patterns and explanations. Annual Review of Sociology 21, 349–386.

Yang, Yang, Lee, Linda C., 2009. Sex and race disparities in health: cohort variations in life course patterns. Social Forces 87, 2093–2124.

Yashin, Anatoli I., Iachine, Ivan A., Harris, Jennifer R., 1999. Half of the variation in susceptibility to mortality is genetic: findings from Swedish twin survival data. Behavior Genetics 29 (1), 11–19.

Zimmer, Zachary, House, James S., 2003. Education, income, and functional limitation transitions among American adults: contrasting onset and progression. International Journal of Epidemiology 32, 1089–1097.