Income inequality and self-rated health in the United States: Does education or race explain the link?

James E. Foster, Ph.D., Professor, Achintya Ray, Ph.D., Assistant Professor

a Department of Economics, Vanderbilt University, and Vanderbilt Institute for Public Policy Studies, Nashville, TN, USA, 37235, email: james.e.foster@vanderbilt.edu

b Department of Economics and Finance and Office of Business and Economic Research, College of Business, College of Business, Tennessee State University, Suite K32, 330 Tenth Avenue North, Nashville, TN, 37203, USA, email: aray4@mytsu.tnstate.edu

Correspondence to: Dr. James Foster james.e.foster@vanderbilt.edu

Abstract
Objective: To evaluate whether the ecological variables of income inequality, educational attainment, and racial composition have independent effects on self-rated health, controlling for individual characteristics.
Design: Multilevel, multivariate logistic regression analysis.
Setting: All US states and the District of Columbia.
Main outcome measure: Low self-rated health status (fair or poor) is the dependent variable. Independent variables at the state level include mean income, Gini coefficient, percentage of adults (age 18 or above) attaining less than 12 years of formal education, and fraction of the population that is black. Independent individual level variables include income, education achievement, race, and age.
Results: The percentage attaining less than 12 years of education is shown to be the only ecological variable that exerts independent impact on self-rated health in the presence of individual level variables.
Conclusions: The distribution of human capital, as measured by the percentage without a high-school diploma, is a significant marker of health whose impact remains strong even in the presence of other ecological and individual variables. Income inequality and racial composition in US states are not found to be independent health hazards.
Introduction

Many ecological studies have found a positive correlation between income inequality and mortality in US states.\(^1\)\(^-\)\(^3\) This relationship has potentially important implications for health policy, and several plausible mechanisms have been proposed to explain it.\(^4\)\(^-\)\(^6\) At the same time questions have arisen as to whether this aggregate relationship might be spurious or due entirely to omitted individual level variables. Gravelle, for example, provided a model in which the health consequences of inequality are due to a curvilinear relationship between individual income and health.\(^7\) To determine whether inequality has salience for health when individual characteristics are taken into account, aggregate methods have given way to multilevel analyses, which employ both ecological and individual variables to explain individual levels of health. Ecological variables are defined to be the ones that affect everybody. While the multilevel evidence is mixed, it appears that there is indeed a modest independent impact of inequality on self-rated health in US states.\(^8\)\(^,\)\(^9\)

Two recent studies have challenged the inequality-health hypothesis as arising from omitted collective variables. Muller\(^10\) introduces a state-level indicator of educational attainment; Deaton and Lubotsky\(^11\) consider racial composition. Each finds that the introduction of a state-level variable causes income inequality to lose its significance as a predictor of health. Both provide credible bases for questioning the inequality-health relationship. However, as group level analyses, they are subject to the same weakness that faced the original ecological inequality-health studies.

Can their conclusions survive a thoroughgoing multilevel analysis? Or will income inequality continue to be viewed as a significant determinant of individual health? To resolve this open question, we conducted a multilevel study to evaluate the independent effects of US state income inequality, educational attainment, and racial composition on morbidity (as indicated by low self-rated health).

Methods

We used data from the 1995 US Current Population Survey’s (CPS) March supplement downloaded using the CPS on Web facility (http://www.unicon.com). CPS is conducted jointly by the US Bureau of Labor Statistics and the Bureau of Census as a representative sample (at national and state levels) of the non-institutionalized, civilian, US population. Some 50,000 households (or 150,000 individuals) are sampled every month. It is widely used in social and economic analyses. The March supplement of CPS includes self-rated health status with five possible values: Excellent, Very Good, Good, Fair and Poor. Following Kennedy et al. we created a dichotomous variable with value 1 if an individual reported fair or poor health.\(^8\) We called this variable low health. We also calculated a state level variable age adjusted health poverty indicating the population percentage reporting low health adjusted by the 1990 Census age structure.

Income data are reported in the CPS at the household level. We excluded households with non-positive reported incomes. To reflect the differential needs of children, we constructed an adult equivalence scale that assigns a weight of 0.5 if a person is below age 16 and 1 otherwise. We constructed the equivalent income for each household member by dividing household income by the number of adult equivalents. This variable was used to calculate the mean income and the Gini coefficient for each state and the District of Columbia (DC). The Gini coefficient is a widely used measure of income inequality that can be interpreted as half the expected (absolute) difference between two randomly drawn incomes, expressed as a share of the mean.\(^12\) It ranges between 0 and 1 with a higher value indicating more income inequality. Equivalent and mean incomes were
measured in thousands of dollars. The Gini coefficient was expressed in percentage points from 0 to 100.

We created a dichotomous *race* variable that takes value 1 if an individual reported race as black. We computed the percentage of black residents in each state and DC and we denoted this as the *fraction black*. This is analogous to the variable used by Deaton and Lubotsky. For persons 18 or older we created a dichotomous variable *education* that takes value 1 if the individual reported completing grade 12. We computed for each state and DC the percentage of residents aged at least 18 not completing grade 12, and we called this variable *education poverty*. It is analogous to the variable used by Muller. Other individual level variables employed in our multilevel analyses are: the *age* of a person in number of completed years and *gender*, a dichotomous variable that takes value 1 if female.

To account for the sampling framework, we used observation weights provided by CPS to construct aggregate variables except for the Gini coefficient. Note that mean income, the Gini coefficient, and the fraction black are constructed from the all-age sample population (148,739 observations), while education poverty includes only the population aged 18 or older (107,285 observations).

The multilevel study was estimated as a logistic regression of *low health* on aggregate and individual variables for the 18 and older population (to ensure that the education variable is meaningful). For comparison with the existing aggregate level results, we also conducted analyses using ordinary least squares regression for the 51 state observations. We used the STATA/SE 8.0 commands “logistic” and “regress” to implement the multilevel and ordinary regressions.

**Results**

Regression coefficients from the state level aggregate analyses are presented in Table 1 with p-values in parentheses. Specification A1 focuses exclusively on state income variables and finds that age adjusted health poverty is higher in states with higher inequality. In A2, inequality is not significant when the racial composition of states is included. When education poverty is added in A3, state level mean income, inequality and racial composition are not significant, while education poverty is significant and harmful for health. The fit of the model (measured by adjusted R-squared) increases somewhat from A1 to A2 and quite significantly from A2 to A3.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean income</td>
<td>-0.89 (0.000)</td>
<td>-0.92 (0.000)</td>
<td>-0.24 (0.180)</td>
</tr>
<tr>
<td>Gini coefficient</td>
<td>0.32 (0.013)</td>
<td>0.12 (0.469)</td>
<td>-0.15 (0.258)</td>
</tr>
<tr>
<td>Fraction black</td>
<td>0.07 (0.063)</td>
<td>-0.002 (0.939)</td>
<td></td>
</tr>
<tr>
<td>Education poverty</td>
<td></td>
<td></td>
<td>0.58 (0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>14.69 (0.014)</td>
<td>22.10 (0.002)</td>
<td>11.14 (0.053)</td>
</tr>
<tr>
<td>Adjusted R-Squared</td>
<td>0.40</td>
<td>0.43</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 1
Aggregate Analysis
(Dependent variable: *Age adjusted health poverty*)
The multilevel analysis results are presented in Table 2, which provides the odds ratios from the logistic equation and corresponding p-values in parentheses. In each specification, individual income and education reduce the relative likelihood of reporting low health, while age, race and gender increase it. M1 includes state income inequality as the key contextual variable and finds that it is indeed significant and harmful. M2 shows that when the fraction black variable is included, the significance of the inequality variable and the magnitude of its effect are maintained, while fraction black is not significant. Our main specification, M3, finds that inequality and racial composition are not significant when the state education variable is included. Moreover, education poverty is significant and its impact is not unsubstantial: a one percentage point decrease in a state’s level of education poverty decreases an average resident’s likelihood of reporting low health by 3%, roughly equivalent in impact to a $750 increase in the resident’s income.

Table 2
Multilevel Analysis
(Dependent Variable: Low Health)

<table>
<thead>
<tr>
<th>Explanatory Variables</th>
<th>Specification M1</th>
<th>Specification M2</th>
<th>Specification M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.05 (0.000)</td>
<td>1.05 (0.000)</td>
<td>1.05 (0.000)</td>
</tr>
<tr>
<td>Race</td>
<td>1.71 (0.000)</td>
<td>1.68 (0.000)</td>
<td>1.69 (0.000)</td>
</tr>
<tr>
<td>Gender</td>
<td>1.09 (0.000)</td>
<td>1.09 (0.000)</td>
<td>1.09 (0.000)</td>
</tr>
<tr>
<td>Income</td>
<td>0.96 (0.000)</td>
<td>0.96 (0.000)</td>
<td>0.96 (0.000)</td>
</tr>
<tr>
<td>Education</td>
<td>0.48 (0.000)</td>
<td>0.48 (0.000)</td>
<td>0.48 (0.000)</td>
</tr>
<tr>
<td>Mean Income</td>
<td>0.94 (0.000)</td>
<td>0.94 (0.000)</td>
<td>0.97 (0.000)</td>
</tr>
<tr>
<td>Gini Coefficient</td>
<td>1.01 (0.003)</td>
<td>1.01 (0.037)</td>
<td>0.99 (0.114)</td>
</tr>
<tr>
<td>Fraction Black</td>
<td></td>
<td>1.002 (0.156)</td>
<td>0.99 (0.417)</td>
</tr>
<tr>
<td>Education Poverty</td>
<td></td>
<td></td>
<td>1.03 (0.000)</td>
</tr>
</tbody>
</table>

Table 3 presents odds ratios and p-values from multilevel analyses for various population subgroups. The independent effect of education poverty on health holds for all income and educational classes considered and for whites as a group. However, education poverty is not a significant ecological variable for black residents, although individual education is significant and protective for all of the groups.
Table 3
Multilevel Analysis for various population subgroups

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Population Cell</th>
<th>Income below $10K</th>
<th>Income $10-20K</th>
<th>Income above $20K</th>
<th>Education below 12 years</th>
<th>Education 12 years &amp; above</th>
<th>Race White</th>
<th>Race Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td>1.05 (0.000)</td>
<td>1.05 (0.000)</td>
<td>1.06 (0.000)</td>
<td>1.05 (0.000)</td>
<td>1.05 (0.000)</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td>0.91 (0.000)</td>
<td>0.95 (0.000)</td>
<td>0.99 (0.000)</td>
<td>0.95 (0.000)</td>
<td>0.97 (0.000)</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td>0.55 (0.000)</td>
<td>0.54 (0.000)</td>
<td>0.51 (0.000)</td>
<td>…</td>
<td>…</td>
<td>0.47</td>
<td>0.55</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td>1.08 (0.011)</td>
<td>1.07 (0.032)</td>
<td>1.08 (0.081)</td>
<td>1.11 (0.000)</td>
<td>1.09 (0.002)</td>
<td>1.08</td>
<td>1.09</td>
</tr>
<tr>
<td>Education Poverty</td>
<td></td>
<td>1.03 (0.000)</td>
<td>1.04 (0.000)</td>
<td>1.02 (0.000)</td>
<td>1.03 (0.000)</td>
<td>1.03 (0.000)</td>
<td>1.03</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Discussion
We found that state educational attainment is a significant marker for self-rated health, even after controlling for individual characteristics, while income inequality and state racial composition are not. The contextual impact of education is evident for population groups stratified by income or education, for whites as a group, but not for black residents.

This was the first multilevel study of the inequality hypothesis to include state variables for education and race. It employed a large US dataset representative at both state and national levels and used as its key outcome variable self-rated health, a well-established indicator of morbidity for multilevel studies. While rich in socioeconomic characteristics, the data do not contain medical and behavioral variables nor do they allow comparisons through time. Certain specifications used in the study – such the education cutoff, the income and age variables – are to some extent arbitrary. Robustness analyses (not reported here) varied the cutoff, included age-squared, and replaced income with log income. In each instance the conclusions about education poverty and racial composition were substantively unchanged, while inequality either remained insignificant or was not harmful.

In a key multilevel study, Kennedy et al.\(^8\) obtained results suggesting that income inequality is a significant marker for health. We replicated these results and noted that they arise due to the exclusion of the contextual variable, education poverty. Additionally, they found that their contextual variable, inequality, has no effect at higher income levels. Our study found education poverty to be significant for each income (and education) range.

Deaton and Lubotsky concluded that it is the racial composition of states and metropolitan areas – not inequality – that is important for health.\(^11\) When we excluded education poverty, we obtained similar findings at the aggregate level, but not in the multilevel study. This suggests that individual level confounders may account for some of the differences between our results – although our use of self-rated health instead of mortality may also be relevant. Another of their main conclusions is that white health is negatively correlated with the variable fraction black. In a
multilevel analysis (not reported here) we found a similar result when we excluded education poverty; however, including this variable renders racial composition insignificant for whites. One difference is that while Deaton and Lubotsky control for education, they do not include a broader contextual variable like education poverty in their analysis.

Finally, our results support Muller’s findings that state education poverty is significantly related to health and income inequality has no independent impact. However, unlike Muller, we obtained this conclusion in the presence of individual level confounders. It is well known that health and educational attainment are related at the individual level. Muller’s aggregate level relationship could simply be a composition of individual effects. Our results show that state education is contextual, a possibility neither emphasized nor established by Muller.

Our results should not be interpreted as saying that inequality and racial composition are unrelated to health. Greater education poverty is likely to be associated with higher income inequality via the labor market’s link between earnings and education. Given the higher dropout rates for black students, education poverty could well reflect the racial composition of a state. However, our results do suggest that education poverty is better at accounting for variations in health achievements and accordingly may have greater relevance then other collective variables.

There are several plausible explanations of our result that education has contextual import. Many studies have shown that education has significant spillover effects and our results provide further evidence of this. Education poverty may be symptomatic of low investment in other public goods and hence a worse environment for health, especially for persons with limited opportunities. Group educational attainment is empirically related to the important – yet difficult to measure – concept of social capital, which in turn has been linked to health. School is where much social capital is formed and education is a strong predictor of subsequent involvement in social organizations. Note, though, that social capital is not simply the connections between people; it depends positively on their collective capabilities. The distribution of education or human capital is an important, if overlooked, dimension of social capital. Our results may be viewed as further evidence of the relationship between social capital and health and suggest that additional research into education’s role may be fruitful. We have no explanation for our finding that black residents do not share in the ecological benefits of higher state educational attainment. Further research into the differential health impact of education across racial groups may help in understanding racial inequalities in health.

Our results provide support for policies to improve educational achievements, especially among those at risk for not completing high school, such as “No Child Left Behind” or “Head Start”. At the other end of the spectrum, it should encourage health professionals to consider the extent of education, both at the individual and group levels, in designing appropriate care strategies.

Acknowledgments
We thank Dr. William Collins for helpful comments.
Funding source: JEF is a recipient of the Robert Wood Johnson Foundation Investigator Award in Health Policy Research and AR was the research assistant on the Award.
Conflict of interest: None
References
13. Pincus T, Callahan LS, Burkhauser RV. Most chronic diseases are reported more frequently by individuals with fewer than 12 years of formal education in the age 18-64 United States population. J Chron Dis 1987; 40: 865-74.
19. Glaeser E. The formation of social capital. ISUMA 2001 (Spring); 2. (http://www.isuma.net/v02n01/glaeser/glaeser_e.shtml)