NBER WORKING PAPER SERIES

MORTALITY, INEQUALITY AND RACE IN AMERICAN CITIES AND STATES

Angus Deaton
Darren Lubotsky

Working Paper 8370
http://www.nber.org/papers/w8370

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2001

We are grateful to Anne Case and Christina Paxson for helpful discussions during the preparation of this paper. We gratefully acknowledge financial support from the John D. and Catherine T. MacArthur Foundation through its Network on Poverty and Inequality in Broader Perspectives, and from the National Institute on Aging through a grant to the National Bureau of Economic Research. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Mortality, Inequality and Race in American Cities and States
Angus Deaton and Darren Lubotsky
NBER Working Paper No. 8370
July 2001
JEL No. I12

ABSTRACT

A number of studies have found that mortality rates are positively correlated with income inequality across the cities and states of the US. We argue that this correlation is confounded by the effects of racial composition. Across states and MSAs, the fraction of the population that is black is *positively* correlated with average white incomes, and *negatively* correlated with average black incomes. Between-group income inequality is therefore higher where the fraction black is higher, as is income inequality in general. Conditional on the fraction black, neither city nor state mortality rates are correlated with income inequality. Mortality rates are higher where the fraction black is higher, not only because of the mechanical effect of higher black mortality rates and lower black incomes, but because *white* mortality rates are higher in places where the fraction black is higher. This result is present within census regions, and for all age groups and both sexes (except for boys aged 1–9). It is robust to conditioning on income, education, and (in the MSA results) on state fixed effects, and cannot plausibly be attributed to variations in the local provision of health care.

Angus Deaton
Center for Health and Wellbeing
347 Wallace Hall
Princeton University
Princeton, NJ 08544
and NBER
Tel: 609-258-5967
Fax: 609-258-5974
deaton@princeton.edu

Darren Lubotsky
Center for Health and Wellbeing
339 Wallace Hall
Princeton University
Princeton, NJ 08544
Tel: 609-258-6917
Fax: 609-258-5974
lubotsky@princeton.edu

1. Introduction and background

In recent years there has been a great deal of interest in whether income inequality is a health hazard in the sense that individuals are less healthy in places where income is more unequally distributed. The strongest advocate of the income inequality hypothesis has been Richard Wilkinson (1992), (1996), (2000), who has put forward a variety of evidence, from individual, area, cross-country, and time-series data. A survey of the subsequent debate over this evidence is given in Deaton (2001b). In this paper, we are concerned with one of the most prominent of these relationships, the ecological association between income inequality and mortality across states and cities in the United States. One version of this correlation is shown in Figure 1 below, which plots (directly) age-adjusted all-cause mortality against the gini coefficient of per adult-equivalent income; the District of Columbia is included and, although it has both higher mortality and higher inequality than any state, it lies on the regression line. (The definitions of these and other data are given in the next section.) The positive correlation between income inequality and mortality across the US states was first shown in studies by Kaplan et al (1996) and Kennedy, Kawachi and Prothrow-Stith (1996a, b). Lynch, Kaplan and Pamuk (1998) reproduced the correlation using data from 282 Metropolitan Statistical Areas (MSAs) in 1990, finding that the loss of life from income inequality "is comparable to the combined loss of life from lung cancer, diabetes, motor vehicle crashes, HIV infection, suicide, and homicide in 1995." These, and other related studies, are collected in Kawachi, Kennedy, and Wilkinson (1999).

In this paper, we investigate the robustness of the connection between income inequality and mortality across states and MSAs, with particular attention to the effects of race as a potential confounder. That the spatial link between income inequality and mortality might be spurious is suggested by several studies in the literature. Most recently, Mellor and Milyo (2001) pool census data on the 48 continental states from 1950, 1960, 1970, 1980, and 1990 and find that the relationship between the gini coefficient and mortality is not robust to the inclusion of various plausibly important controls, such as education, urbanization, and race. There is also a lack of consistently positive evidence on the role of income inequality from follow-up data on individuals, in the National Longitudinal Mortality Study, where there is a relationship neither in the individual data nor in the state-level data (Deaton 2001a), from the National Health and Nutritional Examination Survey (Fiscella and Franks 1997), from the Panel Study of Income Dynamics (Daly et al 1998), and only relatively weak evidence from the National Health Interview Survey (Lochner et al 2001). Furthermore, a study of cities and provinces in Canada failed to find any relationship between income inequality and mortality, Ross et al (2000).

In the results presented below, we show that, once we control for the fraction of the population that is black, there is no relationship between income inequality and mortality across either states or cities. This result does *not* come from the pooling of black and white mortality; as emphasized in the earlier literature, the correlation between income inequality and mortality is present for each race separately. Instead, our results come from the fact that white mortality rates are higher in places where a higher fraction of the population is black. We do not have an explanation for this result, although we explore and rule out a number of possibilities. In particular, the relationship between the fraction black and white mortality rates holds within broad geographical regions, and so is not driven by a comparison of the South with the rest of the country. The correlation is also robust to the inclusion of controls for state fixed effects and for education, holds for nearly all age groups and for males and females, and cannot readily be attributed to variations in local health provision.

The paper is organized as follows. Section 2 discusses data sources and methodology. Section 3 presents the main results from the states and MSAs. Section 4 explores robustness and investigates some alternatives. We focus on mortality among whites though, in line with earlier literature, we also show some results for all races pooled. The results for other races are of considerable interest in their own right, but we confine our attention here to one element of the story, leaving for future work the comparative results. That the fraction black increases mortality rates for blacks is shown in Miller and Paxson (2001). Section 5 concludes.

2. Data and methodology

The data on mortality are taken from the *Compressed Mortality Files*, (CMF) from the National Center for Health Statistics at the Center for Disease Control. The CMF contain a complete census of all deaths by year from 1968 to 1994, by cause of death, race, sex, age group, and county of residence, except for Alaska where only state-level data are available. The CMF files also provide population totals for each cell, which we use to calculate mortality rates as well as racial composition. We use data on deaths in 1980 and 1990. The county identifiers are used to aggregate deaths and populations to the Metropolitan Statistical Area (MSA) level; once again, Alaska is an exception and is excluded from our MSA analysis. We match 287 MSAs in 1980 and 1990. Not everyone lives in an MSA; the 287 MSAs used here contain 79.9 and 80.7 percent of the total population of the U.S. in 1980 and 1990 respectively. The data aggregated by state cover the entire population of the U.S. The CMF data are disaggregated by 13 age groups; we preserve these age groups when aggregating to the state and MSA levels, and then calculate age-adjusted all-cause mortality rates by direct adjustment to the U.S. population in 1990. Age adjustment is done separately by sex, and separately for all races combined and for whites alone. Hence, for example, the age-specific mortality rates for white females in New York City are weighted by the age-distribution of white females in the U.S. population in 1990.

The creation of a consistent set of MSA mortality data requires a mapping of counties into MSAs, as well as a method of handling changes in the definitions of MSAs between the two years. MSAs are defined by the US Office of Management and Budget (OMB) and in some cases, their geographical boundaries changed from 1980 to 1990. MSAs are always collections of counties, except in New England where they are collections of cities and towns, so that counties may be split between multiple MSAs. Because the mortality data come at the county level, aggregating mortality to the MSA level is relatively straightforward outside of New England. Within New England we use New England County Metropolitan Areas, OMB's county-based alternatives to the city- and town-based MSAs.

Data for income and education are taken from the 5 per cent public-use samples of the 1980 (A sample) and 1990 censuses. Income data in the census refer to the previous year, i.e. 1979 or 1989, which is one year earlier than the mortality data from the CMF. Other choices of timing could be investigated, for example by averaging mortality over several years around the censuses, or by using mortality several years after each census, but given the arbitrariness of any choice, the one year lag seems as reasonable as any.

Census data do not come at the county level, but at the level of County Group in 1980 and PUMA ("Public Use Microdata Area") in 1990. The 1990 PUMAs do not necessarily match the 1980 County Groups, nor are they necessarily collections of counties. Instead, they can be parts of counties, single counties, collections of whole counties, or collections of parts of counties. Our procedure is to use the 1990 MSA definitions and create, as closely as possible, consistently defined metropolitan areas in 1980. We begin with the mapping of County Groups and PUMAs into MSA definitions given in Jaeger et al (1998) for cities with populations over 250,000 people. In 1990, 20 of the cities with populations over one million people are designated by OMB as CMSAs, essentially combinations of MSAs, and are treated as units by Jaeger et al. We split these CMSAs into their component cities, technically referred to as Primary Metropolitan Statistical Areas (PMSAs). For example, the Dallas-Fort Worth CMSA is composed of the Dallas PMSA and the Fort Worth-Arlington PMSA, and we treat each as one observation in our analysis of MSAs. We also include 110 smaller cities; these are defined by OMB, and are generally places with populations of at least 100,000 but less than 250,000. In the end, we have 287 MSAs consistently defined in 1980 and 1990. 110 of these are the MSAs in Jaeger et al, 54 come from our disaggregation of CMSAs, 110 are smaller cities that were not included by Jaeger et al, and there are 13 New England County Metropolitan Areas.

In some cases the 1980 County Groups and the 1990 PUMAs contain areas that are partly inside and partly outside of an MSA. For these, it is not possible to create an exact match between an area in 1980 and 1990, nor between Census and mortality data. In these cases, a judgment must be made as to whether to drop the unit, if it is impossible to make a reconciliation by aggregating up to a reasonably sized larger unit, or to include it, if the differences between the two years are small. Of our 287 MSAs, 237 contain identical counties in 1980 and in 1990. Of the 50 others, only a small fraction of the population lives in the areas that are included in only one year. For each MSA, we calculated the sum of the populations in the two years that lived in counties included in both years, and divided it by the sum of the total populations in the two years. The resulting ratio is unity for the 237 consistent MSAs. For the other 50, the mean of the ratio is 93.2 percent, the median is 94.7 percent, and the minimum is 71.9 percent. The definitions of our MSAs, and their relationship to counties, County Groups, and PUMAs is detailed in an Addendum to this paper that is available at http://www.wws.princeton.edu/~chw

Each individual in the census is assigned an MSA according to the rules discussed above. Each is also assigned the adult equivalent household income for the household in which he or she lives, where equivalent income is calculated by dividing total household income by the number of adults plus half the number of children, defined as household members aged 18 and younger. Logarithms of income and of income per equivalent are calculated at the individual level, and averaged over MSAs and states. Income from the 1980 census—which relates to 1979—is converted to 1989 prices using the CPI in order to make it comparable with data from the 1990 census. We make no attempt to deal with top-coding. The gini coefficient is calculated on an individual basis, using income per equivalent adult imputed to each individual. We calculate gini coefficients and income levels separately by race and by sex, as well as over all races and both sexes. Note that if all households consisted of a male and female couple, and because the same per equivalent income is imputed to each, the male and female ginis would be identical. Although this is not the case, the cross-MSA correlation between the (white) male and female ginis is 0.97 in 1980 and 0.95 in 1990. For each individual we also record an indicator for the level of education achieved according to five categories; less than high school, high school, some college (education post high-school, but without a bachelor's degree), completed college, post-graduate education (in 1980, more than 16 years of education, in 1990 holding a master's, professional, or doctoral degree.) The binary indicators are averaged within states and MSAs, for people aged 25 and above, again separately by sex and race. This gives us data, for example, on the fractions of adult men or women in Ohio or in Dallas whose highest education is in each of the five categories.

In the results that follow we use OLS regressions with either state or MSA-level data. The dependent variable is an age-adjusted mortality rate converted to a log odds. The independent variables are area averages of the explanatory variables, such as the logarithm of income per equivalent, or state or MSA-wide estimates of the gini coefficient, racial composition, or the fractions of the population whose highest level of education is in each of the education classes. Each regression is weighted by the square root of the population at risk in each state or MSA.

3. Basic results for states and MSAs

Table 1 shows results from the state data, including the District of Columbia, and pooling data from 1980 and 1990, so that there are 102 observations in each regression. All regressions include a dummy variable for 1990; if there is a decline in mortality rates that is unexplained by the included variables, the regression coefficient on the dummy should be negative, as is always in fact the case. The first two columns in the left-hand panels, for all males and all females irrespective of race, show the results that are typically reported in the literature. In the first regression, with no other variables included, the logarithm of per adult equivalent income has a protective effect that is about twice as large for males as for females, –0.22 versus –0.09. The 1990 dummy has a coefficient of –0.11 for men and –0.07 for women so that there is a background improvement in mortality that is not explained by changes in income, or at least cannot be explained by assigning the same effect to income over time as it is estimated to have over states. The second column shows the effects of including income inequality in the form of the gini coefficient of income per equivalent adult. The gini coefficient attracts large and significant positive coefficients for both males and females. Over the 51 states in 1990, the mean of the gini of per equivalent income was 0.37 with a standard deviation of 0.02, so if we move from one standard deviation below the mean to one above the mean, from Vermont to Mississippi, or from Michigan to Florida, the log odds increases by 0.057 for men and by 0.043 for women, corresponding to relative risks of (approximately) 1.06 for men and 1.04 for women. The coefficient on income is not significantly different from zero in these regressions. The coefficients on the 1990 dummy are larger than before. Mortality declined from 1980 to 1990 while income inequality increased, so that the hazardous effects of inequality that are estimated from the interstate differences must be offset by the time dummy.

That the estimated effects of income inequality are potentially confounded by the effects of race has been recognized since the first papers on the topic. Blacks have higher mortality rates than whites and, on average, have lower incomes, so that in places with a substantial black population, both income inequality and mortality tend to be higher. That there is some such problem is shown by the third column in the first panel. When the fraction of the state population that is black is added to the regressions, it attracts a significantly positive coefficient, and the coefficient on the gini coefficient is no longer significantly different from zero. But this regression does little more than illustrate that there is a problem with the first two columns. Indeed, as noted by Kaplan et al (1996), separate regressions by race find that income inequality is estimated to be a hazard for each.

The results for whites alone are shown in the right-hand panel of Table 1.The coefficient on the gini coefficient in the second column of the right hand panel is a good deal smaller for whites than it was for all races taken together, and for women the effect is no longer significantly different from zero. Once we look only at whites, it is unclear which concept of income inequality is the appropriate one, inequality among whites in the state, or inequality among everyone in the state. The third column shows the effect of replacing the gini coefficient for all incomes with the gini coefficient for white incomes alone. Both coefficients are further reduced, and neither is significantly different from zero. From this, we can deduce that the component of income inequality that matters for mortality is income inequality between races, not income inequality within them. Because blacks are in the minority and have lower incomes, the all-race gini coefficient will be larger where the fraction black is larger, which suggests including it in the regressions. The final columns show the regression containing the fraction black together with the original all-race gini coefficient. The fraction black is estimated to increase white mortality for both males and females. Taking the same example as before, the difference between Vermont and Mississippi, with fractions black of zero and 0.34, gives relative risks of 1.14 for white men, and 1.09 for white women.

The results in Table 1 are important because they show that the effects of income inequality on mortality at the state level are not robust to the inclusion of the fraction of the population that is black. They thus demonstrate that the income inequality hypothesis is incorrect, but they tell us nothing about what actually drives mortality rates. In these state-level data, the fraction black is higher in the southern states, and it is not difficult to think of reasons why mortality, including white mortality, might be higher in the South. But alternative hypotheses are difficult to test with data from only 51 states, so it is useful to move on to the larger number of observations offered by the MSA-level data where it is possible, for example, to look at different regions separately. Quite apart from the fact that there are more of them, cities are more plausibly salient than states for the health of their residents.

Our 1990 MSA data appear to be comparable to those used by Lynch, Kaplan, and Pamuk (1998) (LKP). For example, the correlation coefficients between age-adjusted mortality and the gini coefficient of per equivalent income is 0.28 (compared with 0.25 with the gini of per capita income in LKP) and with the logarithm of per equivalent income is –0.32 (compared with is –0.28 with per capita income in LKP.) Figure 2, which corresponds to Figure 1 for the states, shows the correlation between the gini and the log odds of age adjusted mortality for all persons. Each circle represents an MSA, and the diameters of the circles are in proportion to the population of each, a procedure that makes it clear that the correlation is not driven by a few large MSAs.

Table 2 reproduces Table 1 using MSA data for 1990. As before, the gini coefficient is a significant risk factor when the data are pooled across races, and once again, the effect is removed (indeed reversed) once we control the fraction of each MSA's population that is black. When we restrict the regression to whites, the fraction black is a significant health hazard for both men and women and the coefficients are similar to those estimated from the state data. Once the fraction black is controlled for, income inequality has no effect, whether we use income inequality over everyone, as shown in the final column, or income inequality among whites, not shown. Table 3 repeats the MSA results for 1980. These are shown separately from the 1990 results because, unlike 1990, the gini has no effect on mortality even in the all-race regressions.

Even so, the final regressions for white males and females are similar to those for 1990. Whites die at younger ages in places where a larger fraction of the population is black and, conditional on fraction black, there is no mortality risk associated with income inequality. Indeed, in both Tables 2 and 3, and in the latter significantly so, income inequality appears to be *protective* once we control for the racial composition of the MSAs.

The key to these results is the relationship between income and race across American states and cities. Average incomes for the population as a whole, as well as average incomes among blacks, are negatively correlated with the percent of the population that is black, but the reverse is true for average white incomes. Average incomes of whites are *higher* in cities with a larger fraction of blacks. This divergent behavior of black and white incomes means that the income difference between blacks and whites is larger in cities with larger black populations, which is what induces the relationship between overall income inequality and racial composition. Of course, this *does not* mean that racial composition and income inequality are the same thing, nor that either one is an equally valid marker for the same underlying health risk. In regressions containing both the fraction black and income inequality, the former drives out the latter so that, even if we cannot tell what it is about a high fraction black that drives the mortality results, it is not the associated income inequality.

4. Extensions and further exploration

What is it about the racial composition of places that affects their mortality rates? The previous section demonstrates that it is not the associated income inequality, but nothing beyond that. In this section, we explore a number of other potentially confounding variables in an attempt to learn something about the mechanism that might be driving these results. As explained in the introduction, we focus entirely on the effects of racial composition on white mortality.

One hypothesis concerns education. If the presence of a large black minority results in low levels of education for both blacks and whites, and if education is important for lowering mortality rates, we might find a spurious correlation between racial composition and education. We test this hypothesis by using the census data on individuals' education levels to calculate the fraction of the white population in each MSA whose highest level of education falls into various classes. Table 4 shows the results of the mortality regressions with education included using pooled MSA data from 1980 and 1990.

These results strongly support the view that people with higher education have lower mortality rates, but they do nothing to moderate the estimated effect of the fraction black on white mortality rates. The MSA results are consistent with other results using both regional and individual data; a college education, even some college education, is protective compared with only a high school education. (Though note that for men, postgraduate education adds nothing, and for women, those with postgraduate education are no more protected than high-school graduates or high-school drop outs.) But the main effect of the inclusion of the education variables is not on the estimated effect of racial composition, but on the estimated effect of income, which is now estimated to be mildly hazardous. Such findings are consistent with an earlier literature in economics, Grossman (1975), Fuchs (1989, 1993), and Garber (1989), which argues that it is education, not income, that is protective of health, as well as Ruhm (2000), who argues that business-cycle induced increases in income are hazardous to health. However, they stand in sharp contrast to analyses on individual level data, particularly those using the National

Longitudinal Mortality Study, where income is importantly protective of health even conditional on education, see Elo and Preston (1996) and Deaton and Paxson (2001). The question of whether it is income, education, or some combination that is important for health matters a great deal for policy, especially for arguments about the role of fiscal policy in public health. However, income is not our main concern here, so we do no more than note the puzzle.

Another possible explanation for our main finding is that the provision of public services, especially health services, is poorer in places with a larger black population. Such an explanation would require that the provision of such services is in itself an important determinant of (white) mortality rates, something that goes against an extensive literature that imputes a small or negligible role to access to health care in explaining differences in mortality by socioeconomic status, see for example the review by Adler et al (1994). Moreover, there is evidence against the proposition that health expenditures are indeed lower in places with a larger black minority. Alesina, Baqir, and Easterly (1999) argue on theoretical grounds that racial diversity is likely to decrease the political willingness to provide public goods, but in their empirical analysis find that local public expenditures, including expenditures on health, are *higher* in places where there is more ethnic fractionalization—which in the context of the US means in places where there is a large fraction of the population that is black. In the light of these findings, the provision of local public goods does not seem a promising avenue for explaining our results.

A third line of enquiry is to look at the results by region. In the state-level results with which we began, the correlations between mortality, income inequality, and fraction black had much to do with the South, where all three quantities tend to be higher than in the rest of the country. One of the main advantages of working with the MSA data is the ability to work *within* regions, and thus to eliminate the suspicion that the results are being driven by the South versus the rest of the US. There is also the hypothesis, advanced by Fuchs and McClellan (2001), that the mortality differences might come from selective migration. Migrants are typically healthier than those who stay behind so that, if they migrate from areas with larger to smaller minority black populations, they will increase mortality in the transmitting region and reduce it in the receiving region and, depending on initial conditions, may induce a correlation between white mortality rates and the fraction black. A serious examination of this hypothesis is beyond the scope of this paper, but to the extent that migrations are between regions, intraregional and interregional correlations are likely to differ.

Table 5 shows the results of running a stripped down regression—log odds of white male and female mortality on the mean of the logarithm of per equivalent income, the gini, the fraction black, and the 1990 dummy—for four regions of the US, the North-East, the South, the Mid-West, and the West. The effects of income inequality are inconsistent from region to region, and are more often estimated to be protective than hazardous. There is also some heterogeneity in the effects of income, with income less protective in the West than elsewhere. But the effects of the fraction black are consistently and significantly hazardous in all four regions, though the effects are about twice as large in the North-East and in the South than in the West and Mid-West. In any event, the effect on white mortality does not reflect some unmeasured difference between the South and the rest of the US.

Although there are more MSAs than states, there are not enough to allow us to run cross-MSA regressions state by state. However, it is possible to run the stripped-down regressions with the inclusion of dummy variables, one for each state; when MSAs cross state boundaries, we assign them to the state in which the majority of its population lives. Allowing for state effects allows us to control for unmeasured state-level factors that contribute to mortality rates and that are potentially correlated with the fraction of the population that is black. However, the fraction black remains a hazard to health in these regressions. For white males, and using the same regressions as in Table 5, the coefficient on the fraction black is 0.49 with a *t*–value of 8.7; for females, the coefficient is 0.48 with a *t*–value of 9.9. However, if we go one step further, and include dummies for each of the 287 MSAs, the coefficient on the fraction black becomes small and insignificantly different from zero. Unfortunately, this result is not very informative. When MSA dummies are included, we are essentially running a regression of the changes in the log odds of mortality against the changes in mean income and the racial composition of the MSAs. This regression suffers from a lack of precision because the fraction black does not change much over a decade. Beyond that, all we learn is that the fraction black is standing proxy for some constant or slowly changing factors that are important in the cross-section, but not in the time series. We learn nothing about what those factors might be. We also note that the analysis of changes on changes puts much more strain on the timing—which years of mortality to match with the 1979 and 1989 income and race data from the census—than is the case with the cross-sectional results.

Finally, we look at the age composition of mortality. Because the cause of death differs by age, locating the effects of racial composition in the age distribution may give some clue about the mechanisms involved. The age-specific regressions also protect us against potential artefactual effects associated with age-adjustment, which requires an essentially arbitrary choice of base population. Table 6 presents the estimated effects of income and of fraction black on the mortality rate in thirteen age groups. The form of these regressions differs from before. In a few of the smaller MSAs, there are no recorded deaths in the specific age groups in one or other of the years, and such observations cannot be included in a regression with the log-odds as dependent variable. Dropping them and running the standard regression produces results that are qualitatively similar to those in the table. Even so, we present the results of regressions using the mortality rate itself as the dependent variable. On the right hand side, in addition to the fraction black and the dummy for 1990, we include the mean of income per equivalent, rather than the mean of its logarithm. The table shows the coefficients for the fraction black and for income, by sex and age, with males on the left and females on the right. These are scaled so that the numbers in the left-hand panel are estimates of the effects on the white mortality rate per 1,000 of moving from an MSA with zero to one with 100 percent black population, while those in the right-hand panel are the effects of an additional $1,000 on the mortality rate per 1,000. Average mortality rates across the MSAs (weighted by population) are shown for comparison.

These results do nothing to resolve the puzzle. The effects of racial composition, like those of income, are different at different ages, and varying largely in proportion to the level of mortality itself, so that the effect on the log odds would be roughly the same at all ages. With the exception of males aged 1 through 9, the fraction black is estimated as a significant risk to mortality at all ages. It is particularly high for 15–19 year old males, falling off for 15 years thereafter, but rising rapidly with age thereafter. The effect is always positive, and always significantly different from zero. Miller and Paxson (2001) further show, using PUMA level data, that the fraction black is correlated with the death from a range of diseases; for example, for white males aged 25 to 64, the effect is present for death from infectious disease, cancer, homicide, and cardio-vascular disease, but not for diabetes nor accidents.

Figure 3 shows scatter-plots between the fraction black and the mortality of white males at selected ages using the MSA data for 1990; once again, the diameters of the circles are proportional to population size. Note that each plot has its own scale for the vertical axis. The figures provide an immediate visual counterpart to the results in Table 6, and they also establish that the correlations do not depend on one or two peculiar MSAs. Even in the three central panels, where there is one large MSA in the upper-right (New York City), the significance of the positive correlation is not affected by its exclusion or down-weighting.

5. Conclusions

Cross-section regressions across American states and cities show that, conditional on racial composition, income inequality does not raise the risk of mortality. The fraction of the population that is black is a significant risk-factor for mortality, not only for the population as a whole—which would follow mechanically from the fact that blacks have higher mortality rates than whites—but for both blacks and whites separately. Our empirical results cast remarkably little light on why the fraction black should be associated with higher white mortality. The effect is robust to conditioning on education, it is present for all age-groups except boys aged 1 to 9, and it is present within geographical regions of the country. It cannot plausibly be attributed to variation in local public provision of health services. Further research is needed to identify the mechanisms that generate these effects, as well as to resolve the currently contradictory evidence on the effects of income and education on mortality rates. These are much more likely to be productive avenues for research than further work on the effects of income inequality on health.

6. List of works cited:

Adler, Nancy E., Thomas Boyce, Margaret A. Chesney, Sheldon Cohen, Susan Folkman, and S. Leonard Syme, 1994, "Socioeconomic status and health: the challenge of the gradient", *American Psychologist*, **49**, 15–24.

Alesina, Alberto, Reza Baqir, and William Easterly, 1999, "Public goods and ethnic divisions," *Quarterly Journal of Economics*, 114, 1243–84.

Daly, Mary C., Greg J. Duncan, George A. Kaplan, and John W. Lynch, 1998, "Macro to micro links in the relation between income inequality and mortality," *The Millbank Quarterly*, 76, 315–39.

Deaton, Angus, 2001a, "Relative deprivation, inequality, and mortality," Princeton University, Research Program in Development Studies and Center for Health and Wellbeing, processed, and Cambridge, Mass, NBER Working Paper No. 8099 (January).

Deaton, Angus, 2001b, "Health, inequality, and economic development," Princeton University, Center for Health and Wellbeing, processed, and Cambridge, Mass, NBER Working Paper No. 8318 (June).

Deaton, Angus and Christina Paxson, 2001, "Mortality, education, income, and inequality among American cohorts," in David Wise, editor, *Themes in the Economics of aging*, Chicago University Press for NBER. (NBER Working Paper No. 7140)

Elo, Irma and Samuel H. Preston, 1996, "Educational differentials in mortality: United States, 1979–85," *Social Science and Medicine*, 42, 47–57.

Fiscella, Kevin, and Peter Franks, 1997, "Poverty or income inequality as a predictor of mortality: longitudinal cohort study," *British Medical Journal*, 314, 1724–28.

Fuchs, Victor R., 1989, "Comments," in John P. Bunker, Deanna S. Gombey, and Barbara Kehrer, eds., *Pathways to health: the role of social factors*, Menlo Park, Ca. Kaiser Family Foundation, pp. 226–9.

Fuchs, Victor R., 1993, "Poverty and health: asking the right questions," in David E. Rogers and Eli Ginzburg, eds., *Medical Care and the Health of the Poor*, Boulder, Co. Westview Press.

Fuchs, Victor, and Mark McClellan, 2001, "Area differences in utilization of medical care and mortality among U.S. elderly," Stanford University and NBER, processed. (May).

Garber, Alan M., 1989, "Pursuing the links between socioeconomic factors and health: critique, policy implications, and directions for future research," in John P. Bunker, Deanna S. Gombey, and Barbara Kehrer, eds., *Pathways to health: the role of social factors,* Menlo Park, Ca. Kaiser Family Foundation, pp. 271–315.

Grossman, Michael, 1975, "The correlation between health and schooling," in Nestor E. Terleckyj, ed., *Household production and consumption*, New York, NBER.

Jaeger, David A., Susannah Loeb, Sarah E. Turner, and John Bound, 1998, "Coding geographical areas across census years: creating consistent definitions of Metropolitan Areas," Cambridge, Mass. NBER Working Paper No. 6772 (Oct.)

Kaplan, George A., Elsie R. Pamuk, John W. Lynch, Richard. D. Cohen, and Jennifer L. Balfour, 1996, "Inequality in income and mortality in the United States: analysis of mortality and potential pathways," *British Medical Journal*, 312, 999—1003.

Kawachi, Ichiro, Bruce P. Kennedy, and Richard G. Wilkinson, 1999, *The society and population health reader: Volume 1: Income inequality and health*, New York, New Press.

Kennedy, Bruce P., Ichiro Kawachi, and Deborah Prothrow-Stith, 1996, "Income distribution and mortality: cross-sectional ecological study of the Robin Hood index in the United States," *Britich Medical Journal*, 312, 1004–07.

Kennedy, Bruce, P., Ichiro Kawachi, and Deborah Prothrow-Smith, 1996, "Important correction," *British Medical Journal*, 312, 1194.

Lochner, Kim, Elsie R. Pamuk, Diane Makuc, Bruce P. Kennedy and Ichiro Kawachi, 2001, "State-level income inequality and individual mortality risk: a prospective multilevel study," *American Journal of Public Health*, 91(3), 385–391.

Lynch, John W., George A. Kaplan, and Elsie R. Pamuk, 1998, "Income inequality and mortality in metropolitan areas of the United States," *American Journal of Public Health*, 88, 1074–80.

Mellor, Jennifer M., and Jeffrey Milyo, 2001, "Re-examining the evidence of an ecological association between income inequality and health," *Journal of Health Politics, Policy, and Law*, forthcoming.

Miller, Douglas, and Christina Paxson, 2001, "Relative income, race, and mortality," Center for Health and Wellbeing, Princeton University, processed.

Ross, Nancy A., Michael C. Wolfson, James R. Dunn, Jean-Marie Berthelot, George A. Kaplan, and John W. Lynch, 2000, "Relationship between income inequality and mortality in Canada and in the United States: cross sectional assessment using census data and vital statistics," *British Medical Journal*, 319, 989–902.

Ruhm, Christopher, 2000, "Are recessions good for your health?" *Quarterly Journal of Economics*, 115, 617–50.

Wilkinson, Richard G., 1992, "Income distribution and life-expectancy," *British Medical Journal*, 304, 165–68.

Wilkinson, Richard G., 1996, *Unhealthy societies: the affliction of inequality*, London. Routledge.

Wilkinson, Richard G., 2000, *Mind the gap: hierarchies, health, and human evolution,* London, Weidenfeld and Nicolson.

Table 1: Log odds of mortality regressions: 50 US states plus DC, 1980 and 1990 pooled

	All Males			White Males only			
Equivalent income	–0.22	–0.11	–0.14	–0.11	–0.09	–0.09	–0.16
	(4.1)	(1.9)	(3.3)	(2.5)	(2.2)	(2.1)	(4.2)
Gini coefficient		1.42	–0.24		0.92		0.01
		(3.9)	(0.8)		(3.6)		(0.0)
Gini among whites						0.62	
						(1.8)	
Fraction black			0.71				0.42
			(10.3)				(5.7)
1990 dummy	–0.11	–0.16	–0.11	–0.13	–0.16	–0.15	–0.12
	(6.5)	(7.9)	(7.3)	(9.4)	(10.3)	(8.3)	(8.7)
	All Females			White Females only			
Equivalent income	–0.09	0.00	–0.02	0.02	0.03	0.03	–0.02
	(2.0)	(0.1)	(0.5)	(0.4)	(0.6)	(0.6)	(0.5)
Gini		1.08	–0.36		0.38		–0.30
		(3.4)	(1.1)		(1.6)		(1.0)
Gini among whites						0.24	
						(0.7)	
Fraction black			0.51				0.26
			(7.4)				(3.2)
1990 dummy	–0.07	–0.11	–0.66	–0.08	–0.09	–0.09	–0.07
	(5.0)	(6.2)	(4.1)	(6.5)	(6.2)	(5.1)	(4.2)

Notes: Equivalent income is the mean of the logarithm of income per adult equivalent, calculated on an individual basis with 1979 repriced to 1989 using the CPI. The gini coefficient relates to income per equivalent, again on an individual basis. The Gini coefficient among whites is calculated using white incomes only. Gini coefficients are calculated separately for males and for females, after imputing household income per equivalent adult to each individual. There are 102 observations in all regressions. The dependent variable is the log odds of age-adjusted mortality; mortality is adjusted to the 1990 US population; age adjustment is done separately by sex, and separately for all groups, and for whites. The figures in brackets are absolute *t*-values. All regressions are weighted by the square root of the relevant population.

Table 2: Log odds of mortality regressions: 287 Metropolitan Statistical Areas, 1990

	All Males			White Males only			
Equivalent income	–0.16	–0.12	–0.17	–0.10	–0.08	–0.09	–0.15
	(5.6)	(3.7)	(7.2)	(4.1)	(3.4)	(3.5)	(6.8)
Gini coefficient		0.55	–0.38		0.46		–0.09
		(2.7)	(2.5)		(3.0)		(0.6)
Gini among whites						0.16	
						(0.9)	
Fraction black			0.83				0.50
			(16.7)				(9.0)
	All Females			White Females only			
Equivalent income	–0.09	–0.05	–0.09	–0.01	–0.01	–0.02	–0.06
	(3.8)	(2.0)	(4.1)	(0.7)	(0.3)	(0.7)	(2.8)
Gini		0.44	–0.42		0.26		–0.22
		(2.6)	(2.8)		(1.9)		(1.5)
Gini among whites						–0.03	
						(0.2)	
Fraction black			0.54				0.31
			(13.0)				(6.4)

Notes: Equivalent income is the mean of the logarithm of income per adult equivalent, calculated on an individual basis with 1979 repriced to 1989 using the CPI. The gini coefficient relates to income per equivalent, again on an individual basis. The Gini coefficient among whites is calculated using white incomes only. Gini coefficients are calculated separately for males and for females, after imputing household income per equivalent adult to each individual. There are 287 observations in all regressions. The dependent variable is the log odds of age-adjusted mortality; mortality is adjusted to the 1990 US population; age adjustment is done separately by sex, and separately for all groups, and for whites. The figures in brackets are absolute *t*-values. All regressions are weighted by the square root of the relevant population.

Table 3: Log odds of mortality regressions: 287 Metropolitan Statistical Areas, 1980

	All Males			White Males only			
Equivalent income	–0.16	–0.15	–0.19	–0.05	–0.05	–0.10	–0.17
	(4.8)	(3.7)	(6.3)	(1.8)	(1.6)	(3.1)	(5.9)
Gini coefficient		0.08	–1.15		0.05		–0.90
		(0.4)	(6.3)		(0.3)		(5.3)
Gini among whites						–0.64	
						(3.2)	
Fraction black			0.74				0.54
			(15.7)				(10.2)
	All Females			**White Females only**			
Equivalent income	–0.07	–0.04	–0.08	0.05	0.05	0.00	–0.04
	(2.1)	(0.9)	(2.2)	(1.6)	(1.5)	(0.1)	(1.3)
Gini		0.25	–0.91		–0.02		–0.75
		(1.1)	(4.3)		(0.1)		(3.7)
Gini among whites						–0.54	
						(2.5)	
Fraction black			0.56				0.34
			(11.2)				(5.7)

Notes: Equivalent income is the mean of the logarithm of income per adult equivalent, calculated on an individual basis with 1979 repriced to 1989 using the CPI. The gini coefficient relates to income per equivalent, again on an individual basis. The Gini coefficient among whites is calculated using white incomes only. Gini coefficients are calculated separately for males and for females, after imputing household income per equivalent adult to each individual. There are 287 observations in all regressions. The dependent variable is the log odds of age-adjusted mortality; mortality is adjusted to the 1990 US population; age adjustment is done separately by sex, and separately for all groups, and for whites. The figures in brackets are absolute *t*-values. All regressions are weighted by the square root of the relevant population.

Table 4: Education, income, inequality and white mortality across MSAs in 1980 and 1990

	White males		White females	
Equivalent income	0.052	(2.1)	0.081	(3.3)
Gini coefficient	0.263	(2.0)	–0.084	(0.6)
Fraction black	0.388	(10.6)	0.227	(5.9)
No high school	0.059	(0.7)	0.060	(0.8)
Some college	–0.266	(3.0)	–0.204	(2.3)
College graduate	–0.530	(3.9)	–0.883	(6.1)
Post-graduate	–0.512	(4.2)	0.277	(1.5)
1990 dummy	–0.108	(10.3)	–0.020	(2.1)

Notes: Pooled data, 1980 and 1990, 574 observations. OLS regressions with the log odds of age-adjusted mortality as the dependent variable; age-adjustment is to the 1990 US population and is done separately for males and females. Equivalent income is the mean in the MSA of log income per adult equivalent at 1989 prices. The gini coefficient is calculated on an individual basis from income per equivalent adult over all races. The schooling variables are the fractions of people (white men or women, respectively) in the MSA whose highest education is as shown. The omitted category is high school graduate. Absolute *t*-values are shown in parentheses. All regressions are weighted by the square root of the relevant population.

Table 5: Regional regressions of mortality across MSAs in 1980 and 1990

	NORTH EAST				SOUTH			
	White males		White females		White males		White females	
Equivalent income	–0.242	(7.4)	–0.118	(4.0)	–0.209	(6.3)	–0.149	(4.3)
Gini coefficient	0.198	(0.8)	–0.363	(1.4)	–1.252	(5.5)	–0.833	(3.5)
Fraction black	0.401	(3.3)	0.498	(4.6)	0.438	(7.1)	0.386	(6.4)
1990 dummy	–0.113	(9.8)	–0.096	(7.3)	–0.075	(5.7)	–0.028	(2.1)
	MID-WEST				WEST			
Equivalent income	–0.300	(7.3)	–0.172	(3.8)	–0.076	(1.5)	–0.079	(2.1)
Gini coefficient	–0.414	(1.5)	–0.682	(2.0)	0.662	(2.2)	0.141	(0.6)
Fraction black	0.993	(10.0)	0.920	(8.3)	0.843	(3.5)	0.956	(5.7)
1990 dummy	–0.129	(9.8)	–0.054	(3.2)	–0.149	(8.0)	–0.067	(4.7)

Notes: Pooled data, 1980 and 1990. OLS regressions with the log odds of age-adjusted mortality as the dependent variable; age-adjustment is to the 1990 US population as a whole, but is done separately for males and females. Variables as defined in previous tables. Each column represents a regression. There are 98 observations for the North East, 216 in the South, 158 in the Mid-West, and 102 in the West. The standard Census regions are: North-East: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont; South: Delaware, District of Columbia, Florida, Georgia, Maryland, N. and S. Carolina, Virginia, W. Virginia, Alabama, Kentucky, Mississippi, Tennessee, Arkansas, Louisiana, Oklahoma, and Texas; Mid-West: Illinois, Indiana, Michigan, Ohio, Wisconsin, Iowa, Kansas, Minnesota, Missouri, Nebraska, N. and S. Dakota; West: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, California, Hawaii, Oregon, and Washington. All regressions are weighted by the square root of the relevant population.

Table 6: Age, racial composition, and white mortality across MSAs in 1980 and 1990

	Males					Females				
Age group	Mort-ality	Fraction Black* 10^3		Mean Income * 10^7		Mort-ality	Fraction Black* 10^3		Mean Income * 10^7	
0 to 1	8.4	3.06	(2.9)	–2.25	(6.0)	6.5	2.31	(2.5)	–1.30	(3.8)
1–4	0.4	0.06	(0.5)	–0.18	(3.6)	0.4	0.30	(2.8)	–0.24	(5.1)
5–9	0.2	0.11	(1.4)	–0.09	(2.6)	0.2	0.14	(2.3)	–0.05	(1.8)
10–14	0.3	0.27	(3.3)	–0.15	(4.5)	0.2	0.24	(3.6)	–0.16	(5.4)
15–19	1.1	1.00	(5.0)	–0.38	(4.4)	0.4	0.48	(4.9)	–0.17	(3.8)
20–24	1.4	0.61	(2.4)	–0.14	(1.4)	0.4	0.27	(2.7)	–0.01	(0.3)
25–34	1.8	0.85	(3.9)	–0.15	(2.0)	0.6	0.32	(3.8)	–0.06	(1.9)
35–44	2.7	1.29	(3.7)	0.26	(2.4)	1.2	0.60	(4.6)	–0.12	(2.5)
45–54	5.4	4.09	(9.0)	–0.95	(7.4)	3.0	0.93	(4.1)	–0.32	(4.1)
55-64	14.5	9.96	(12.5)	–3.13	(14.4)	8.2	2.38	(5.6)	–0.65	(4.0)
65–74	33.7	19.72	(11.5)	–6.91	(12.4)	19.3	4.98	(4.9)	–1.65	(4.1)
75–84	78.1	29.93	(8.0)	–9.48	(7.3)	48.6	13.71	(6.6)	–2.12	(2.2)
85+	182.0	20.49	(2.4)	–0.85	(0.3)	144.1	23.25	(4.8)	0.09	(0.0)

Notes: Pooled data, 1980 and 1990. Each number comes from an OLS regression with probability of death on the left hand side and the fraction black, mean income per equivalent in 1989 prices, and a dummy for 1990 on the right-hand side. The coefficients on fraction black are multiplied by 1,000 and are therefore the effect of a unit change (from 0 to 1.0) on the mortality rate per 1,000. The coefficient on income is multiplied by 10,000,000, and so represents the effects of an additional $1,000 of per equivalent income on the mortality rate per 1,000. All regressions are weighted by the square root of the relevant age and sex specific population.

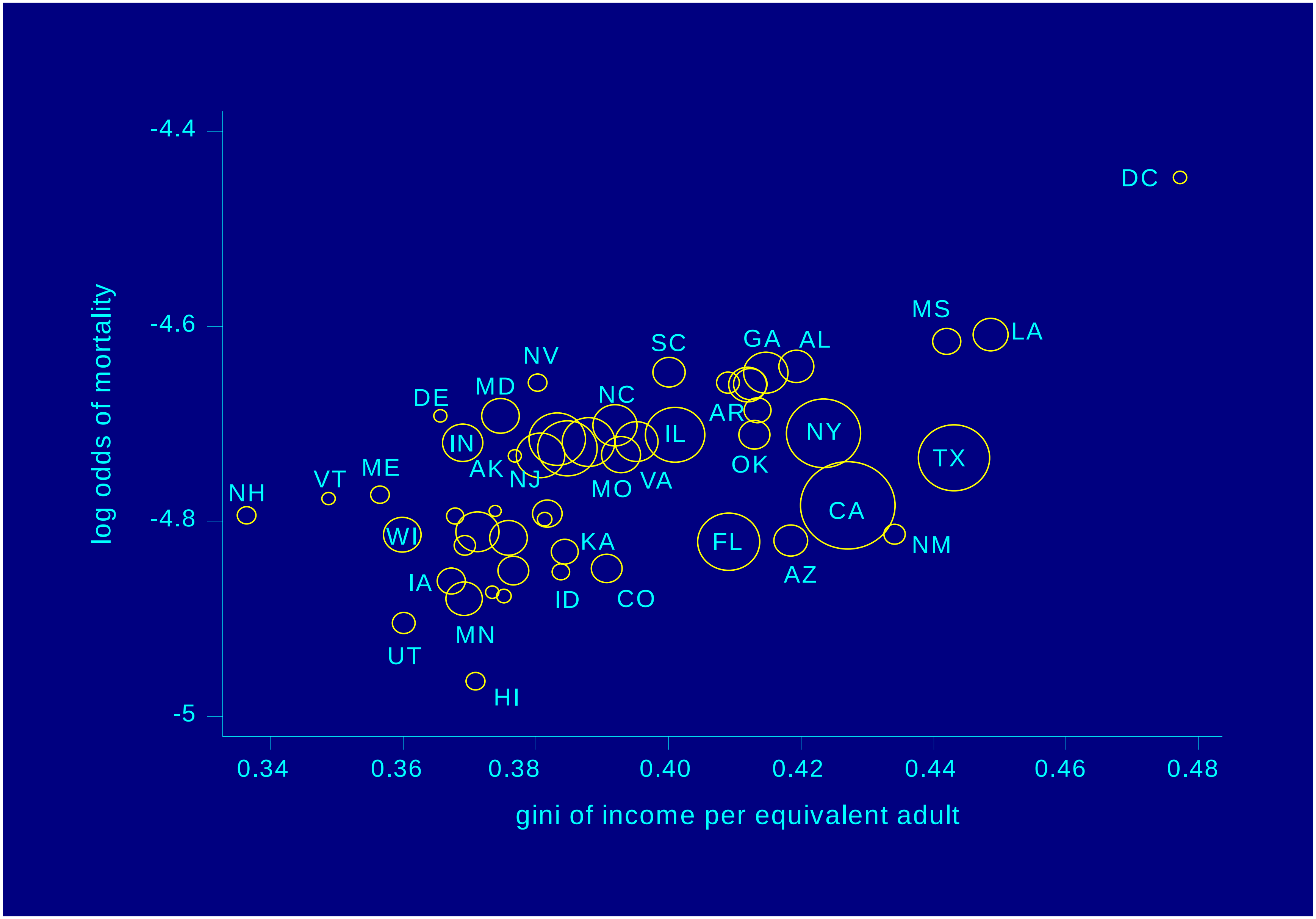

Figure 1: Inequality and age-adjusted mortality across US states, 1990. (Circles are proportional to population.)

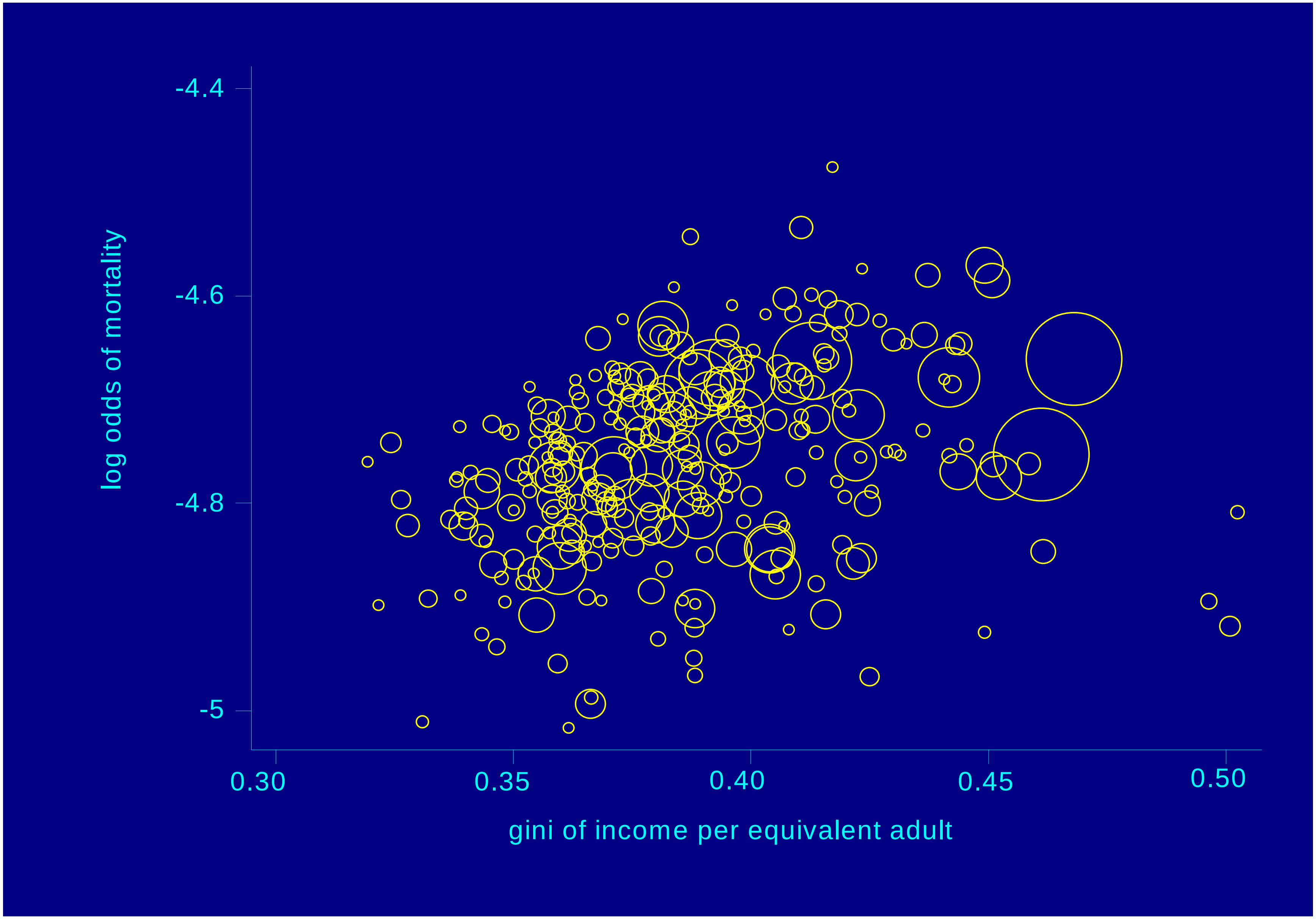

Figure 2: Income inequality and mortality across US MSAs, 1990. (Circles are proportional to population.)

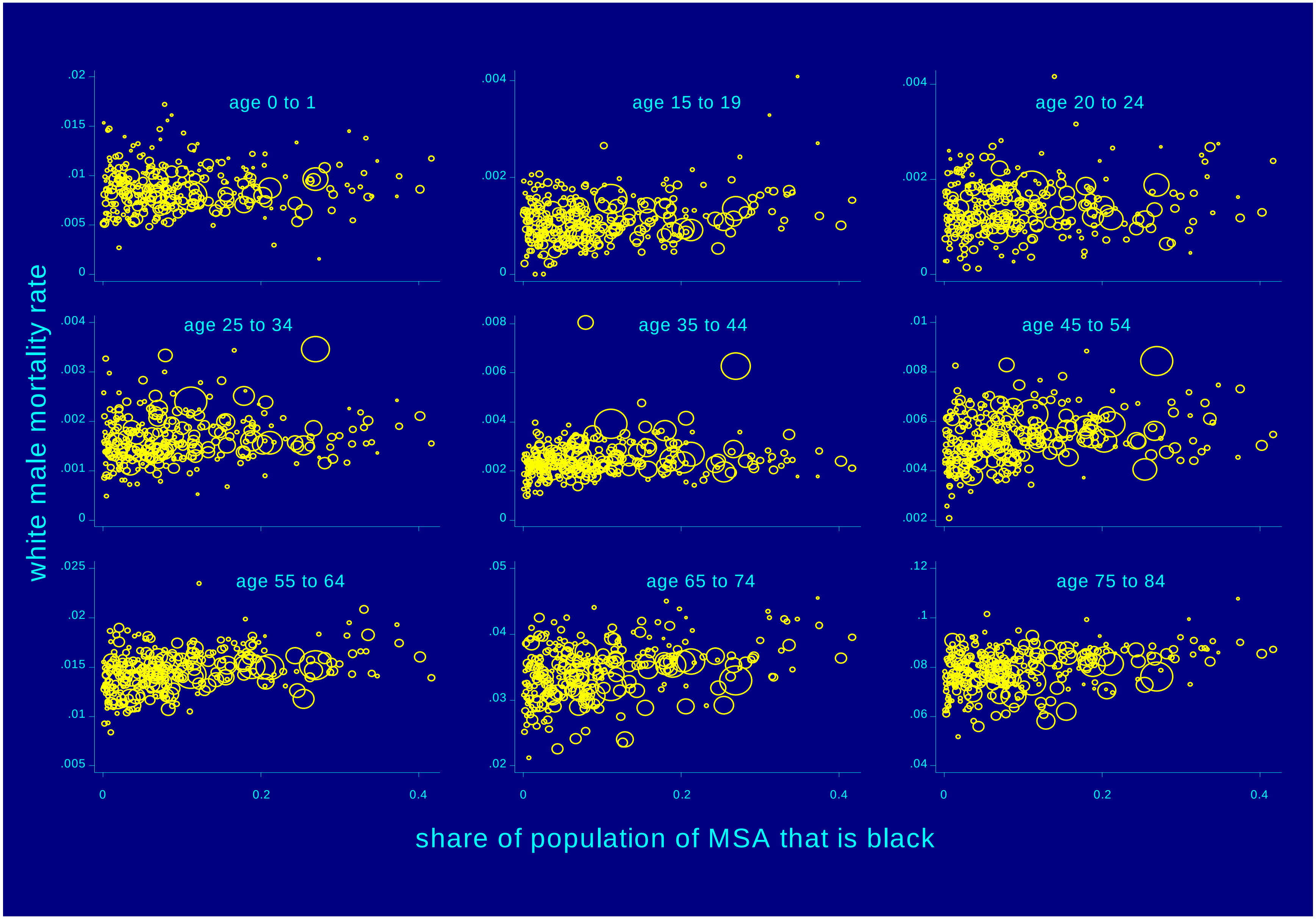

Figure 3: White male mortality at selected ages versus fraction black, 287 MSAs, 1990

MSA definitions used in
"Mortality, inequality, and race in America cities and states"
Angus Deaton and Darren Lubotsky
June, 2001

See note at foot of page for definitions

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Abeline, TX	40	Texas	26	3000	South
Akron, OH	80	Ohio	23-25	900 4701-4704	Midwest
Albany-Schenectady-Troy, NY	160	New York	5-9 33	800-1200 3800	Northeast
Albuquerque, NM	200	New Mexico	4-5	201-204	West
Alexandria, LA	220	Louisiana	7	600	South
Allentown-Bethlehem-Easton, PA	240	Pennsylvania	51-54	2001-2002 2101-2102	Northeast
Altoona, PA	280	Pennsylvania	36	1600	Northeast
Amarillo, TX	320	Texas	1	100	South
Anaheim-Santa Ana, CA	360	California	43-46	4200-4700 4801-4808	West
Anderson, IN	400	Indiana	20	2200	Midwest
Anderson, SC	405	South Carolina	1	2100	South
Ann Arbor, MI	440	Michigan	34-35	3100-3200	Midwest
Anniston, AL	450	Alabama	7	1900	South
Appleton-Oshkosh-Neenah, WI	460	Wisconsin	5-6	400-500	Midwest
Asheville, NC	480	North Carolina	2	200-300	South
Atlanta, GA	520	Georgia	4-12 27	800 1500-1700 1801-1803 1901-1904 2001-2005 2300 3000	South
Atlantic City, NJ	560	New Jersey	38	100-200	Northeast
Augusta, GA-SC	600	Georgia*	25-26	500 2200	South
		South Carolina	16	1800	

Note: The first and second columns give the MSA name and FIPS code. The third through fifth columns give the state, and the codes of the County Groups (1980) and PUMA's (1990) that comprise the MSA. When we include state fixed effects or run the models separately by region, we assign MSA's that cross state boundaries to the state indicated with an asterisk. The final column gives the region we assign to each MSA. See main text for additional details.

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Aurora-Elgin, IL	620	Illinois	35	3501-3502	Midwest
Austin, TX	640	Texas	44-45 66	4901-4904 5000-5200	South
Bakersfield, CA	680	California	36-37	4900 5001-5002	West
Baltimore, MD	720	Maryland	4-8 12	201-204 301-306 500-600 1000-1100 1501-1504	South
Bangor, ME	733	Maine	3	200	Northeast
Baton Rouge, LA	760	Louisiana	14-17	1200 1301-1302 1400-1500	South
Battle Creek, MI	780	Michigan	29	2600	Midwest
Beaumont-Port Arthur, TX	840	Texas	54-56	5900-6100	South
Beaver County, PA	845	Pennsylvania	22-23	3800	Northeast
Bellingham, WA	860	Washington	1	100	West
Benton Harbor, MI	870	Michigan	25	2300	Midwest
Bergen-Passaic, NJ	875	New Jersey	1-9	400-1200	Northeast
Billings, MT	880	Montana	3	600	West
Biloxi-Gulfport, MS	920	Mississippi	16	1600	South
Binghamton, NY	960	New York	17 29-30	3500 3601-3602	Northeast
Birmingham, AL	1000	Alabama	9-11	1700 1801-1806	South
Bloomington, IN	1020	Indiana	27	2100	Midwest
Bloomington-Normal, IL	1040	Illinois	29	2700	Midwest
Boise City, ID	1080	Idaho	3	301-302	West
Boston-Lawrence-Salem- Lowell-Brockton, MA	1123	Massachusetts	11-37	1400-1900 2001-2005 2100-3800 4100-4300	Northeast
Boulder-Longmont, CO	1125	Colorado	10	701-702	West
Bradenton, FL	1140	Florida	27	4700	South
Brazoria, TX	1145	Texas	62	6400	South
Bremerton, WA	1150	Washington	16	1600	West
Brownsville-Harlingen, TX	1240	Texas	38	4201-4202	South
Buffalo, NY	1280	New York	23-24	2500-2900 3001-3003	Northeast

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Canton, OH	1320	Ohio	27-28	5001-5003	Midwest
Cedar Rapids, IA	1360	Iowa	14	1200	Midwest
Champaign-Urbana-Rantoul, IL	1400	Illinois	28	2600	Midwest
Charleston, SC	1440	South Carolina	19-21	1100 1201-1202	South
Charleston, WV	1480	West Virginia	8-10	700	South
Charlotte-Gastonia- Rock Hill, NC-SC	1520	North Carolina*	36-40	801-804 900-1200	South
		South Carolina	6	500	
Charlottesville, VA	1540	Virginia	3	1300	South
Chattanooga, TN-GA	1560	Georgia Tennessee*	1 15-17	300 1000 1200-1300	South
Chicago, IL	1600	Illinois	32-34 37	3001-3019 3101-3114 3201-3206 3300	Midwest
Chico, CA	1620	California	5	600	West
Cincinnati, OH-KY-IN	1640	Indiana Kentucky Ohio*	36 1-2 52-55	400 1300-1400 3000 3300 5401-5406	Midwest
Clarksville-Hopkinsville, TN-KY	1660	Tennessee	7	2200	South
Cleveland, OH	1680	Ohio	8-18	400 800 3901-3905 4000-4600	Midwest
Colorado Springs, CO	1720	Colorado	14	1000-1100	West
Columbia, MO	1740	Missouri	6	600	Midwest
Columbia, SC	1760	South Carolina	13-15	1601-1602 1700	South
Columbus, GA-AL	1800	Alabama Georgia*	15 15-17	1200 200 2900	South
Columbus, OH	1840	Ohio	41-44 56	1700 2300 2700-2800 5101-5107	Midwest

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Corpus Christi, TX	1880	Texas	35 39	4000 4301-4302 4800	South
Cumberland, MD-WV	1900	Maryland West Virginia*	1 6	100 500	South
Dallas, TX	1920	Texas	13-18 22-24	1500 2201-2202 2300-2400 2501-2509 2600-2800 2901-2904	South
Danville, VA	1950	Virginia	12	1500	South
Davenport-Rock Island-Moline, IA-IL	1960	Illinois* Iowa	7 10-11 16	600 1000 1400	Midwest
Dayton-Springfield, OH	2000	Ohio	45-50	2100-2200 2600 5201-5205	Midwest
Daytona Beach, FL	2020	Florida	14-15	900	South
Decatur, IL	2040	Illinois	16	1500	Midwest
Denver, CO	2080	Colorado	6-12	101-104 201-202 301-302 400-500 601-602	West
Des Moines, IA	2120	Iowa	8-11	700-1000	Midwest
Detroit, MI	2160	Michigan	24 33 36-59	2200 3000 3301-3308 3401-3405 3500-3800 3901-3903 4000 4101-4107 4200-4400	Midwest
Dubuque, IA	2200	Iowa	15	1300	Midwest
Duluth, MN-WI	2240	Minnesota* Wisconsin	3-4 1	300 100	Midwest
Eau Claire, WI	2290	Wisconsin	11	1000	Midwest
El Paso, TX	2320	Texas	32	3701-3705	South
Elkhart-Goshen, IN	2330	Indiana	8	1400	Midwest
Erie, PA	2360	Pennsylvania	1-2	101-102	Northeast
Eugene-Springfield, OR	2400	Oregon	9-10	700-800	West
Evansville, IN	2440	Indiana	31-33	500 2400	Midwest

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Fayetteville, NC	2560	North Carolina	32-33	3000-3100	South
Fayetteville-Springdale, AR	2580	Arkansas	1	200	South
Flint, MI	2640	Michigan	21-23	2000 2101-2102	Midwest
Florence, SC	2655	South Carolina	9	800	South
Fort Collins-Loveland, CO	2670	Colorado	3	800	West
Fort Lauderdale-Hollywood-Pompano Beach, FL	2680	Florida	36-42	3200-3400 3501-3506	South
Fort Myers-Cape Coral, FL	2700	Florida	31	3001-3003	South
Fort Pierce, FL	2710	Florida	30	2700-2800	South
Fort Smith, AR	2720	Arkansas	13	800	South
Fort Walton Beach, FL	2750	Florida	2-3	300-400	South
Fort Wayne, IN	2760	Indiana	9-12	1700-1800 2600-2700	Midwest
Fort Worth-Arlington, TX	2800	Texas	19-21 25	1800 1901-1904 2001-2002 2101-2104	South
Fresno, CA	2840	California	27-28	4000-4100	West
Gadsden, AL	2880	Alabama	6	2000	South
Gainesville, FL	2900	Florida	10	700	South
Galveston-Texas City, TX	2920	Texas	57	6301-6302	South
Gary-Hammond, IN	2960	Indiana	1-4	800-1000 1200	Midwest
Glens Falls, NY	2975	New York	4	700	Northeast
Grand Rapids, MI	3000	Michigan	13-15 17	1300 1401-1402 1500	Midwest
Greeley, CO	3060	Colorado	4	900	West
Green Bay, WI	3080	Wisconsin	4	300	Midwest
Greensboro--Winston-Salem-High Point, NC	3120	North Carolina	7-13	1300-1700 2000-2200	South
Greenville-Spartanburg, SC	3160	South Carolina	2-5	100 201-202 301-302	South
Hagerstown, MD	3180	Maryland	2	1400	South
Hamilton-Middletown, OH	3200	Ohio	51	5301-5302	Midwest

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Harrisburg-Lebanon-Carlisle, PA	3240	Pennsylvania	41-45	2200 3601-3602 3700	Northeast
Hartford-New Britain- Middletown-Bristol, CT	3283	Connecticut	11 14-21	200-1000 2500	Northeast
Hickory-Morganton, NC	3290	North Carolina	4-5	600-700	South
Honolulu, HI	3320	Hawaii	1-2	301-307	West
Houma-Thibodaux, LA	3350	Louisiana	20	1800	South
Houston, TX	3360	Texas	50-51 58-64	5500-5600 6200 6501-6502 6601-6615 6700-6800 6901-6908 7200	South
Huntington-Ashland, WV-KY-OH	3400	Kentucky Ohio West Virginia*	11 59 10-11	1100 3500 800	South
Huntsville, AL	3440	Alabama	3-5	2200-2300	South
Indianapolis, IN	3480	Indiana	21-24	101-107 3300-3500	Midwest
Iowa City, IA	3500	Iowa	12-13	1100	Midwest
Jackson, MI	3520	Michigan	31	2800	Midwest
Jackson, MS	3560	Mississippi	8-10	800-1000	South
Jacksonville, FL	3600	Florida	7-9	200 1000-1100	South
Jacksonville, NC	3605	North Carolina	26	3600	South
Jamestown-Dunkirk, NY	3610	New York	26	3100	Northeast
Janesville-Beloit, WI	3620	Wisconsin	19	1800	Midwest
Jersey City, NJ	3640	New Jersey	10-12	1301-1302 1400-1500	Northeast
Johnson City-Kingsport- Bristol, TN-VA	3660	Tennessee* Virginia	25-27 9-10	100-300 200 3200	South
Johnstown, PA	3680	Pennsylvania	35	3400	Northeast
Joliet, IL	3690	Illinois	38	3700-3900	Midwest
Joplin, MO	3710	Missouri	14	1800	Midwest
Kalamazoo, MI	3720	Michigan	27-28	2501-2502	Midwest

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Kansas City, MO-KS	3760	Kansas	9-11	901-903 1000-1100	Midwest
		Missouri*	7-12	700-900 1001-1005	
Kenosha, WI	3800	Wisconsin	21	2000	Midwest
Killeen-Temple, TX	3810	Texas	46-47	5300-5400	South
Knoxville, TN	3840	Tennessee	21-24	600-900	South
Kokomo, IN	3850	Indiana	14-15	3000	Midwest
La Crosse, WI	3870	Wisconsin	14	1300	Midwest
Lafayette, LA	3880	Louisiana	11	900	South
Lafayette-West Lafayette, IN	3920	Indiana	16	2000	Midwest
Lake Charles, LA	3960	Louisiana	8-9	700	South
Lake County, IL	3965	Illinois	36	3401-3404	Midwest
Lakeland-Winter Haven, FL	3980	Florida	20	4600	South
Lancaster, PA	4000	Pennsylvania	46-48	3501-3503	Northeast
Lansing-East Lansing, MI	4040	Michigan	18-20	1700-1900	Midwest
Laredo, TX	4080	Texas	34	3900	South
Las Cruces, NM	4100	New Mexico	7	900	West
Las Vegas, NV	4120	Nevada	1-2	201-205	West
Lawrence, KS	4150	Kansas	15	1500	Midwest
Lexington-Fayette, KY	4280	Kentucky	6-8	1600-1700 1801-1802	South
Lima, OH	4320	Ohio	33	1100	Midwest
Lincoln, NE	4360	Nebraska	5	800	Midwest
Little Rock-North Little Rock, AR	4400	Arkansas	7-10	1500-1600	South
Longview-Marshall, TX	4420	Texas	10	1100	South
Lorain-Elyria, OH	4440	Ohio	6-7	3700-3800	Midwest
Los Angeles-Long Beach, CA	4480	California	40-42	5200-6300 6401-6424 6501-6521 6600	West
Louisville, KY-IN	4520	Indiana Kentucky*	34-35 3-5	200-300 1900 2001-2002 2101-2103	South
Lubbock, TX	4600	Texas	3	301-302	South

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Lynchburg, VA	4640	Virginia	11 13	1400 1600	South
Macon-Warner Robins, GA	4680	Georgia	13-14	400 3100	South
Madison, WI	4720	Wisconsin	17-18	1600-1700	Midwest
Mansfield, OH	4800	Ohio	30	1300	Midwest
McAllen-Edinburg-Mission, TX	4880	Texas	36-37	4101-4103	South
Medford, OR	4890	Oregon	14	900	West
Melbourne-Titusville-Palm Bay, FL	4900	Florida	16	1200-1400	South
Memphis, TN-AR-MS	4920	Arkansas Mississippi Tennessee*	5 1 1-3	600 100 1800-2000	South
Merced, CA	4940	California	29	3000	West
Miami-Hialeah, FL	5000	Florida	43-53	3601-3602 3700-3800 3901-3909	South
Middlesex-Somerset-Hunterdon, NJ	5015	New Jersey	26-31 51-52	1600-2300 4000	Northeast
Milwaukee, WI	5080	Wisconsin	23-26	2201-2206 2300-2400	Midwest
Minneapolis-St. Paul, MN-WI	5122	Minnesota* Wisconsin	14-25 12	900 1100-2400 1100	Midwest
Mobile, AL	5160	Alabama	19-21	700-800	South
Modesto, CA	5170	California	22-23	2400-2500	West
Monmouth-Ocean, NJ	5190	New Jersey	32-37	2400-3100	Northeast
Montgomery, AL	5240	Alabama	17	1300 1500	South
Muncie, IN	5280	Indiana	19	2300	Midwest
Muskegon, MI	5320	Michigan	12	1200	Midwest
Nashua, NH	5350	New Hampshire	3	503	Northeast
Nashville, TN	5360	Tennessee	8-11	501-505 2300-2500	South
Nassau-Suffolk, NY	5380	New York	43-44	2401-2412 4601-4609 4700-4900	Northeast
New Bedford-Fall River-Attleboro, MA	5403	Massachusetts	32 34-37	3800-4300	Northeast

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
New Haven-Waterbury- Meriden, CT	5483	Connecticut	6-10 22-23	1800-2400	Northeast
New London-Norwich, CT	5523	Connecticut	11-14	2600-2700	Northeast
New Orleans, LA	5560	Louisiana	19 21-26	1700 1901-1904 2000-2300	South
New York, NY	5600	New York	34 36-42	4000 4301-4302 4401-4405 4500 5001-5010 5101-5110 5201-5203 5301-5318 5401-5414	Northeast
Fairfield County, CT	5602	Connecticut	1-6 21 23-25	1100-1700	Northeast
Newark, NJ	5640	New Jersey	13-25	2800 3200-3500 3601-3602 3700-4400	Northeast
Niagara Falls, NY	5700	New York	25	2301-2302	Northeast
Norfolk-Virginia Beach- Newport News, VA	5720	Virginia	16-21 26	2300 2500-3100 3300	South
Oakland, CA	5775	California	16-18	1700 1801-1805 2000 2101-2109	West
Ocala, FL	5790	Florida	12	4100	South
Odessa, TX	5800	Texas	30	3300	South
Oklahoma City, OK	5880	Oklahoma	10-13	900-1100	South
Olympia, WA	5910	Washington	12	1200	West
Omaha, NE-IA	5920	Iowa Nebraska*	19 6-7	1700 900 1001-1004	Midwest
Orange County, NY	5950	New York	35	4100-4200	Northeast
Orlando, FL	5960	Florida	17-19	1600-2300	South
Owensboro, KY	5990	Kentucky	20	300	South
Oxnard-Ventura, CA	6000	California	38-39	6701-6705	West
Parkersburg-Marietta, WV-OH	6020	Ohio West Virginia*	61 4	2900 300	South
Pascagoula, MS	6025	Mississippi	15	1500	South

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Pensacola, FL	6080	Florida	1	100	South
Peoria, IL	6120	Illinois	8-9	700-900	Midwest
Philadelphia, PA-NJ	6160	New Jersey Pennsylvania*	41-48 55-72	4500-5300 2601-2611 2701-2703 2801-2805 2901-2903 3001-3003	Northeast
Phoenix, AZ	6200	Arizona	8-9	101-116	West
Pine Bluff, AR	6240	Arkansas	16	1300	South
Pittsburgh, PA	6280	Pennsylvania	24-34	1301-1312 1400 3301-3303 3901-3902	Northeast
Pittsfield, MA	6323	Massachusetts	1	100	Northeast
Portland, ME	6400	Maine	6	600	Northeast
Portland, OR	6440	Oregon	2-7	1000-1500	West
Stafford County, NH	6450	New Hampshire	6	400	Northeast
Providence-Pawtucket-Woonsocket, RI	6483	Rhode Island	1-6	100-800	Northeast
Provo-Orem, UT	6520	Utah	6	500	West
Pueblo, CO	6560	Colorado	15	1200	West
Racine, WI	6600	Wisconsin	22	2100	Midwest
Raleigh-Durham, NC	6640	North Carolina	15-19	2301-2303 2400-2500 2700	South
Rapid City, SD	6660	South Dakota	1	100	Midwest
Reading, PA	6680	Pennsylvania	49-50	3101-3102	Northeast
Redding, CA	6690	California	3	300	West
Reno, NV	6720	Nevada	3-4	100 300-400	West
Richland-Kennewick-Pasco, WA	6740	Washington	7-8	700-800	West
Richmond-Petersburg, VA	6760	Virginia	22-25	1800-2100 2400	South
Riverside-San Bernardino, CA	6780	California	47-50	3700 6800 6901-6905 7000-7100 7201-7207	West

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Roanoke, VA	6800	Virginia	5-6	400-500	South
Rochester, NY	6840	New York	19-22	1900-2200	Northeast
Rockford, IL	6880	Illinois	2-3	300-400	Midwest
Sacramento, CA	6920	California	7-11	700 1000-1200 2801-2803 2901-2906	West
Saginaw-Bay City-Midland, MI	6960	Michigan	8-9 11	800-900 1100	Midwest
St. Cloud, MN	6980	Minnesota	26	800	Midwest
St. Joseph, MO	7000	Missouri	2	200	Midwest
St. Louis, MO-IL	7040	Illinois Missouri*	18-22 22-26	1700-2100 1101-1104 1201-1203 1300-1500	Midwest
Salinas-Seaside-Monterey, CA	7120	California	33	3800 3901-3902	West
Salt Lake City-Ogden, UT	7160	Utah	1-5	100-400	West
San Angelo, TX	7200	Texas	29	3500	South
San Antonio, TX	7240	Texas	40-42	4401-4408 4501-4502 4600	South
San Diego, CA	7320	California	51-52	3301-3313	West
San Francisco, CA	7360	California	14-15 19	1501-1502 1901-1906 2201-2206	West
San Jose, CA	7400	California	30-31	3401-3411	West
Santa Barbara-Santa Maria-Lompoc, CA	7480	California	35	3201-3202	West
Santa Cruz, CA	7485	California	32	3600	West
Santa Rosa-Petaluma, CA	7500	California	13	1300 1401-1402	West
Sarasota, FL	7510	Florida	28	3101-3102	South
Savannah, GA	7520	Georgia	21-22	100 2400	South
Scranton--Wilkes-Barre, PA	7560	Pennsylvania	7-15	500-700 1800 3200	Northeast

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Seattle, WA	7600	Washington	17-19	1001-1004 1701-1703 1801-1808	West
Sharon, PA	7610	Pennsylvania	4	1200	Northeast
Sheboygan, WI	7620	Wisconsin	27	300	Midwest
Sherman-Denison, TX	7640	Texas	7	700	South
Shreveport, LA	7680	Louisiana	1-2	100	South
Sioux City, IA-NE	7720	Iowa* Nebraska	6-7 9	600 500	Midwest
Sioux Falls, SD	7760	South Dakota	5	500	Midwest
South Bend-Mishawaka, IN	7800	Indiana	6-7	1500-1600	Midwest
Spokane, WA	7840	Washington	5-6	500-600	West
Springfield, IL	7880	Illinois	15	1400	Midwest
Springfield, MO	7920	Missouri	15-16	1900 2400	Midwest
Springfield, MA	8000	Massachusetts	2-6	200-700	Northeast
State College, PA	8050	Pennsylvania	18	1700	Northeast
Steubenville-Weirton, OH-WV	8080	Ohio* West Virginia	38-39 1	2000, 2500 100	Midwest
Stockton, CA	8120	California	20-21	2301-2304	West
Syracuse, NY	8160	New York	13-16	400 1400-1700	Northeast
Tacoma, WA	8200	Washington	13-14	1301-1304	West
Tallahassee, FL	8240	Florida	5	600	South
Tampa-St. Petersburg- Clearwater, FL	8280	Florida	11 21-26	4000 4200-4500	South
Terre Haute, IN	8320	Indiana	28-29	1900 3200	Midwest
Texarkana, TX-AR	8360	Arkansas Texas*	14 8	1100 800-900	South
Toledo, OH	8400	Ohio	1-4	100-200 3601-3604	Midwest
Trenton, NJ	8480	New Jersey	49-50	5400-5500	Northeast
Tucson, AZ	8520	Arizona	5-6	201-205	West
Tulsa, OK	8560	Oklahoma	4-7	100 600-700	South
Tuscaloosa, AL	8600	Alabama	12	400	South

MSA Name	FIPS Code	State	County Group Codes, 1980	PUMA Codes, 1990	Region
Tyler, TX	8640	Texas	12	1200	South
Utica-Rome, NY	8680	New York	10-12	501-502 600	Northeast
Vallejo-Fairfield-Napa, CA	8720	California	12	900 1601-1603	West
Vancouver, WA	8725	Washington	10	1901-1902	West
Victoria, TX	8750	Texas	67	7100	South
Vineland-Millville-Bridgeton, NJ	8760	New Jersey	40	5600	Northeast
Visalia-Tulare-Porterville, CA	8780	California	26	3500	West
Waco, TX	8800	Texas	48	1700	South
Washington, DC-MD-VA	8840	Washington, DC Maryland* Virginia	1 3 9-11 27-31	101-105 400 700 900 1201-1206 1301-1307 800-1100 2200	South
Waterloo-Cedar Falls, IA	8920	Iowa	4	400	Midwest
Wausau, WI	8940	Wisconsin	10	900	Midwest
West Palm Beach-Boca Raton- Delray Beach, FL	8960	Florida	33-35	2901-2906	South
Wichita, KS	9040	Kansas	4-6	400-600	Midwest
Wichita Falls, TX	9080	Texas	5	500	South
Williamsport, PA	9140	Pennsylvania	16	800	Northeast
Wilmington, DE-NJ	9160	Delaware* New Jersey	1 39	301-304 300	South
Wilmington, NC	9200	North Carolina	29	3400	South
Worcester-Fitchburg- Leominster, MA	9243	Massachusetts	7-10 25-26	800-1300 3000 3200-3300	Northeast
Yakima, WA	9260	Washington	9	900	West
York, PA	9280	Pennsylvania	38-40	2300 2501-2503	Northeast
Youngstown-Warren, OH	9320	Ohio	19-22	4801-4802 4901-4902	Midwest