The Impact of Rising Inequality on Health at Birth

Anna Aizer
Brown University and NBER

Florencia Borrescio-Higa
Brown University

Hernan Winkler
World Bank

Abstract: The income distribution has widened significantly over the past 40 years, primarily due to skill biased technological change which has disproportionately raised the income of the most highly skilled. There are a number of reason why, in theory, we might expect that increase in inequality to result in worse health outcomes. These include increases in psychosocial stress associated with low relative position and "pecuniary externalities" which refers to the fact that the price of medical goods and services might rise in an area with high average income, leaving the relatively poor worse off. We estimate the impact of rising inequality over the period 1970-2010 on health at birth. We define inequality two ways: as a group-level measure (the Gini coefficient for each county) and as an individual-level measure of relative deprivation which is the distance between an individual woman's own income and the average income of her reference group (all other women in her county). Unlike most previous empirical work, we use individual level data and instrument for inequality, exploiting skill biased technological change over this period which resulted in large increases in inequality in areas with a wide distribution of skill, but much smaller increases in inequality in areas with a much more narrow skill distribution.

I. Introduction

Income inequality has been on the rise in most industrialized nations since the 1970s. In the US, for example, the Gini coefficient increased from .39 in 1970 to .47 by 2010. There has been considerable discussion of the causes of the rise in income inequality. Most research based on developed countries points to the increase in skill-biased technological change and globalization as the most important factors (add cites).

While there has been considerable work trying to understand the causes of rising income inequality, there is less research that attempts to understand its consequences. In this paper, we estimate the impact of rising inequality over the past 40 years on health at birth. We focus on health at birth for multiple reasons. First, newborn health is largely, though not entirely, a function of short term conditions. This makes it easier to isolate the economic conditions affecting health. Second, health at birth has been shown to be an important determinant of long term outcomes such as educational attainment, IQ and earnings (Black, Devereaux and Salvanes, 2007). Third, individual-level data on birth outcomes has been collected and reported consistently at a local geographic level (county) for the period 1970-2010, allowing one to estimate the impact of changes in inequality and relative income at a local level on individual outcomes.

In our estimates of the impact of rising inequality on health at birth, we define inequality in two ways. First, we define it as the Gini coefficient for the local area (county). This measure is common to all individuals in the area. Second, we calculate an individual measure of relative deprivation which is the distance between an individual's income and the average income of the individual's reference group. In this case, the reference group is other women in the same county. We then estimate whether changes in either the Gini or relative deprivation affect birth outcomes controlling for one's own (absolute) income level. However, this is not sufficient to establish a causal relationship. Differences in an individual's income and the income of one's neighbors could simply reflect differences in underlying maternal characteristics that may affect newborn health directly.

One of the main contributions of this paper is that we address the potential endogeneity of income (both relative and absolute). Previous work on this topic has not. To do so, we instrument for income using IV strategies based on those developed by Bartik (1991) and Boustan, Ferreira, Winkler and Zolt (2012).[1] The instruments allow us to isolate the change in both individual income and the local income distribution that is driven by national shifts in the return to skill over time. More specifically, we construct an instrument for local (county-level) distribution of income by holding the local county income (education) distribution fixed at the

[1] The Bartik instrument of labor demand has been employed by Blanchard and Katz (1991), Bound and Holzer (2000), Autor and Dugger (2003) among others.

1970 distribution and match this initial distribution to national patterns in income growth for different points in the distribution of income (education).[2]

In this very preliminary draft, we present only OLS results for birth weight based on the period 1970-2000 for large counties. We have applied for and been granted access by the CDC to natality and linked mortality data with finer geographic identifiers (ie, smaller counties identified) for the period 1970-2010. We anticipate receiving these data in the coming month. Future work will incorporate the additional data, include the IV results, experiment with different reference groups and measures of relative deprivation, and also include results for additional outcomes: mortality, gestation and behavioral changes (initiation of prenatal care).

II. Background

A. Inequality and relative deprivation

Inequality and relative deprivation are closely related but distinct concepts. Areas characterized by greater inequality have higher relative deprivation. However, inequality characterizes an entire group of individuals, whereas relative deprivation is specific to an individual within a group.

There are three reasons why inequality could be related to health. The first has to do with non-linearities in the production of health. If maternal income produces child health (see Case, Lubotsky and Paxson, 2002), but is marginally more productive at low levels of income than at high levels, then an increase in inequality would reduce average health. This could explain the link between income inequality and mortality that is observed in aggregate data at the country, state or metropolitan level. We refer to this channel as an "indirect" effect.

The second has to do with relative status and the higher levels of psychological stress that characterize those who are relatively worse off than their peers. In this formulation, relative deprivation may be measured in terms of income, education, social or political status, all of which are correlated. Individuals who feel greater stress because of their lower status may be stressed and/or depressed and also more likely to engage in behaviors that negatively affect health (poor eating and exercise habits, smoking, etc). In previous work, Aizer, Stroud and Buka

[2] For example, consider two counties, A and B. In 1970, county A had a disproportionate share of women in the bottom quartile of the (national) education distribution at the time, while county B had a disproportionate share of women in the top quartile of the (national) education distribution. By 1980, the distribution of education county A had changed so that it is more similar to the distribution of education in county B. However the instrument for the distribution of income in county A is calculated by holding the distribution of education fixed at the 1970 level and then predicting the income distribution based on national trends in income growth for the initial distribution of education.

(2012) found that the stress hormone cortisol is higher in poor women and that higher levels of cortisol during the prenatal period is associated with worse outcomes for the offspring (lower IQ, worse health and less completed schooling). These data, however, did not allow for the examination of the direct role or impact of inequality on cortisol levels and offspring outcomes.

Much of the evidence regarding a causal impact of social status on health is drawn from research on primates (Sapolsky et al, 1997; Cohen et al, 1997). In this work, researchers have documented a strong relationship between social status within a group and health outcomes. This relationship has been found to persist even after experimental manipulation of a primate's rank or status within a group, something that is impossible to do with humans.

A third and final reason why inequality and health might be related "pecuniary externalities" associated with having neighbors who are relatively more wealthy. Depending on the nature of the externality, however, having wealthy neighbors could be either beneficial or harmful to the less wealthy. Examples of a positive pecuniary externality includes an increase in the availability of certain medical services that have high fixed costs and as a result only locate in counties with high average levels of income (assuming income and demand are positively correlated). As a result, relatively poor individuals in the same county benefit from the increased availability of this service. An example of a negative externality is when an increase in average income in an area is associated with an increase in private expenditures on health and offsetting reduction in public expenditures on health, making the relatively poor worse off. [3]

Of the three potential mechanisms by which inequality might be associated with health, the first is relevant only in the context of aggregate data. Because we have individual data that can be aggregated, we can explore the extent to which non-linearities in the production of health can explain the relationship between the Gini and average health observed in aggregate data.

With respect to the other two mechanisms, stress and pecuniary externalities, we will be able to shed some light on the relative importance of the latter by examining how inequality affects pollution levels (a public good), NICU availability (a service with high fixed costs) and the presence of local clinics (public expenditures).

B. Previous Evidence

The first evidence of a relationship between inequality and health was presented by Rodgers in 1979 who conducted cross-country analyses of the relationship between gini coefficients, GNP and multiple measures of health (infant mortality, life expectancy of birth) and found a strong relationship between both income and inequality and health. Since then, several others have

[3] The policy implications of the first and the last two mechanisms differ. With respect to the first mechanism (heterogeneity and non-linearities in the production of health), increasing the wealth of the wealthiest individuals would not affect health. In contrast, for the latter two mechanisms, it would.

conducted within country analysis of the relationship between inequality and health. Kennedy (1996) provided evidence of a strong relationship between mortality and inequality across the 50 US states.

However, Deaton and Lubotsky (2002) show that for the cross-state and cross-MSA level analyses for the US, once one controls for the share of the local area that is black, the negative relationship between inequality and average mortality disappears. This is due to the fact that inequality is higher in areas with a large share black and so is mortality (for both blacks and whites), though for reasons unknown. As such, the share black is an important omitted variable in analyses of the relationship between inequality and average health.

There is much less work examining the impact of relative deprivation. Evans and Eibner (2005) represent the first serious attempt to do so. They estimate the impact of relative deprivation on mortality from cardio-vascular disease using NHIS survey data linked with mortality data from the period 1988-1991. They define a reference group as those of the same race, age and education class in one's own state of residence. This is largely because the NHIS data only includes state identifiers and nothing more local. They find that relative deprivation is predictive of higher mortality, worse self-reported death, higher BMI and riskier behavior, controlling for own income. Evans and Eibner (2005) employ multiple measures of relative deprivation and include a discussion of the advantages and disadvantages of each. For this proposal, we use only one measure of relative deprivation (the z-score) but will also use a second measure of relative deprivation developed by Runciman (1966) and refined by Yizhaki (1979).

Miller and Paxson (2006) also estimate the impact of relative income on mortality in an effort to explain black-white differentials in mortality. Specifically, they assess the extent to which blacks' lower relative income can explain their higher rates of mortality. Using mortality data at the county level and census data on income at the puma (county-group) level, they find that within a county, the average income of blacks affects the black mortality rate, but that the average income of whites in the county also affects the black mortality rate. In fact, for black men, an increase in the average income of white residents of the same county increases the black mortality rate by about half as much as a similar-size decrease in black income.

III. Data and Empirical Strategy

A. Data

1. Data on birth outcomes and maternal characteristics

The data on birth outcomes (birth weight, LBW, gestation, mortality) and intermediate health behavior (initiation of prenatal care) will come from vital statistics birth and matched mortality records for the period 1970-2010. These data include the education, age, race and marital status of the mother and father as well as county-identifiers for counties with a population of at least

100,000. Data on smoking was not collected until 1990 and data on Hispanic origin was not collected until 1980. Prior to 1980, there were few births to Hispanic women, but because the gradient in health among Hispanic women is very different (eg, there is either no gradient or a reverse gradient for Hispanic immigrants), for some analyses we drop Hispanic women from the data as well as all births in states with large Hispanic populations (California, Arizona, Florida and Texas).

2. Data on "pecuniary externalities"

In order to assess the extent to which any estimated effect of inequality or relative deprivation on health is operating through pecuniary externalities, we are collecting data on potential externalities in the context of newborn health, including the number of NICUs (from multiple sources, including the AHA, and the Maternal and Child Health Bureau), local pollution levels and the availability of public clinics (federally qualified health centers, or FQHCs).

3. Constructing measures of own income, the Gini coefficient, and relative income

Own income

The vital statistics natality data do not include information on income. Rather, we will use the information on maternal education to impute income based on the relationship between education and income in that county and year for women between the ages of 22 and 40 from decennial census data. There are some advantages to using this measure of imputed income. First, educational attainment is arguably a better predictor of permanent income than a spot measure of income. Second, income imputed based on education is less likely to suffer the endogeneity bias of an actual income measure, though we will instrument for this measure anyway.

Thus, the measure of own income (y_i) will be the average income of all women with the same level of education in the county and year of birth.

$$y_i = S_i * \beta_{sc}$$

where S_i is the level of schooling of mother I and β_{sc} is the return to level of schooling (s) in county c. We construct two measures of income. The first one is based on personal income, which measures each respondent's total pre-tax personal income from all sources. The second one is based on per capita income, which is defined as total pre-tax income earned by the family divided by the number of members of the family.

Gini coefficient

We construct Gini coefficients for each county in 1970, 1980, 1990, 2000 and 2010. The Gini coefficient is derived from household income data collected as part of the decennial census data

for 1970, 1980, 1990 and 2000. The Gini for 2010 comes from household income data collected as part of the American Community Survey (ACS), three year average, 2008-2010 due to the absence of a 2010 decennial census long form.

Relative Income

The first measure of relative income we use will be the individual's z-score. This measure was used previously by Eibner and Evans (2005). The z-score is the number of standard deviations away from the group mean. The group in this context is other women living in one's county.

$$\text{The z-score is: } \frac{(y_i - \mu_c)}{\sigma_c}$$

Where y_i is the individual mother's own imputed income, μ_c is the average income of mothers in the county (the reference group) and σ_c is the standard deviation of income of mothers in the county. Data on the average/standard deviation of income in the county and year come from census data and is therefore not imputed.

B. Empirical Strategy

1. IV Analysis

One of the main contributions of this work is to address the potential endogeneity of own income, relative income and the Gini by instrumenting for these measures. Our identification strategy takes advantage of the large increase in the return to skill witnessed over the past 40 years which resulted in large increases in inequality, as those with higher levels of education saw disproportionately greater gains in the returns to education over this period. However, the increase in inequality over this period was not uniform, with some local areas witnessed much larger increases in inequality than others. Boustan et al (2012) show that between 1970 and 2000, the Gini coefficient in the average municipality (defined as a city or school district) increased from 0.32 to 0.38, but that in one third of municipalities, the increase was less than one point (or even decreased), and that in another third of municipalities the increase was more than five points. For identification, we exploit the variation in the growth of the Gini over this period that derives from differences in the initial (1970) distribution of income or skill and national trends in income growth or returns to skill over this period.

Constructing the instruments for own income, the Gini and relative income

The instruments for income are constructed so that they reflect variation in income derived from national trends in the returns to skill. For own income, the instrument is the average income that

corresponds to an individual woman's level of education, for the entire nation outside of county c.

$$Z_{Ii} = S_i * \beta_{s\,-c}$$

Where S_i is schooling level of mother i and $\beta_{s\,-c}$ is the return to schooling level s in all counties except mother i's own county.

To instrument for the distribution of income and relative income, we follow Boustan et al (2012) and hold the income distribution of the county fixed at its 1970 level and predict changes in the distribution of income based on the initial income distribution in 1970 and national trends in income growth over this period. In this way, to borrow the language of Boustan et al (2012), we construct a "synthetic" version of the income distribution in each county. We will also implement a modified version of this instrument that will be based on the initial distribution of education (not income) in the county to which we will apply national trends in the return to education. In practice, we anticipate the two instruments will be highly correlated, but the latter better captures the skill-biased technological change driving the increase in inequality.

IV. Very Preliminary Results

Table 1 shows only descriptive statistics on the distribution of the Gini coefficient across counties and over time. The main thing to take away from the table is the considerable variation both over time and across counties within each year in the Gini coefficient.

We beign with OLS estimates of the following equation:

$$\text{Birth Weight}_{ict} = \beta_0 + \beta_1 \text{Income}_i + \beta_2 \text{Gini}_{ct} + \beta_3 X_i + \beta_4 G_{ct} + \beta_5 \text{Year}_t + \beta_6 \text{County}_c + \varepsilon$$

Where income is maternal personal income imputed based on her education level and the average income of women in her county and education level. The Gini coefficient is defined at the level of the county and year, X_i is a vector of maternal controls including age and marital (note that the first set of tables are restricted to white women only, so no controls for race are included). G_{ct} is vector of county-level controls (population size, share black, share Hispanic, poverty rate and share over 65), following Boustan et al (2012). We also include year FE and county FE.

OLS estimation of the above equation is presented in Table 2. As you can see, even conditional on personal income, which has appositive and significant effect on birth weight, the Gini coefficient has a negative effect on birth weight. An increase in the Gini of 0.08 (the average increase between 1970 and 2000), is associated with a 30 gram decline in birth weight (Table 2, column 1). In column 2 of Table 2, we present the same results but including an alternative

measure of maternal income (median per capita income based on family income) and the results are similar, though the effect of an increase in the Gini is greater (the same increase leads to a 40 gram decline).

In table 3, we present estimates of the above equation but including a measure of relative income (the Z-score) instead of the Gini. Recall that the Gini is the same for all women in the county whereas the Z-score is an individual-level measure of relative income or deprivation. OLS estimates suggest that both personal or absolute income and relative income matter. Conditional on an absolute level of income, a woman with income one standard deviation above the mean of her county gives birth to a child weighing 20 grams more, on average. In column 2 of Table 3 we present IV estimates, which are smaller, but still precise. The first stage results, not shown, are quite strong, with t statistics of 25.

While these results are very preliminary, they are highly suggestive of a significant and economically meaningful relationship between inequality and health at birth. The results will have important implications for policy makers interested in understanding the consequences of rising inequality, in general, but also more specifically for the intergenerational transmission of income and subsequent inequality in human capital of the next generation.

References (incomplete)

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Black, Sandra, Devereaux, Paul and Kjell Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics.*

Boustan, Leah, Ferreiro, Fernanado, Winkler, Hernan and Eric Zolt. 2012. "Income Inequality and Local Government in the United States, 1970-2000" *Review of Economics and Statistics*.

Cohen, Sheldon, Line, Scott, Manuck, Stephen, Rabin, Bruce, Heise, Eugene and Jay Kaplan. 1997. "Chronic Social Stress, Social Status and Susceptibility to Upper Respiratory Infections in Non-Human Primates." *Psychosomatic Medicine* 59(3): 213-212.

Deaton, Angus & Lubotsky, Darren, 2003. "Mortality, inequality and race in American cities and states," Social Science & Medicine, Elsevier, vol. 56(6), pages 1139-1153, March. Elsevier, vol. 56(6), pages 1139-1153, March.

Eibner, Christine and William Evans. 2005. "Relative Deprivation, Poor Health Habits and Mortality" Journal of Human Resources.

Miller, Douglas and Christina Paxson. 2006. "Relative income, race, and mortality," *Journal of Health Economics*. Vol 25. pp. 979-1003.

Runciman, Walter Garrison. 1966. Relative Deprivation and Social Justice. California: University of California Press.

Yizhaki (1979).

Sapolsky, Robert, ALberts, Susan, and Jeanne Altmann. 1997. "Hyocotisolism Associated with Socail Subordinance or Social Isolation among Wild Baboons" Archives of General Psychiatry. 54(12): 1137-43.

Table 1: Variation in County Level Gini Coefficients

	1970	1980	1990	2000
Average	0.323	0.342	0.378	0.404
Standard Deviation	0.034	0.027	0.032	0.032
10th percentile	0.281	0.314	0.341	0.366
25th percentile	0.302	0.323	0.356	0.384
median	0.318	0.339	0.374	0.400
75th percentile	0.342	0.360	0.401	0.424
90th percentile	0.370	0.376	0.419	0.446

Table 2: Local Gini and Birth Weights for White Women aged 22-40, OLS

	(1) Birth Weight	(2) Birth Weight
Gini Coefficient	-369.199**	-497.862***
	[168.704]	[174.331]
Log median personal income	7.324***	
	[1.497]	
Log median per capita income		49.555***
		[6.422]
Dummy: year 1980	71.860***	81.567***
	[7.903]	[6.508]
Dummy: year 1990	86.702***	93.854***
	[12.011]	[10.922]
Dummy: year 2000	78.220***	89.528***
	[16.446]	[15.326]
Dummy: over 35 years old	2.298	-1.379
	[3.042]	[3.074]
Dummy: married	130.635***	123.579***
	[6.237]	[6.127]
Population (county level)	0.000***	0.000***
	[0.000]	[0.000]
Share of black population (county level)	-114.739***	-135.252***
	[23.995]	[23.166]
Share of hispanic population (county level)	-85.214***	-78.604***
	[28.887]	[29.256]
Poverty rate (county level)	109.417	289.102***
	[76.722]	[77.939]
Share oover 65 years of age (county level)	-1.794	56.314
	[68.968]	[70.947]
Constant	3,281.310***	2,892.136***
	[49.614]	[72.147]
Observations	529,887	529,887
R-squared	0.008	0.009

Robust standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Impact of Relative Income (Z-Score) on Birth Weight, White Women aged 22-40, OLS and IV

	(1) OLS	(2) IV
Log median personal income	29.222**	22.543*
	[12.735]	[13.567]
Z-score median personal income (county reference group)	17.697***	16.827**
	[2.486]	[6.866]
Dummy: over 35 years old	-2.159	2.792
	[2.805]	[4.000]
Dummy: married	118.886***	133.269***
	[5.605]	[3.897]
Population (county level)	-0.000	0.000
	[0.000]	[0.000]
Share of black population (county level)	-88.296	-265.509**
	[88.192]	[105.951]
Share of hispanic population (county level)	107.491	260.495***
	[77.737]	[79.293]
Poverty rate (county level)	-74.051	-75.529
	[96.091]	[87.428]
Share oover 65 years of age (county level)	198.818	-45.758
	[166.491]	[204.592]
Constant	3,231.067***	3,064.239***
	[29.037]	[86.263]
Observations	528,721	258,821
Number of countyfips	2,004	890

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Year dummies and county dummies also included