

NBER WORKING PAPER SERIES

INCOME INEQUALITY, SOCIAL MOBILITY, AND THE DECISION TO DROP
OUT OF HIGH SCHOOL

Melissa S. Kearney
Phillip B. Levine

Working Paper 20195
<http://www.nber.org/papers/w20195>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2014

The authors thank Miles Corak, Sue Dynarski, Nora Gordon, Caroline Hoxby, Judy Hellerstein, Robin McKnight, and Lesley Turner for helpful conversations and comments on this paper. We also acknowledge helpful comments from seminar participants at American University School of Public Policy, Notre Dame University, Stanford University, University of New Hampshire, University of Texas at Austin, the Federal Reserve Bank of Cleveland Workshop on Income Distribution, and the NBER Universities Research Conference on Poverty, Inequality, and Social Policy. We are grateful to the Smith Richardson Foundation for providing financial support for this project. Any views expressed are those of the authors alone. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2014 by Melissa S. Kearney and Phillip B. Levine. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Income Inequality, Social Mobility, and the Decision to Drop Out of High School
Melissa S. Kearney and Phillip B. Levine
NBER Working Paper No. 20195
June 2014
JEL No. D31,I24,J24

ABSTRACT

This paper considers the role that high levels of income inequality and low rates of social mobility play in driving the educational attainment of youth in low-income households in the United States. Using high school degree status from five individual-level surveys, our analysis reveals that low-socioeconomic status (SES) students, and particularly boys, who grow up in locations with greater levels of lower-tail income inequality and lower levels of social mobility are relatively more likely to drop out of high school, conditional on other individual characteristics and contextual factors. The data indicate that this relationship does not reflect alternative characteristics of the place, such as poverty concentration, residential segregation, or public school financing. We propose that the results are consistent with a class of explanations that emphasize a role for perceptions of one's own identity, position in society, or chances of success. In the end, our empirical results indicate that high levels of lower-tail income inequality and low levels of social mobility hinder educational advancement for disadvantaged youth.

Melissa S. Kearney
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
kearney@econ.umd.edu

Phillip B. Levine
Department of Economics
Wellesley College
106 Central Street
Wellesley, MA 02481
and NBER
plevine@wellesley.edu

I. INTRODUCTION

International comparisons show that the United States is a country that ranks high in its level of income inequality and low in its level of social mobility. Corak (2006) built on the theoretical contributions made by Solon (2004) and was the first to show empirically that this relationship is part of a broader pattern that exists across countries. Those with high inequality also tend to exhibit lower social mobility, as measured by greater intergenerational income persistence. In a speech on January 12, 2012, Council of Economic Advisors Chairman Alan Krueger popularized this relationship as “The Great Gatsby Curve.” Using our own data that we describe later along with that from Chetty, et al. (2014a), we construct a similar Gatsby Curve across states in the United States: Figure 1 shows an upward sloping relationship between state level income inequality and intergenerational income persistence that resembles the international pattern.

The combination of high income inequality and low economic mobility leads to concerns about the prospects for today’s disadvantaged youth.¹ The 2012 *Economic Report of the President* stated, “The confluence of rising inequality and low economic mobility over the past three decades poses a real threat to the future of the United States as a land of opportunity. Social and economic mobility across generations are at risk of declining unless concerted efforts are devoted to providing more opportunities for those born into lower-income households (p. 181).”

This is an important concern that requires a careful empirical analysis. Its premise is not necessarily true. Corak (2013) notes the possibility that the combination of high inequality and low mobility might in part reflect something about the composition of the population. Places like

¹Social mobility is a concept that includes the likelihood of moving up or down in the income distribution, which is specifically labeled as economic mobility, but also may include changes in position in other distributions as well, like educational attainment, occupational status, and health. We restrict our attention to economic mobility in this paper, but adopt the common approach of using the terms social and economic mobility interchangeably.

the U.S. tend to be more demographically diverse relative to places like Denmark, which has considerably less inequality and greater mobility. Mankiw (2013) similarly observes that low social mobility could occur even if there was equality of opportunity because of the inheritability of talent, intellect, and interpersonal skills. If a population is comprised of individuals with a large degree of variation in talents and abilities, then we might expect both inequality in income and high persistence in income between parents and children. As such, disadvantaged youth might fare worse in the U.S., but that would simply be a reflection of the nature of its population, rather than a consequence of its level of inequality and mobility.

It is not even clear the extent to which the empirical discussion regarding these points accurately captures the issues at stake. It is common to see references regarding the impact of inequality on mobility.² In reality, though, the data available to us indicate whether economic inequality that exists today (in period t) is related to social mobility that has taken place in the recent past (between, say, period $t-1$ and period t). What we are able to document by these Gatsby Curves is that these outcomes are positively correlated; there is no empirical way to attribute causation to their relationship. What we focus on in this paper, though, is the relationship between the inequality that exists today, the mobility that has occurred up to today, and its impact on the outcomes of today's children going forward (between periods t and $t+1$). We reference these concepts independently as inequality, mobility, and opportunity.

Furthermore, it is unclear what the transmission mechanism is between high inequality, low mobility and subsequent outcomes for today's youth, should such a relationship exist. Greater residential segregation could further restrict access to better jobs. Discrimination against

² For example, Jason Furman, Chair of the Council of Economic Advisers, stated "I think we think all else being equal, more inequality will lead to less relative mobility," in *The Atlantic's* Economy Summit 2014 (available at: http://fora.tv/2014/03/18/conversation_with_jason_furman_-_7:37 into the interview).

the poor could hinder subsequent outcomes. Those facing greater inequality and less mobility may have a more difficult time advancing through the educational system, among other possible mechanisms.

This paper considers one specific, albeit important, potential outcome of high inequality and low mobility: the likelihood that disadvantaged youth drop out of high school. Cross-sectional comparisons indicate that places with higher levels of inequality and higher levels of income persistence tend to have higher rates of high school drop-out behavior. But this correlation cannot be taken to imply a causal relationship; for instance, places with higher levels of inequality and lower levels of mobility might have more economically disadvantaged youth, who are more likely to drop out of school than youth from higher-income families. To determine whether this relationship is causal, we use individual-level data pooled from five national surveys to investigate how income inequality and income persistence affect the rate at which low socioeconomic status (SES) youth drop out of high school, controlling for individual background and aggregate-level contextual factors.

We find that both higher levels of lower-tail income inequality and lower levels of social mobility lead teens from low-SES households to drop out of high school with greater frequency, controlling for a rich set of individual and aggregate level characteristics. Because income inequality and mobility are so strongly correlated, we are unable to distinguish whether one of the two factors is a more important determinant. Our analysis rules out that this link is being driven by a number of potential confounding factors, such as other features of the income distribution (including upper-tail income inequality), aggregate poverty rates, or incarceration rates. The data indicate that up to a third of the effect can be explained by differences in ability or achievement, as measured by AFQT scores. We discuss alternative interpretations of that

finding -- noting that the AFQT captures both innate and learned cognitive skills – and emphasize that the bulk of the effect remains unexplained either way.

The second goal of our paper is to consider the potential mechanisms that drive the empirical relationship between lower-tail income inequality, low economic mobility and the decision of low-SES youth to drop out of school. The data do not offer support for a number of alternative channels, including residential segregation and public school funding. Instead, we speculate that the empirical link we are finding is explained by the effect that high inequality and low mobility has on low-income individuals' *perceptions*. This class of explanations consists of models that highlight a role for perceptions of identity (Akerlof and Kranton, 2000; Watson and McLanahan, 2011), relative societal position (e.g., Luttmer, 2005), and one's chance at economic success (Kearney and Levine, 2014; Genicot and Ray, 2014). Although we are ultimately unable to establish a precise mechanism, the empirical relationship we document is consequential, implying that greater levels of income inequality and lower rates of social mobility are perpetuating, in so far as they lead low-income youth to engage in more drop-out behavior.

II. BACKGROUND

A. Facts about Inequality and Mobility

Much attention has been paid recently to trends in inequality and mobility. It is now widely known that overall income inequality has increased over recent decades. This increase has been driven by a steady climb in incomes at the top of the distribution. The gap in income between the 90th and the 50th percentiles of the distribution has largely continued on a steady climb. Lower-tail income inequality, as measured by the ratio between the 50th and 10th percentiles of the distribution, has generally plateaued or compressed since the mid-1980s (see Autor, Katz, and Kearney, 2008). This is an important distinction because our analysis focuses on the 50/10 ratio. We do so because the gap between being poor and being in the middle of the

distribution is arguably more relevant to the lives of the economically disadvantaged than is the gap between being poor and the top 10 or 1 percent. We empirically test this proposition later in the paper.

Data on social mobility also indicates that rates of intergenerational income persistence in the United States remained steady for cohorts born between 1971 and 1993, albeit at a low level (Chetty, et al., 2014b). These combined facts suggest that there have not been substantial changes in either lower-tail income inequality or economic mobility, such that they would have led to a first-order effect on high school drop-out rates over the past few decades.

In contrast, differences across places in lower-tail income inequality and social mobility are large and persistent. A critical component of this analysis is our focus on longstanding differences in inequality and mobility, not transitory differences.³ This distinction is implied in a measure of intergenerational mobility; by nature it is constructed over decades. Inequality can be measured at much more frequent intervals, but the measure we use is intended to capture long-standing inequality. We do so because we are trying to capture something about the permanent or semi-permanent economic and cultural landscape in the place where an adolescent lives, as opposed to short-term fluctuations. If a state experiences a temporary decrease in lower-tail income inequality, it is unlikely, for example, that neighborhoods will change sufficiently quickly and sufficiently visibly that either economic opportunities or the perceptions thereof will be altered. It is also the case that there is much more cross-sectional variation in lower-tail income inequality across states as compared to within a state over time. In the income data we

³This contrasts with a typical empirical approach to conduct panel analyses controlling for state and year fixed effects and exploiting transitory variation in the explanatory variable of interest. Mayer (2001), for instance, uses this approach in her study of whether state level income inequality affects individual level educational outcomes. She uses the 1993 PSID data to exploit over-time variation in state-level income inequality, as measured by the GINI coefficient. In her regression models that include both state and year fixed effects, the results indicate a statistically insignificant relationship.

describe below that includes three Censuses and five years of data from the ACS, we find that the standard deviation in the 50/10 ratio across states averaged over time is 0.43. Using the same data, we find that the average standard deviation within states over time in the 50/10 ratio is 0.16.

Using a conceptual distinction we described earlier, our empirical analysis is meant to capture economic inequality that exists today (in period t), economic mobility that has taken place in the past (between, say, period $t-1$ and period t), and their impact on youths' opportunities to succeed in the future, as reflected by their economic advancement between periods t and $t+1$. The measure of income inequality we use has changed little over time; using its average value over multiple decades is a stable proxy for current inequality. Measures of mobility are, by nature, backward looking and also roughly constant over time. Economic decisions that children make, like dropping out of high school, alter subsequent well-being and capture our broad concept of opportunity to advance economically.

B. High School Dropout Rates and the Correlation with Inequality and Mobility

For at least the past 20 years, roughly one-quarter of students who begin 9th grade in the United States fail to graduate high school within four years (Snyder and Dillow, 2012). Recent empirical work by Heckman and LaFontaine (2010) finds that the high school graduation rate peaked at around 80 percent in the late 1960s, and has since declined by 4-5 percentage points. Their work also suggests that there has not been a convergence in majority/minority graduation rates, and that only 65 percent of black and Hispanic teens graduate high school today. Murnane (2013) reports large differences between children from high and low socioeconomic groups; he reports that 36 percent of 8th grade students in 1988 who were in the lowest socio-economic status quartile failed to graduate from high school, as compared to 5 percent of students in that cohort whose families were in the top quartile.

The relationships between inequality, mobility and high school dropout rates are striking: places with higher levels of income inequality and lower levels of economic mobility tend to have higher dropout rates.⁴ Figure 2 displays this relationship across states, where income inequality is measured by the 50/10 ratio of total household income, as described subsequently. The figure shows clearly that more students fail to graduate from high school on time in states with greater inequality. Around 40 percent of those who start high school in Louisiana and the District of Columbia fail to graduate in that time period, as compared to only around 15 percent of students in Vermont and Wisconsin; inequality is much greater in the former states. Figure 3 reports the analogous relationship replacing income inequality with the level of intergenerational income persistence, which is also positively related to high school dropout rates.⁵ Of course, the relationships in these figures are just correlations and do not necessarily imply a causal relationship.

C. Relevant Previous Research

Rich, descriptive information about the relationship between income inequality and educational attainment is provided in Duncan and Murnane's (2011) edited volume, *Whither Opportunity*. The chapter by Bailey and Dynarski (2011) details growing inequality in college attainment over time as income inequality has risen. Reardon (2011) documents growing inequality in test scores over time, driven mainly by the increased performance of children at the top of the income distribution. Reardon goes on to find that the specific timing of that increase does not quite match the timing in income inequality growth, suggesting that inequality may not

⁴Wilkinson and Pickett (2009) promote the idea that levels of income inequality across countries are related to negative social outcomes, including lower levels of education. The empirical evidence presented in that book, however, is purely correlational and based entirely on aggregate-level data.

⁵Our measure of intergenerational income persistence is the Chetty, et al. (2014a,b) measure that they label "relative mobility," which we describe subsequently.

be the main contributing factor. An important distinction between these papers and ours is that we focus on persistent differences in inequality across states, not changes in income inequality in the U.S. over time.

Virtually all of the most relevant empirical research focuses on the relationship between educational outcomes and income inequality, not social mobility. Data containing geographic variation in economic mobility rates across the U.S. simply did not exist before Chetty, et al. (2014a). This paper made an enormous contribution in this area by creating such measures using detailed administrative IRS data. The authors document substantial variation in mobility across the U.S. and demonstrate that places with higher rates of mobility also have less income inequality, better primary schools, greater social capital, and a lower share of single-parent households. All of these relationships merely reflect correlations in the data, as the authors recognize.

Focusing on the link between income inequality and educational attainment, alternative perspectives provide conflicting predictions regarding the direction of the relationship. From a Beckerian perspective, greater income inequality would *increase* the incentive to make human capital investments because the returns to doing so would be greater, *ceteris paribus* (Becker and Murphy, 2007). In contrast, an “ecological model” – described by Duncan and Murnane (2011) – presents a number of pathways through which income inequality could hinder the educational attainment of disadvantaged students. Such effects could manifest through the effect of income inequality on neighborhoods, families, and labor markets directly, or indirectly through the educational system. They do not address mobility, but extending their framework along these lines is straightforward.

A particularly useful feature of the ecological framework is that it easily incorporates existing insights. Residential segregation, for instance, has been described in a number of papers as a factor that affects the attitudes and behaviors of the poor. To the extent that higher income inequality is associated with increased residential segregation – as empirically demonstrated by Watson (2009) – this could be a pathway through which income inequality affects the educational attainment of disadvantaged youth. Greater residential segregation can affect social and labor market networks, the presence of high achieving role models, and the establishment of peer groups and norms.

The influential work of Wilson (1987) emphasizes the role of “social isolation” in driving rates of urban joblessness and non-marital childbearing. He hypothesizes that the lack of exposure to mainstream middle class role models plays an important role. Case and Katz (1991) provide an early example of non-experimental empirical research suggesting significant neighborhood peer effects for criminal behavior as well as the likelihood that youth are out of school and out of work.⁶

There also exists a body of work considering the political economy implications of income inequality for the public financing of public goods. Predictions from these models are ambiguous. More money in the hands of the rich may reduce transfers of resources to the poor or it may increase transfers if the rich become more socially fearful of the poor agitating for social change. Alternatively, the median voter model implies that increased inequality will lead to increased public good provision. With greater inequality, the median falls relative to the mean, and the preferences of the median voter for more distribution from the rich prevail. Recent

⁶The notion that neighborhood effects are important in driving economic outcomes implies that housing assistance could be an effective public policy tool. The Moving to Opportunity (MTO) experiment was designed to test this idea; the evaluations of that housing assistance demonstration project found no effect on labor market outcomes or children’s educational test scores (see, Kling, Ludwig, Katz, 2005; and Kling, Liebman, Katz, 2007).

empirical evidence on the relationship between income inequality and public revenue for school spending finds public school spending increases as the level of local income inequality rises (Boustan et al. 2013; Corcoran & Evans, 2010; Gordon, 2013). We will draw on all of these models in our empirical analysis and consider such pathways as potential mediating factors between income inequality/low mobility and educational attainment of low SES youth.

III. EMPIRICAL STRATEGY

The goal of our econometric analysis is to determine whether individuals from disadvantaged backgrounds who live in areas with high rates of income inequality and low rates of economic mobility experience higher high school dropout rates.⁷ We have shown that cross-sectional differences in the level of inequality/mobility and high school dropout rates are sizable and correlated. In our empirical analysis, we incorporate this by including geographic (state or MSA) fixed effects. What we want to know is whether low SES youth in high inequality locations are even more likely to drop out high school. We capture this effect with interactions between inequality (or mobility) and indicators of SES status for an individual. We include other individual and state level controls in the model so that the estimated effect of inequality for low-SES adolescents is the net of effects driven by other factors that might be correlated with measures of inequality and mobility.

For ease of exposition, in this discussion of the empirical model we refer to inequality as the explanatory variable of main interest, rather than repeatedly noting that we are interested in both income inequality and economic mobility. The formal econometric model takes the following form:

⁷This empirical strategy is conceptually identical to that which we used in our analysis of early, non-marital childbearing in Kearney and Levine (2014).

$$\begin{aligned}
Outcome_{isc} = & \beta_0 + \beta_1(I_s \cdot LS_{is}) + \beta_2(I_s \cdot MS_{is}) + \beta_3LS_{is} + \beta_4MS_{is} \\
& + \beta_5X_{is} + \beta_6E_{sc} + \gamma_s + \gamma_c + \varepsilon_{isc}
\end{aligned} \tag{1}$$

where the outcome is some measure of educational attainment (high school dropout, GED, or high school graduate), I is our measure of inequality or mobility as appropriate, LS and MS are indicators of low and middle SES, respectively. The subscripts i , s , and c index individuals, states, and birth cohorts, respectively; γ_s and γ_c represent state and cohort fixed effects. Cohort variation comes from the different datasets. The vector X consists of additional personal demographic characteristics – gender, race/ethnicity, and an indicator for living with a single parent at age 14. The vector E captures environmental factors including relevant public policies and labor market conditions in the state and year in which the respondent was age 16.⁸ It is important to note that inequality is a long-run average (not subscripted by c), so we are estimating the impact of persistent differences in inequality, not transitory differences. We have specified this model focusing on state-level variation, but we also consider variation at the MSA level as well.

Ideally we would examine the effect of income inequality and economic mobility as two distinct concepts. In a place with high levels of income inequality and high rates of mobility, the effects of income inequality might be less consequential. Alternatively, in places with high levels of income inequality and low rates of upward mobility, any negative effects of income inequality

⁸These variables include the state unemployment rate at age 16, the state minimum wage, state education policies (compulsory schooling age and indicators for high school exit exam requirements), state welfare policies (family cap and maximum AFDC/TANF benefit for a family of 3), state abortion policies (Medicaid funding, parental notification/consent, and mandatory delay laws), and an indicator variable for SCHIP implementation and Medicaid family planning waiver implementation. Information on exit exam requirements by state and year is taken from Dee and Jacob (2007) and Center on Education Policy (2010). Information on compulsory school laws by state and year are obtained from the National Center for Education Statistics, *Digest of Education Statistics* (various years). Detailed source information and notes about the construction of the other variables in this list are provided in Kearney and Levine (2012). We have also experimented with interacting all of the policy variables with SES indicators and found that the results were unaltered by doing so.

might be particularly strong. In practice, our state-level measure of income inequality is strongly negatively correlated with economic mobility, making it empirically difficult to identify separate effects. Thus, we consider inequality and mobility as highly correlated empirical constructs that are essentially interchangeable in our regression analyses.

The main shortcoming with this empirical strategy is that any omitted, state-specific factor that is fixed over time and correlated with long-term measures of income inequality may generate biased results if it has disproportionate impacts on the educational attainment of low SES youth. We have no definitive approach to resolve this problem, but we do implement a method designed to determine whether potentially likely alternatives are playing this role. To do so, we estimate regression models of the form:

$$\begin{aligned} Outcome_{is} = & \beta_0 + \beta_1 (I_s \cdot LS_{is}) + \beta_2 (I_s \cdot MS_{is}) + \beta_3 (A_s \cdot LS_{is}) + \beta_4 (A_s \cdot MS_{is}) \\ & + \beta_5 LS_{is} + \beta_6 MS_{is} + \beta_7 X_{is} + \beta_8 E_s + \gamma_s + \gamma_c + \varepsilon_{is} \end{aligned} \quad (2)$$

In essence, our approach involves including potential alternative state factors (A_s) that could plausibly affect the relative educational attainment of low SES youth and examining whether the results change when we include them in the same manner in which we have included the inequality/SES interactions. If the coefficients on the interaction terms of primary interest change when we add the additional interactions between SES and these alternatives, then it would suggest the results generated from equation (1) are biased. It is impossible to rule out this form of bias unless we try including every possible alternative, but if what we believe to be important alternatives have no impact, then we can be more confident in a causal interpretation of our findings.

When we implement this approach, we consider four categories of these other state factors that are designed to examine four alternative sets of hypotheses. The first set of factors addresses the measurement of income inequality. As noted earlier, we use the 50/10 ratio as our

primary measure of inequality. In our past work on early, non-marital childbearing, we found that the 50/10 ratio was the most empirically relevant measure to determining rates of early, non-marital childbearing. But, we recognize that there are reasons why upper tail inequality might be particularly important for educational outcomes; for example, upper-tail inequality might lead to increases (or decreases) in the level of public school funding available in low-income neighborhoods. We empirically explore the impact of including the 90/50 ratio, as well as the 10th and 50th percentiles of the income distribution.⁹

The second set of alternative factors we include are measures of the returns to education. This is important because it enables us to identify the incentive effect of higher returns (as in a standard Becker model) separately from any offsetting discouragement effect of the type we propose. Third, we consider a set of alternatives that could be considered mediating factors to determine the mechanism by which increased inequality alters educational attainment. If we include these mechanisms in the model as we express in Equation (2), we should see a change in the estimated impact of the 50/10 ratio. Finally, we include a set of potential confounding factors that would be more typically addressed when thinking about problems of omitted variable bias.

We also seek to determine the extent to which differences in distributions of underlying ability would alter the interpretation of our findings. We discuss this possibility and our approach to addressing it below.

IV. DATA

A. Education Microdata

⁹We also examined 50/10 ratios constructed just among high school graduates to rule out the possibility that changes in educational attainment associated with inequality are altering 50/10 ratios. This analysis yielded similar findings to those reported below.

To estimate these models, we use five sources of individual-level data. Three of the sources are available from the National Center for Education Statistics (the National Educational Longitudinal Survey – NELS, High School and Beyond – HSB, and the Educational Longitudinal Survey – ELS) and the other two are the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth (NLSY79 and NLSY97).¹⁰ Each of these datasets has the distinct advantage of including detailed measures of educational attainment, including the ability to separately identify those who receive a degree through passing a general educational development (GED) test and those who receive a traditional high school degree. Their combination also generates a sample of tens of thousands of teens who are moving through (or just recently completed) their high school years. The NLSY79 originally surveyed 12,686 respondents born between 1957 and 1964 (age 14-22 in 1979). HSB originally surveyed over 30,000 high school sophomores in 1980, of whom around half were invited to participate and 13,682 of whom did so in the second follow up four years later.¹¹ We measure high school completion in that year. NELS surveyed 14,915 8th graders in 1988 who were also surveyed in 1994, when we can determine whether they completed high school. NLSY97 surveyed 8,984 respondents born between 1980 and 1984 (age 12-18 in 1997). ELS surveyed 15,300 10th graders in the spring of 2002 who were also surveyed in 2006, when high school completion can be measured. In combination, a maximum of 65,567 respondents are available. In reality, mainly because of missing state identifiers, missing information regarding SES (defined below as level

¹⁰For all datasets other than High School and Beyond, geographic identifiers are only available for those with restricted use data agreements. This means that we are not able to share our data with other researchers, although we are happy to provide our programs so that those who are able to obtain their own agreement can follow our steps. Formal state identifiers are not available at all for High School and Beyond, but researchers have identified ways to provide educated guesses of state of residence for survey respondents (cf. Grogger, 1996). We are grateful to Jeff Grogger for providing us with his data indicating state identifiers for these data.

¹¹This survey also included over 28,000 seniors in 1980, but we do not use them because many high school dropouts never make it to be seniors in high school; using these data would introduce substantial selection bias.

of maternal education), and sample attrition we have available 53,150 teens for our analysis.¹² Limited time variability is available when we combine these datasets, but our analysis relies on long-term geographic variability, as we described earlier.

A critical feature of these data, as captured in our econometric models, is a measure of the youth's socioeconomic status. The measure that is available in each of these datasets is mother's level of education. We distinguish students according to whether their mother dropped out of high school, graduated from high school, or attended college (regardless of their graduation status). Although maternal education does not perfectly predict economic status, we take advantage of the fact that it is strongly correlated with SES.

Although the availability of all five of these datasets provides a unique opportunity to generate a large sample of high school students and follow them through the completion (or not) of their degree, their combination also presents challenges. In particular, identifying a consistently selected sample and outcome measure is somewhat complicated. Sample selection is an issue because individuals entered the samples at different ages and grades. For instance, the NELS initially surveyed 8th graders and the ELS and HSB initially surveyed 10th graders. Survival to 10th grade represents a degree of success that changes the composition of the sample since more poorly performing students may drop out before they make it to 10th grade. We discuss issues like these in the attached data appendix; we account for this in our econometric specification by including data set dummy variables, which we have labeled in the model as cohort fixed effects since datasets identify cohorts. We focus on three consistent measures of educational attainment across datasets. In each of these datasets, we are able to determine (a) whether a student completed high school and received a traditional diploma, (b) whether the

¹² Sample attrition reduces the sample size to 61,067. Missing educational attainment reduces it further to 59,286. Missing maternal education brings the final sample size down to 53,150.

student received a GED, or (c) whether the student never obtained a high school degree via either route.

B. Data on Income Inequality

We calculate our measures of income inequality – the 50/10 ratio and the 90/50 ratio – by state and survey year using microdata from the 1980, 1990, and 2000 Censuses, which capture details of the income distribution over a comparable period as our microlevel datasets.¹³ These data are available from IPUMS-USA (Ruggles et al., 2010). We then take the long-term average over all years for a state after adjusting for inflation. As we described earlier, we take this approach because we are trying to capture something about the permanent or semi-permanent economic and cultural landscape in the place where an adolescent lives, as opposed to short-term fluctuations.

C. Data on Economic Mobility

To measure intergenerational mobility, we rely on the important data contributions of Chetty, et al. (2014a), which reports mobility statistics using data for the 1980 through 1982 birth cohorts obtained from the Internal Revenue Service. Income data for the parents of this cohort were measured in 1996-2000 when the children were teenagers. The children's own income data were measured in 2011-2012 when the children were around 30 years old.¹⁴

Using these data, Chetty, et al. (2014a) construct two measures of mobility. *Absolute mobility* is defined as the average percentile in the income distribution of a child who is born to

¹³Total household income in the census is defined as the sum of eight categories: Wages, salary, commissions, bonuses, or tips from all jobs; Self-employment net income; interest, dividends, net rental income, royalty income, or income from estates and trusts; Social Security or Railroad Retirement; Supplemental Security Income (SSI); any public assistance or welfare payments from the state or local welfare office; retirement, survivor, or disability pensions other than social security; any other sources of income received regularly such as Veterans' (VA) payments, unemployment compensation, child support, or alimony. It is a pre-tax, post-transfer measure of income.

¹⁴The data itself can be obtained at <http://www.equality-of-opportunity.org/>. These data are presented at the level of Census commuting zones; we use the provided population estimates to aggregate the data to the state level.

parents at the 25th percentile in the income distribution; larger values indicate greater mobility. *Relative mobility* is defined as the change in the child's percentile rank for a one percentile increase in the parents' rank; larger values indicate less mobility or greater intergenerational income persistence. The correlation coefficient between these measures and the 50/10 ratio of household income across states is 0.49 and 0.59, respectively. The correlation between the two mobility measures is 0.68. We have explored using both measures in our analysis and they generate similar qualitative results. But, the results based on relative mobility are easier to interpret relative to inequality, and so we focus on those results in the text. We seek to examine whether greater income inequality and less economic mobility – as captured by higher values for these measures – lead to higher dropout rates for low SES youth.

V. EMPIRICAL RESULTS

A. Descriptive Analysis

To highlight the identification strategy that we use, we initially present the results of a descriptive analysis of educational outcomes for teens by their socioeconomic status and the level of income inequality and economic mobility that exists in their state. Figures 4 and 5 present the results of this descriptive analysis for dropping out of high school for inequality and mobility, respectively. Foreshadowing the results from our subsequent formal econometric analysis, we present these results just for boys. In Figure 4, we classify states into those in the top, bottom and middle two quartiles of inequality as measured by the 50/10 ratio.¹⁵ In Figure 5

¹⁵States fall into the following categories (with the 50/10 ratio in parentheses). Low inequality: UT (3.40), NV (3.49), VT (3.54), ID (3.59), NH (3.61), NE (3.71), IA (3.72), WI (3.72), AK (3.75), OR(3.77), WY (3.78), ME (3.80), IN (3.80). Middle inequality: CO (3.81), AZ (3.81), ND (3.82), HI (3.82), SD (3.84), FL (3.85), MT (3.86), DE (3.87), KS (3.88), MN (3.90), WA (3.92), MD (3.98), VA (4.03), PA (4.03), CT (4.06), MO (4.07), OH (4.08), CA (4.15), OK (4.19), NC (4.19), NM (4.21), NJ (4.22), MI (4.22), WV (4.25), AR (4.28). High inequality: IL (4.29), RI (4.38), TX (4.40), TN (4.44), SC (4.45), MA (4.52), KY (4.54), MS (4.59), GA (4.66), NY (4.77), AL (4.85), LA (5.03), DC (5.66).

we take a similar approach categorizing mobility in states by the level of intergenerational income persistence (“relative mobility”).¹⁶ In both figures, the bars represent the percentage of boys who dropped out of high school. Boys are separated into categories according to their mother’s educational attainment to proxy for the SES, along with the level of inequality/mobility that exists in their state.

Each of these figures groups SES categories so that the pattern in educational outcomes by inequality status within SES category is readily apparent. For instance, in Figure 4, we see that around 5 percent of boys from higher SES families drop out of high school regardless of the level of income inequality in their state. No obvious pattern is evident among the middle SES boys in different inequality categories either. Among low SES boys, however, higher inequality is associated with higher rates of dropping out of high school. The magnitude of the difference is sizeable. Low SES boys in high inequality states are almost 6 percentage points more likely to drop out of high school than low SES boys in low inequality states. Figure 5 tells a similar story. Boys from low SES households in low mobility states are more likely to drop out of high school relative to those in states with more mobility. This pattern is not observed for boys from higher SES households.

B. State-Level Analysis

These findings from our descriptive analysis are affirmed when we estimate the regression models described in equation 1. In essence, these regressions are analogous to the data reported in Figures 4 and 5 with the exception that the 50/10 ratio and intergenerational income

¹⁶ States fall into the following categories (with the level of relative mobility listed in parentheses). : Low mobility: MS (0.419), MD (0.408), LA (0.399), OH (0.397), AL (0.391), DE (0.391), NC (0.387), SC (0.397), IL (0.384), VA (0.382), IN (0.373), TN (0.373), MO (0.372). Middle mobility: GA (0.371), PA (0.366), AR (0.365), KY (0.363), CT (0.359), MI (0.356), NJ (0.346), WI (0.346), WV (0.343), OK (0.339), NY (0.338), RI (0.333), DC (0.330), MA (0.327), KS (0.323), TX (0.322), NE (0.320), MN (0.316), FL (0.313), IA (0.307), NM (0.301), ME (0.301), NH (0.293), AZ (0.291), VT (0.289). High Mobility: SD (0.286), CO (0.282), OR (0.281), WA (0.280), AK (0.271), ND (0.264), NV (0.263), WY (0.259), MT (0.251), ID (0.249), UT (0.243), CA (0.242), HI (0.236).

persistence are treated continuously rather than in categories and additional explanatory variables are included. As such, it is perhaps not surprising that the estimation results mimic those obtained in the graphical analysis we just presented.

Table 1 presents those results for all students in the sample and then separately for boys and girls.¹⁷ Columns 1 through 3 are identical except they focus on our three different measures of educational outcomes (high school dropout – Column 1; GED – Column 2; and high school graduation – Column 3). The percentage of students in each category is displayed just above the regression results to aid in interpretation. When we focus on dropping out of high school for all students (the top panel of the table), we see that a one point increase in the 50/10 ratio increases the likelihood of dropping out by 2.3 percentage points for students from low SES families. This estimate is not quite statistically significant, with a p-value of 12.3 percent. When we explore differences in estimates by gender, however, we see that boys, in particular, are more likely to drop out of high school when they grow up in a low SES household in an area marked by high inequality. Moving from a relatively low inequality to high inequality state represents perhaps a one point increase in the 50/10 ratio. This means that a making such a move for a boy from a low SES family increases the likelihood of dropping out of high school by age 20 by 4.1 percent. The analogous estimate for girls is considerably smaller, statistically insignificant, and marginally significantly different than the estimate for boys (p-value = 8.6 percent). Estimates for the other two outcomes, receiving a GED or graduating from high school, are too imprecise to determine whether the increase in dropping out for boys came mainly from either of them.

We observe a similar gender disparity in the results when we examine the impact of changes in economic mobility rather than inequality. Table 2 provides these results, where

¹⁷ We have also estimated these models separately by race and ethnicity, but the data were not sufficiently powerful to yield statistically significant differences across groups.

mobility is measured as “relative mobility,” which is an indicator of intergenerational income persistence. These results indicate that low SES teens overall are more likely to drop out of high school when rates of income persistence are higher in their state. When we separate teens by gender, though, we see that this effect is, again, largely driven by the impact on boys. Moving from a high persistence state to a low persistence state changes the relative mobility measure by about 0.15 (from around 0.25 to around 0.4). Multiplying that change by the low SES*persistence coefficient of 0.405 for boys indicates that moving from a high to low income persistence state increases the likelihood of dropping out of high school for a low SES youth by 6 percent, which is somewhat larger than the analogous estimate from our inequality analysis.

Again, for low SES girls we do not observe a statistically significant impact on their likelihood of dropping out of high school; the p-value of a test of equality between the coefficient for boys and girls is 5.7 percent. Although imprecision makes it difficult to draw strong conclusions here, the results in columns (2) and (3) suggest that much of the increase in high school dropout rates among low SES boys in low mobility locations substitutes for high school graduation, not GED completion.

C. MSA-Level Analysis

It is also instructive to consider the appropriate level of geography. The way we are thinking about the possible effects of income inequality and mobility implies that the appropriate unit is a fairly broad area, such as a state or an MSA. These would allow for the effects of any type of residential or institutional segregation that might occur as a result of widened income inequality and affect perceptions of success. Alternatively, a model of relative deprivation (as described below), might imply that the relevant level of income inequality for individual attitudes is more local.

Table 3 focuses on the outcome of dropping out high school and repeats the analysis of the impact of inequality and mobility by MSA rather than state. The models reported here are analogous to those in Table 1 except that these regressions exclude policy variables set at the state level. Omitting those variables from the state-level models has virtually no impact on the results. We are also forced to omit the NELS and HSB data from our analysis because we are not able to identify geography below the state-level in the base year in those data sets. MSA-level results are similar to state level results. Lower SES teens, and particularly boys, who grow up in MSAs with greater lower-tail income inequality and higher rates of income persistence are considerably more likely to drop out of high school. The p-value on the gender difference in impacts on dropping out of high school is 0.0004 when considering inequality and 0.17 when considering income persistence.

The general pattern in the data showing that boys are more likely to be affected by inequality and mobility than girls leads us to focus the remainder of the analysis on boys. Also, because the conclusions based on an analysis of inequality and mobility are so similar, for the purposes of brevity we have chosen to report subsequent results just for inequality. Finally, as in Table 3, we focus the remainder of our reported results solely on the outcome of dropping out of high school.

VI. AN EXAMINATION OF POTENTIAL EXPLANATIONS

In the next set of tables, we estimate models of the form of Equation 2 that are designed to examine the extent to which other state-specific factors may matter and alter our interpretation of a causal impact of income inequality. In each of these tables, we also include the results of our base specification from Table 1 in the first column to facilitate comparison.

A. Alternative Measures of the Income Distribution

Table 4 reports the results of estimating the main equation of interest using different measures of the income distribution. The alternatives we consider are the 90/50 ratio, and, separately, the 10th and 50th percentiles of the income distribution. Each of the alternative measures of the income distribution capture different attributes. The 90/50 ratio represents income inequality at the top of the income distribution. This is the part of the distribution that has grown over time. We have argued that the 50/10 ratio is a better measure of inequality for the low SES population because it may more realistically indicate what is available to them if they were able to move up the ladder, but this is an empirical question. We also include the 10th and 50th percentiles of the income distribution separately to understand whether our findings based on their ratio are actually attributable to one of the two components separately.

As described earlier, we include the interaction of the 50/10 ratio and SES status along with interactions between SES status and these other measures. We can directly interpret the coefficients on those interactions and we can also observe whether substantive changes occur in the coefficients on the 50/10*SES status interaction. The estimates reported in Table 4 provide no substantive reason for changing our earlier conclusions that the 50/10 ratio is the appropriate measure of income inequality to consider. If anything, including the 90/50 ratio strengthens the relationship between the 50/10 ratio among low SES boys and dropping out of high school. Interactions with the other measures are generally statistically insignificant and have no impact on the interaction between the 50/10 ratio and low SES.

B. Wage Inequality and the Returns to Education

Recall from our earlier discussion that if greater inequality reflects a greater return to investment in human capital, the Beckerian framework predicts that all else equal, students should invest more when income inequality is greater. Solon (2004) formalizes this concept in a

model where parents make human capital investments in their children; building on the theoretical foundation of Becker and Tomes (1979), he shows that parental investment in a child's human capital increases when the payoff to that return is higher, that is, when there is more wage inequality. In our framework this would entail a reduction in the likelihood of dropping out of high school. The specifications reported in Table 5 address this possibility directly by considering a distinct offsetting role from the incentive effect of wage differentials. In Column 2 we estimate a regression model that includes separate interaction terms for low SES and (a) lower-tail inequality and (b) the wage premium for high school graduates relative to high school dropouts. The high school graduate wage premium is calculated from the same Census data that we used to estimate measures of inequality except that the sample is restricted to those between ages 21 and 64.

The results of this specification indicate that, even with this additional interaction term in the model, the point estimate on the interaction term between low-SES status and lower tail inequality is virtually unchanged from the initial specification. The data indicate a positive effect of income inequality on the likelihood that a disadvantaged youth drops out of school, conditional on the high school wage premium. The high school graduate to dropout wage premium itself is estimated to reduce the likelihood of dropping out for low SES boys, although it is insufficiently precise to be statistically significant.

C. Potential Mediating Factors

Next we attempt to investigate whether we can empirically identify mediating factors that might play a role in altering educational outcomes in the presence of greater inequality. Two candidate factors are neighborhood structure and quality of educational institutions, as described in Section II above. These explanations would be consistent with the "ecological model"

described by Duncan and Murnane (2011), which posits that higher levels of income inequality may lead to changes in family structure, neighborhoods, and other labor market features that may hinder a child's educational development.

One possible channel through which income inequality can affect social outcomes is through increased levels of residential and institutional segregation. For the poor, greater residential segregation can affect social and labor market networks, the presence of high achieving role models, and the establishment of peer groups and norms. We explore these possibilities by considering the following alternative state characteristics – an index of racial segregation and an index of income segregation.¹⁸ To the extent that any of these factors, when interacted with SES, have a statistically significant effect and/or alter the estimated impact of the SES*50/10 ratio interactions, one could conclude that they are an important mediating factors. The results reported in Table 6 provide no evidence of this sort of effect. None of the coefficients on the interactions with these factors in columns (2) and (3) are statistically significant, and their inclusion has a negligible impact on the SES*inequality interactions.

Another possible channel through which income inequality could affect educational outcomes is through its effect on the public financing of public goods. As described earlier, the political economy theory is ambiguous, but the most recent empirical evidence finds support for the prediction of the median voter theorem that revenue for public school spending increases in the level of local income inequality (Boustan et al., 2013; Corcoran and Evans, 2010; and

¹⁸ The racial segregation index was obtained from the University of Michigan Population Studies Center website: <http://www.psc.isr.umich.edu/dis/census/segregation.html>, accessed on April 25, 2012. The source is William H. Frey, Brookings Institution and University of Michigan Social Science Data Analysis Network's analysis of 2005-9 American Community Survey and 2000 Census Decennial Census tract data. The index is constructed as a Dissimilarity Index that measures the degree to which the minority group is distributed differently than whites across census tracts. They range from 0 (complete integration) to 100 (complete segregation) where the value indicates the percentage of the minority group that needs to move to be distributed exactly like whites. The income segregation index was obtained from Chetty, et al. (2014a,b). We obtained those data at the commuting zone level and calculated the population-weighted, state-level values for our analysis.

Gordon, 2013). If greater levels of income inequality increase public school expenditures, this likely is not the empirical explanation for the link we observe. Nonetheless, we explore this possibility by considering per capita educational expenditures and pupil-teacher ratios.¹⁹

In our data, we see that per capita educational expenditures and pupil teacher ratios are only weakly correlated with state-level lower-tail income inequality (0.14 and -0.23, respectively), making it unlikely that these are omitted variables driving the observed link between income inequality and drop-out behavior. The regression results confirm that this is not the case. As reported in Table 6, columns (4) and (5), the data do not indicate a direct effect of these measures on the rate at which low-SES individuals drop out of high school. Nor does the inclusion of these measures alter the conclusion that greater lower-tail income inequality leads to higher rates of high school drop-out behavior among low SES individuals. The lack of empirical support for these mechanisms suggests that an alternative perspective may be warranted. We offer such a perspective below.

D. Remaining Potential Confounding Factors

In the last set of “horse race” specifications, Table 7 presents the results of including one additional set of interactions with other state-specific factors that could simply represent confounding factors. These include the percentage of the state’s population that is minority, the poverty rate in the state, and the state’s incarceration rate.²⁰ The goal here is to determine whether one of these state-specific factors is a contextual factor related to state-level income inequality and driving the differential high school dropout rates. The results reported in Table 7

¹⁹We thank Liz Cascio for generously sharing the historical data she compiled on per pupil expenditures and per pupil teacher ratios.

²⁰The incarceration data are compiled by the U.S. Bureau of Justice Statistics, Office of Justice programs, downloaded from www.ojp.usdoj.gov. Poverty rate data comes from the United States Census Bureau and were downloaded from <http://www.census.gov/hhes/www/poverty/data/historical/people.html>.

do not indicate that to be the case. Interactions between each of these factors and socioeconomic status are universally insignificant and their inclusion in the regression model has no substantive impact on the estimated effect of the interactions between lower-tail inequality and individual SES.

E. The Role of Underlying Differences in Ability

As described earlier, a potential alternative explanation for the link between high inequality and low mobility is that in locations with greater demographic diversity, there will be a mechanical correlation that links the two. The more similar the underlying populations, the lower the inequality (by definition) and the greater the mobility because chance will play a greater role in determining who succeeds in any given period. In essence, this is an argument about the underlying distribution of “ability.”²¹

We explore this alternative within the context of educational outcomes using test scores as a proxy for underlying ability. Specifically, we use data from scores on the Armed Forces Qualifying Test (AFQT), which was administered to participants in the NLSY79 and NLSY97 surveys. The AFQT is used by the military to determine eligibility and placement. The score is reported as a standardized percentile ranking. These data have been used by empirical researchers in the past for similar purposes (e.g., Herrnstein and Murray, 1994; Neal and Johnson, 1996; Belley and Lochner, 2007). We hasten to note that the AFQT is not a direct measure of innate ability; on this point, Cascio and Lewis (2006) show that exogenous increases in educational attainment lead to increases in AFQT scores, especially for minorities. It is most

²¹Mankiw makes this point clearly in a 2013 post on his widely-read blog by offering as an example the skill of chess players. If we have one group of chess players who are all of roughly comparable ability, then who wins and loses the matches will be closer to a random draw and mobility through the rankings will be high. If another group of chess players has some with greater ability and others who are weaker, then inequality in wins/losses will be higher and mobility lower. (“Observations on the Great Gatsby Curve,” <http://gregmankiw.blogspot.nl/2013/07/some-observations-on-great-gatsby-curve.html>, accessed April 20, 2014.)

appropriately considered a cumulative measure of ability, reflecting innate endowments, environmental influences, and the result of formal and informal human capital investment. Still, these test scores provide information about cognitive ability at the time the exam was taken.

The purpose of the empirical analysis reported in Table 8 is to determine whether these differences in the AFQT measure of cognitive ability can explain any share of the greater relative rate of drop out behavior among low-SES boys in high inequality places. As in past tables, the first column is included for the purpose of comparison; it reports the results from a model analogous to our main specification taken from Table 1 for boys, with the estimated point estimate on the interaction of primary interest being 0.042 (standard error 0.016).²² Since the AFQT is only available in NLSY79 and NLSY97, the second column presents the same regression just for those two datasets. The results indicate a somewhat larger point estimate (0.067), for the effect of inequality on dropping out, but the smaller sample size leads to greater imprecision as well (standard error 0.029). The third column of this table examines what happens if we control for AFQT as an explanatory variable in a specification that is otherwise identical to that in Column 2. We find that doing so does reduce the point estimate by about one-third, to 0.045 from 0.067. This is not statistically different from the estimated effect in column one, but the standard error is now 0.028 (owing to the smaller sample size coming from having to restrict the analysis to just two datasets), and so this estimate is no longer statistically significant from zero.

In Column 4, we treat AFQT as the dependent variable and estimate a model that is otherwise equivalent to those estimated earlier. The point estimates indicate that low-SES youth

²²The only minor difference between this specification and that in Table 1 is that we omit all policy variables since we will subsequently be restricting the sample to just two datasets leaving us with very limited variation across states over time. As the results indicate, dropping those variables has virtually no impact on the findings.

in high inequality areas have lower AFQT scores; that relationship is marginally statistically significant (p-value = 8.3 percent). This result helps explain why the estimated impact of inequality for low SES boys fell when we added AFQT: it appears that low SES boys who live in high inequality locations have AFQT scores that are even lower than those for low SES boys overall.²³

There are two possible interpretations of these results. For readers inclined to interpret the AFQT as measuring innate ability, one could conclude that the exclusion of the AFQT variable in previous analyses leads to an upwardly biased estimate of the causal impact of inequality on dropout rates; still, two-thirds of the effect remains. An alternative interpretation is that part of the effect of income inequality is captured by decreased educational investment before the actual drop out event. This corresponds to a leading view of drop out behavior as a “process” rather than a discrete event: a student begins to demonstrate irregular attendance, then multiple failed courses, and eventually the obstacles to graduation feel overwhelming and the student drops out (Rumberger, 2011). In other words, discouraged students stop applying themselves early. This could show up as a lower AFQT score, consistent with the finding of Cascio and Lewis (2006) that an exogenous increase in education leads to higher AFQT scores. Their finding would imply that decreased effort in school and in learning more broadly would result in a lower AFQT score.

Regardless of interpretation, the impact of greater inequality on dropout behavior is substantial, albeit somewhat smaller if one accepts the interpretation that the AFQT measures innate ability. The question that remains is why does this occur and through what channels?

²³ Multiplying the point estimate of -4.38 from the low-SES interaction term in column (4) with the point estimate of -0.005 on the AFQT variable from column (3) yields 0.022, suggesting that the lower AFQT scores of boys in high inequality states would lead to a 0.022 percentage point relative increase in drop-out rates, which is exactly the difference we see between columns (2) and (3). This is another way to see that differences in AFQT capture about one-third of the estimated effect of inequality on the drop-out rates of low-SES boys.

VII. INTERPRETATION AND ALTERNATIVE EXPLANATIONS

The empirical results presented in Table 6 above failed to provide empirical support for a number of potential explanations for the empirical relationship between income inequality/income mobility and high school dropout rates – namely income segregation, racial segregation, or quality of educational institutions as measured by per pupil expenditures and teacher/pupil ratios. We offer an alternative explanation – perhaps income inequality has a direct impact on educational attainment through perceptions of self, relative position in society, or potential economic success, which are not transmitted through the readily measured contextual channels.

We describe three such theories here. First, an influential theory in social science posits a role for *relative deprivation* – as distinct from absolute deprivation – in leading to acts of social unrest. Luttmer (2005) conducts an empirical investigation of this idea and documents that people are less happy when they live around other people who are richer than themselves. Along these lines, the relative position of individuals could lead to feelings of alienation from society that lead them to want to engage in rebellious types of behaviors, perhaps including dropping out of school.

A somewhat related alternative theory is that one's location in the income distribution matters in shaping one's identity construct, which affects one's decisions. Watson and McLanahan (2011) present evidence that relative income matters for the marriage decision of low-income men. They interpret their model within the framework of an identity construct, based largely on the identity model of Akerlof and Kranton (2000). Specifically, Watson and McLanahan hypothesize that individuals perceive a threshold income required for marriage, and that this threshold is influenced by an individual's local reference group. One could imagine an

extension of this theory that applies to educational attainment. Perhaps individuals perceive a threshold type of person who completes higher levels of education; youth at the bottom of the income distribution in more unequal places may be more likely to view themselves as the low achievers in their reference group.

An alternative explanation for the drop out behavior induced by higher rates of income inequality is captured in the “economic despair” model of Kearney and Levine (2014). That paper focused on the decision to become a young, unmarried mother, but the insight of the model applies to the decision to become a high school drop-out. The idea is as follows: greater income inequality might negatively affect the perceived returns to human capital investment from the perspective of an economically disadvantaged adolescent. The concept of intergenerational income persistence – or lack of economic mobility – would fit within this model in the same way as income inequality would (this is less true of the previous two models discussed). For ease of exposition, we focus on income inequality when describing the main ideas of this model.

The notion we have in mind is that a greater gap between the bottom and the middle of the income distribution might lead to a heightened sense of economic marginalization such that an adolescent at the bottom of the income distribution does not see much value in investing in his/her human capital. This framework offers an explanation within the standard human capital framework of decision-making of why greater inequality – which might reflect in part a greater return to human capital investment – does not necessarily lead to greater rates of educational attainment for certain segments of the population.

To clarify these ideas, we summarize here the stylized framework presented in Kearney and Levine (2014) for the context of the high school drop-out decision. An individual chooses to drop out of school in the current period if expected lifetime utility along that path exceeds the

expected lifetime utility associated with staying enrolled. If one does not experience an immediate utility boost from dropping out, then it is never optimal to drop out. But if the current period utility from dropping out exceeds the current period utility from continued enrollment – i.e., if $u^d > u^e$ – which would be the case if the student experiences substantial utility costs from remaining in school (e.g. psychic costs) – then that current period utility boost needs to be compared to the potential option value lost.

Dropping out of school negatively affects expected future utility by leading to lower levels of consumption in the future, which for simplicity, we characterize as taking on high and low values. We assume that dropping out of school is deterministic, and leads to a lifetime of lower income, with a present discounted value of V^{low} . If the adolescent remains enrolled, there is some positive probability p that s/he will achieve the “high” utility position in future periods, with a present discounted value of V^{high} .

The change in lifetime utility from remaining enrolled in school comes from two opposite-signed sources: (1) the loss of current period enjoyment of being out of school and (2) a positive probability p of achieving the high- utility state in the future. Of course, the student does not perfectly observe p , as in Manski (1993).²⁴ Instead, the student bases the decision on his perception of p , in particular, on his perception of his individual-specific p . Let us call this subjective probability of one’s individual likelihood of success conditional on investment q . The condition for deciding not to drop out can be expressed as follows:

²⁴Jensen (2010) highlights this point and suggests that the perception of returns might be particularly inaccurate in developing country settings, which could potentially explain (in part) why rates of educational attainment remain low despite high measured returns. He conducts an experiment among 8th grade boys in the Dominican Republic whereby students at randomly selected schools are given information about the measured returns to completing school. He finds that male students at schools who receive this information complete an average of 0.20-0.35 additional years of school over the subsequent four years as compared to male students in the comparison schools.

$$qV^{high} + (1-q)V^{low} > V^{low} + (u_o^d - u_o^e) \quad (3)$$

If an adolescent perceives that s/he has a sizable chance at achieving economic success -- and thereby capturing V^{high} -- by investing in education, the adolescent is more likely to stay enrolled. On the other hand, if the student perceives that even if s/he stays enrolled, his/her person-specific chances of economic success are sufficiently unlikely -- in other words, if q is very low -- then the comparison is more likely to favor dropping out in the current period.

We speculate that for an adolescent at the bottom of the income distribution, a greater gap between one's position and the middle of the distribution might reduce one's subjective q . If the middle class is sufficiently far from one's own experience, then the student's perceived chances of getting there -- even if he/she does stay in school -- may be sufficiently low.²⁵

Along these lines, the Kearney and Levine (2014) model of "economic despair" might be interpreted as predicting a perpetuation of low income, since diminished expectations are formed in response to high inequality/low mobility, leading low income individuals to higher rates of drop out behavior, which in turn leads to lower lifetime income. In a recent working paper, Genicot and Ray (2014) describe a theoretical model that leads to the same prediction. Their model proposes that society-wide economic outcomes affect individual aspirations. Aspirations that are slightly above one's position lead to increased human capital investment; but if aspirations get too far from one's current position, that could lead to frustration and lower levels of human capital investment.

²⁵In Kearney and Levine (2014) we offered empirical support for the proposition that low SES adolescents growing up in relatively more unequal places actually do have a lower chance of achieving higher income in later life. To test that idea, we examine data from the restricted-use NLSY79 Geocode data. The regression results show that children who grow up in low SES households and who live in a state with high lower-tail inequality are estimated to have permanent incomes that are over 30 percent lower than similar children in low lower-tail inequality states. High and low inequality states are distinguished by a one point increase in the 50/10 ratio. If perceptions of economic opportunity are gauged on actual outcomes, then these findings are consistent with this potential explanation.

VI. DISCUSSION

This paper has presented empirical evidence that economically disadvantaged youth in places characterized by high income inequality and low economic mobility are more likely to drop out of high school, as compared to their counterparts in more equal and more mobile places. In regression models that test the sensitivity of these results to a number of alternative observable mechanisms – such as other features of the income distribution or aggregate poverty rates – we find strong and consistent evidence that it is inequality and mobility and not these other factors driving the observed relationship. These results are also robust to including the high school graduate to drop-out wage premium, despite the finding that the wage premium itself reduces the dropout rate. In an additional set of models that examine potential mediating factors – including residential segregation and school financing – we do not find an impact of these contextual factors and including them does not change our primary finding. We do find that about one-third of the effect is captured by differences in cognitive ability as measured by the AFQT score; we discussed potential interpretations of this finding.

We discuss a set of possible explanations for the main finding that allow for greater income inequality and lower income mobility to directly affect the perceptions of disadvantaged individuals in ways that are distinct from the mechanisms captured in an ecological perspective. Possible models in this set of explanations include those focusing on perceptions of identity (Akerlof and Kranton, 2000; Watson and McLanahan, 2011), relative societal position (e.g., Luttmer, 2005), and one's chance at economic success (Kearney and Levine, 2014; Genicot and Ray, 2014). Of course, it may also simply be that a more contextual transmission mechanism exists, but that we are unable to identify that mechanism with the observed measures we have

used. Though we are ultimately unable to establish a precise mechanism, the empirical finding is striking and further investigation of mechanisms is warranted.

Our analysis has demonstrated that income inequality and lack of income mobility hinders the educational attainment of disadvantaged youth, boys in particular. In previous work, we documented that low-SES girls are more likely to become young, unmarried mothers if they live in a place characterized by high levels of income inequality. These findings have real implications for the potential of disadvantaged youth to achieve economic progress or even sufficiency in the years ahead. We argue that high inequality and low mobility play a critical role in molding the perceptions of low-income youth. The evidence suggests that there may be substantial effects on economic mobility of policies that provide disadvantaged youth with reasons to believe that they have the opportunity to climb the economic ladder and to make those opportunities real.

REFERENCES

- Akerlof, George A. and Rachel E. Kranton (2000). "Economics and Identity." *Quarterly Journal of Economics*, 115(3): 715-753.
- Autor, David H., Lawrence F. Katz, and Melissa S. Kearney (2008). "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90(2), 300-323.
- Bailey, Martha J. and Susan M. Dynarski (2011). "Inequality in Postsecondary Education," in Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, New York: Russell Sage Foundation.
- Becker, Gary S. and Kevin M. Murphy (2007). "The Upside of Income Inequality." *The American*. May/June.
- Becker, Gary S., and Nigel Tomes (1979). "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*. Vol. 87, pp. 1153-89.
- Belley, Philippe, and Lance Lochner (2007). "The Changing Role of Family Income and Ability in Determining Educational Achievement," *Journal of Human Capital*, 1(1): 37-89.
- Boustan, Leah, Fernando Ferreira, Hernan Winkler, and Eric Zolt (2013). "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from US Municipalities and School Districts, 1970-2000." *Review of Economics and Statistics*, 95(4): 1291-1302.
- Cascio, Elizabeth U. and Ethan G. Lewis (2006). "Schooling and the Armed Forces Qualifying Test Evidence from School-Entry Laws." *Journal of Human Resources*, 41(2): 294-318.
- Case, Anne C. and Lawrence F. Katz (1991). "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," NBER Working Paper No. 3705.
- Center on Education Policy (2010). *State High School Tests: Exit Exams and Other Assessments*. Washington, DC: Center on Education Policy.
- Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014a). "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." NBER Working Paper 19843.
- Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014b). "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility" NBER Working Paper 19844.

Corak, Miles (2006). "Do Poor Children Become Poor Adults? Lessons from a Cross-Country Comparison of Generational Earnings Mobility." *Research on Economic Inequality*, 13: 143-188.

Corak, Miles (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27(3): 79-102.

Corcoran, Sean and William N. Evans (2010). "Income Inequality, the Median Voter, and the Support for Public Education." Cambridge, MA: NBER working paper 16097.

Council of Economic Advisers (2012). *Economic Report of the President*. Washington, DC: Government Printing Office.

Dee, Thomas S. and Jacob, Brian A. (2007). "Do High School Exit Exams Influence Educational Attainment or Labor Market Performance?" In Adam Gamoran (ed.), *Standards-Based Reform and the Poverty Gap: Lessons for No Child Left Behind*. Washington, D.C.: Brookings University Press.

Duncan, Greg J., and Richard J. Murnane (2011). "Introduction: The American Dream, Then and Now," In Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York: Russell Sage Foundation.

Genicot, Garance, and Debraj Ray (2014). "Aspirations and Inequality." Cambridge, MA: National Bureau of Economic Research working paper 19976.

Gordon, Nora (2013). "High School Graduation in the Context of Changing Elementary and Secondary Education Policy and Income Inequality: The Last Half Century," National Bureau of Economic Research working paper 19049.

Grogger, Jeffrey T. (1996). "School Expenditures and Post-Schooling Earnings: Evidence from High School and Beyond." *Review of Economics and Statistics* 78(4): 628-637.

Heckman, James J. and Paul A. LaFontaine (2010). "The American High School Graduation Rate: Trends and Levels," *Review of Economics and Statistics* 92 (2): 244-262.

Herrnstein, Richard and Charles Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Simon and Schuster.

Jensen, Robert (2010). "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 125(2): 515-548.

Kearney, Melissa S. and Phillip B. Levine (2012). "Investigating Recent Trends in the U.S. Teen Birth Rate," National Bureau of Economic Research working paper 17964.

Kearney, Melissa and Phillip Levine (2014). "Income Inequality and Early Non-Marital Childbearing," *Journal of Human Resources*, 49(1): pp. 1-31.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007). "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83-119.

Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz (2005). "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics*. 120(1):87-130.

Luttmer, Erzo. F. P. (2005). "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics* 120(3): 963-1002.

Mankiw, N Gregory (2013). "Defending the One Percent." *Journal of Economic Perspectives*, 27(3): 21-34.

Manski, Charles F. (1993). "Adolescent Econometricians: How Do Youth Infer the Returns to Education?" in Charles T. Clotfelter and Michael Rothschild (eds.), *Studies of Supply and Demand in Higher Education*. Chicago: University of Chicago Press.

Mayer, Susan (2001). "How did the Increase in Economic Inequality between 1970 and 1990 Affect Children's Educational Attainment." *American Journal of Sociology*, 107(1):1-32.

Murnane, Richard J. (2013) "U.S. High School Graduation Rates: Patterns and Explanations." *Journal of Economic Literature* 51(2): 370-422.

Neal, Derek and William R. Johnson (1996). "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104(5): 869-895.

Reardon, Sean (2011). "The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations," in Greg J. Duncan and Richard J. Murnane (eds.), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York: Russell Sage Foundation.

Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. (2010) *Integrated Public Use Microdata Series: Version 5.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Rumberger, Russell W. (2011). *Dropping Out : Why Students Drop Out of High School and what can be done about it*. Cambridge, Mass.: Harvard University Press.

Solon, Gary (2004). "A model of intergenerational mobility variation over time and place," in ed. Miles Corak, *Generational Income Mobility in Europe and North America*, Cambridge University Press: 2004: 38-47.

Snyder, Thomas D., and Sally A. Dillow (2012). *Digest of Education Statistics 2011* (NCES 2012-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Watson, Tara (2009). "Inequality and the Measurement of Residential Segregation by Income." *Review of Income and Wealth*, 55: 820-844.

Watson, Tara, and Sara McLanahan (2011). "Marriage Meets the Joneses: Relative Income, Identity, and Marital Status." *Journal of Human Resources*, 46: 568-586.

Wilkinson, Richard and Kate Pickett (2009). *The Spirit Level: Why Greater Equality Makes Societies Stronger*. London: Bloomsbury Press.

Wilson, William J (1987). *The Truly Disadvantaged: The Inner City, the Underclass and Public Policy*. Chicago: University of Chicago Press.

Table 1: Impact of Long-Term Inequality by State
on Educational Attainment by Age 20, by Socioeconomic Status and Gender

| | High School Dropout (1) | GED Receipt (2) | High School Graduate (3) |
|---------------------------------|----------------------------|--------------------|-----------------------------|
| All | | | |
| Percent in Category | 10.1 | 4.8 | 85.1 |
| 50/10 Ratio* Mom HS Dropout | 0.023 (0.015) | -0.006 (0.010) | -0.017 (0.016) |
| 50/10 Ratio* Mom HS Graduate | 0.018 (0.014) | 0.010 (0.008) | -0.028 (0.013) |
| Boys | | | |
| Percent in Category | 11.2 | 5.5 | 83.3 |
| 50/10 Ratio* Mom HS Dropout | 0.041 (0.015) | -0.018 (0.015) | -0.022 (0.018) |
| 50/10 Ratio* Mom HS Graduate | 0.025 (0.017) | 0.013 (0.009) | -0.037 (0.016) |
| Girls | | | |
| Percent in Category | 9.1 | 4.1 | 86.8 |
| 50/10 Ratio* Mom HS Dropout | 0.007 (0.019) | 0.005 (0.010) | -0.012 (0.022) |
| 50/10 Ratio* Mom HS Graduate | 0.009 (0.014) | 0.006 (0.011) | -0.015 (0.016) |

Notes: Reported standard errors (in parenthesis) are adjusted for clustering at the state level. Additional explanatory variables in each regression include maternal educational attainment, gender (where appropriate) race/ethnicity, an indicator variable for living with a single parent at age 14, the state unemployment rate at age 16, the state minimum wage, state education policies (compulsory schooling age and indicators for high school exit exam requirements), state welfare policies (family cap and maximum AFDC/TANF benefit for a family of 3), state abortion policies (Medicaid funding, parental notification/consent, and mandatory delay laws), and indicator variables for SCHIP implementation and a Medicaid family planning waiver program, along with state and cohort fixed effects. The p-value of a test comparing the equality of coefficients in column (1) by gender in response to a change in the 50/10 ratio*mom HS dropout is 0.086. Sample includes data from the NELS, HSB, ELS, NLSY79, and NLSY97. The total sample size is 53,150, with 25,816 boys and 27,334 girls.

Table 2: Impact of Intergenerational Income Persistence by State
on Educational Attainment by Age 20, by Socioeconomic Status

| | High School Dropout (1) | GED Receipt (2) | High School Graduate (3) |
|---|----------------------------|--------------------|-----------------------------|
| All | | | |
| Percent in Category | 10.1 | 4.8 | 85.1 |
| Intergenerational Persistence* Mom HS Dropout | 0.261 (0.095) | 0.028 (0.102) | -0.288 (0.086) |
| Intergenerational Persistence* Mom HS Graduate | 0.008 (0.066) | 0.040 (0.054) | -0.048 (0.067) |
| Boys | | | |
| Percent in Category | 11.2 | 5.5 | 83.3 |
| Intergenerational Persistence* Mom HS Dropout | 0.405 (0.121) | 0.032 (0.136) | -0.437 (0.121) |
| Intergenerational Persistence* Mom HS Graduate | 0.065 (0.094) | 0.067 (0.073) | -0.132 (0.077) |
| Girls | | | |
| Percent in Category | 9.1 | 4.1 | 86.8 |
| Intergenerational Persistence* Mom HS Dropout | 0.146 (0.113) | 0.022 (0.087) | -0.168 (0.104) |
| Intergenerational Persistence* Mom HS Graduate | -0.047 (0.060) | 0.010 (0.075) | 0.037 (0.102) |

Notes: See notes to Table 1. The p-value of a test comparing the equality of coefficients in column (1) by gender in response to a change in intergenerational persistence*mom HS dropout is 0.057.

Table 3: Impact of Long-Term Inequality and Intergenerational Income Persistence by MSA
On Likelihood of Dropping Out of High School, by Socioeconomic Status and Gender

| | All (1) | Boys (2) | Girls (3) |
|--|------------|-------------|--------------|
| Percent in Category | 12.3 | 14.1 | 10.6 |
| Long-Term Inequality (50/10 Ratio) | | | |
| 50/10 Ratio* | 0.036 | 0.073 | 0.002 |
| Mom HS Dropout | (0.013) | (0.018) | (0.016) |
| 50/10 Ratio* | 0.020 | 0.028 | 0.009 |
| Mom HS Graduate | (0.011) | (0.016) | (0.012) |
| Sample size | 22,304 | 11,042 | 11,262 |
| Intergenerational Income Persistence (Relative Mobility) | | | |
| Intergenerational Persistence* | 0.361 | 0.524 | 0.245 |
| Mom HS Dropout | (0.149) | (0.198) | (0.171) |
| Intergenerational Persistence* | 0.018 | 0.064 | -0.024 |
| Mom HS Graduate | (0.092) | (0.134) | (0.107) |
| Sample Size | 22,247 | 11,013 | 11,234 |

Notes: Reported standard errors are adjusted for clustering at the MSA level. Additional explanatory variables in each regression include maternal educational attainment, race/ethnicity, and an indicator variable for living with a single parent at age 14, along with MSA and cohort fixed effects. The p-value of a test comparing the equality of coefficients by gender for HS dropout mothers is 0.0004 for inequality and 0.17 for persistence. Sample includes data from the ELS, NLSY79 and NLSY97.

Table 4: Impact of Alternative Income Distribution Measures on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status

| | 50/10 ratio (1) | 90/50 ratio (2) | 10 th Percentile of Income (in \$10,000s) (3) | 50 th Percentile of Income (in \$10,000s) (4) |
|--|--------------------|--------------------|---|---|
| Correlation between 50/10 ratio and characteristic: | | 0.67 | -0.63 | -0.20 |
| 50/10 Ratio* | 0.041 | 0.058 | 0.041 | 0.041 |
| Mom HS Dropout | (0.015) | (0.025) | (0.024) | (0.016) |
| 50/10 Ratio* | 0.025 | 0.023 | 0.024 | 0.024 |
| Mom HS Graduate | (0.017) | (0.019) | (0.021) | (0.017) |
| State Characteristic* | --- | -0.069 | 0.0002 | -0.0001 |
| Mom HS Dropout | --- | (0.072) | (0.005) | (0.001) |
| State Characteristic* | --- | 0.004 | -0.0001 | -0.0003 |
| Mom HS Graduate | --- | (0.050) | (0.003) | (0.001) |

Notes: see notes to Table 1. Interacted state characteristic is listed in column headers.

Table 5: Impact of Educational Wage Premiums on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status

| | 50/10 ratio (1) | HS Grad to HS Dropout Wage Premium (2) | College Grad to HS Grad Wage Premium (3) |
|--|--------------------|---|---|
| Correlation between 50/10 ratio and characteristic: | | 0.27 | 0.35 |
| 50/10 Ratio* Mom HS Dropout | 0.041 (0.015) | 0.046 (0.015) | 0.037 (0.017) |
| 50/10 Ratio* Mom HS Graduate | 0.025 (0.017) | 0.023 (0.018) | 0.022 (0.019) |
| State Characteristic* Mom HS Dropout | --- --- | -0.117 (0.076) | 0.039 (0.043) |
| State Characteristic* Mom HS Graduate | --- --- | 0.029 (0.062) | 0.024 (0.043) |

Notes: see notes to Table 1. Interacted state characteristic is listed in column headers.

Table 6: Impact of Potential Mediating Factors on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status

| | 50/10 ratio (1) | Racial Segregation Index (2) | Income Segregation Index (3) | Per Capita Educational Expenditures (x 1,000) (4) | Pupil Teacher Ratio (x10) (5) |
|--|--------------------|---------------------------------------|---------------------------------------|---|-------------------------------------|
| Correlation between 50/10 ratio and characteristic: | | 0.05 | 0.47 | 0.17 | -0.24 |
| 50/10 Ratio* Mom HS Dropout | 0.041 (0.015) | 0.040 (0.017) | 0.040 (0.016) | 0.036 (0.015) | 0.029 (0.015) |
| 50/10 Ratio* Mom HS Graduate | 0.025 (0.017) | 0.025 (0.015) | 0.025 (0.017) | 0.019 (0.012) | 0.020 (0.015) |
| State Characteristic* Mom HS Dropout | --- | 0.0008 (0.0008) | 0.050 (0.396) | -0.001 (0.003) | -0.003 (0.002) |
| State Characteristic* Mom HS Graduate | --- | -0.0008 (0.0004) | 0.0001 (0.204) | -0.005 (0.002) | 0.004 (0.002) |

Notes: see notes to Table 1. Interacted state characteristic is listed in column headers.

Table 7: Impact of Potentially Confounding State Characteristics on Boys' Likelihood of Dropping Out of High School, by Socioeconomic Status

| | 50/10 ratio (1) | Percent Minority (2) | Poverty Rate (3) | Incarceration Rate (x1,000) (4) |
|--|--------------------|----------------------------|------------------------|---------------------------------------|
| Correlation between 50/10 ratio and characteristic: | | 0.41 | 0.63 | 0.44 |
| 50/10 Ratio* Mom HS Dropout | 0.041 (0.017) | 0.053 (0.018) | 0.056 (0.026) | 0.043 (0.021) |
| 50/10 Ratio* Mom HS Graduate | 0.024 (0.012) | 0.021 (0.017) | 0.021 (0.021) | 0.008 (0.014) |
| State Characteristic* Mom HS Dropout | --- | -0.0007 (0.0004) | -0.003 (0.004) | -0.047 (0.092) |
| State Characteristic* Mom HS Graduate | --- | 0.0001 (0.0003) | 0.001 (0.002) | 0.066 (0.045) |

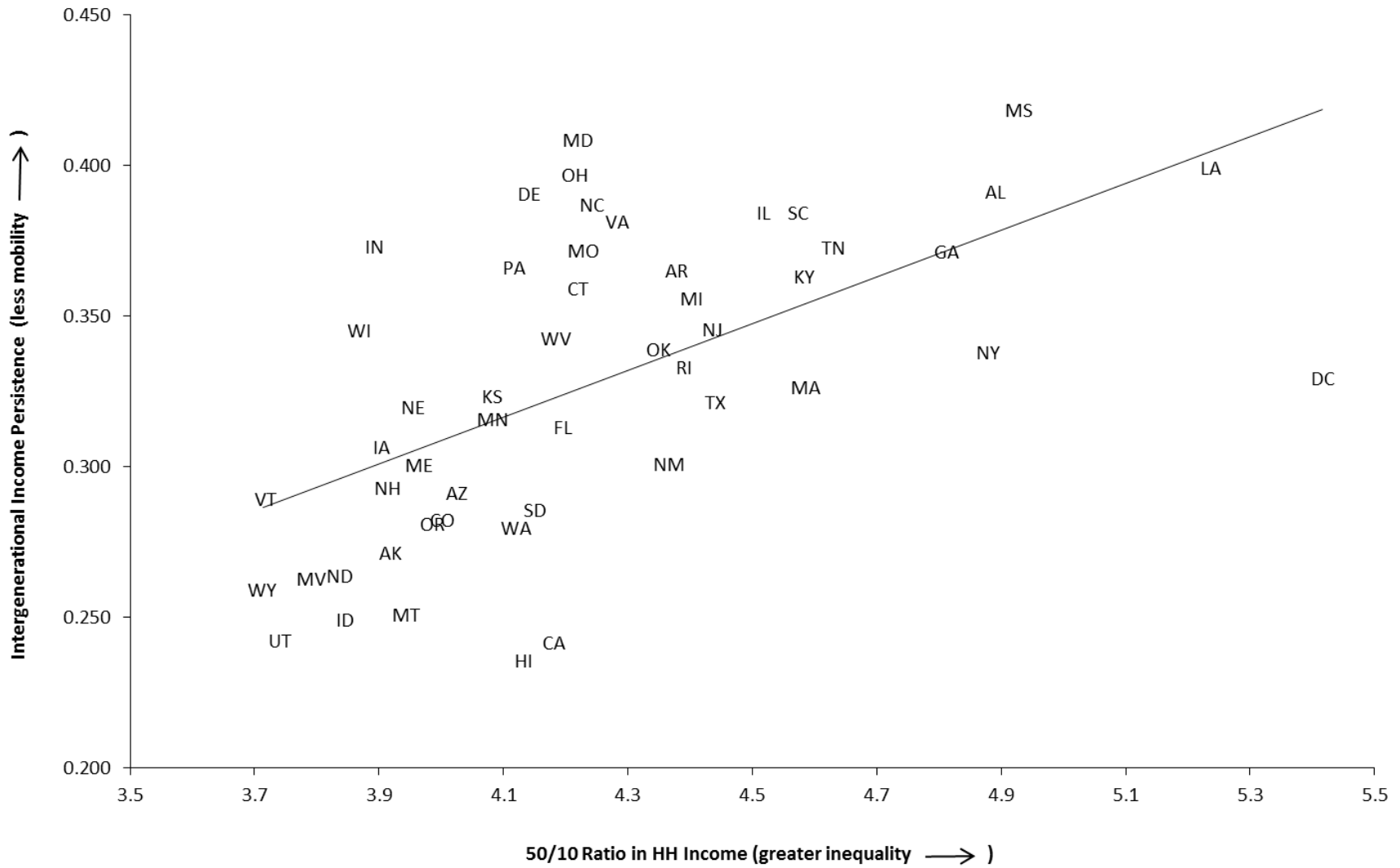
Notes: See notes to Table 1. Interacted state characteristic is listed in column headers.

Table 8: Relationship between Socioeconomic Status, Inequality, and AFQT Scores for Boys

| Sample | All 5 Datasets (1) | NLSY79 and NLSY97 (2) | NLSY79 and NLSY97 (3) | NLSY79 and NLSY97 (4) |
|-----------------------------|-----------------------|-----------------------------|-----------------------------|-----------------------------|
| Dependent Variable | HS Dropout | HS Dropout | HS Dropout | AFQT |
| Mean of Dependent Variable | 11.2 | 17.7 | 17.7 | 50.7 |
| Mom HS Dropout*50/10 Ratio | 0.042 (0.016) | 0.067 (0.029) | 0.045 (0.028) | -4.38 (2.48) |
| Mom HS Graduate*50/10 Ratio | 0.024 (0.018) | 0.077 (0.025) | 0.057 (0.023) | -4.05 (2.29) |
| AFQT | --- | --- | -0.005 (0.0002) | |

Notes: Reported standard errors in parenthesis are adjusted for clustering at the state level. Estimates in Column 1 differ slightly from previous estimates because no state level policy variables are included. The sample used in Column 2 is restricted to those observations with available AFQT scores to compare to Column 3. The samples used for the regressions in Columns 2 through 4 are the NLSY79 and NLSY97 and the sample size is 7,955.

Figure 1: Great Gatsby Curve in the United States



Notes: Income persistence is the relative mobility measure obtained from Chetty, et al. (2014). The 50/10 ratios are calculated by the authors.

Figure 2: Relationship between Inequality and the Rate of High School Non-Completion

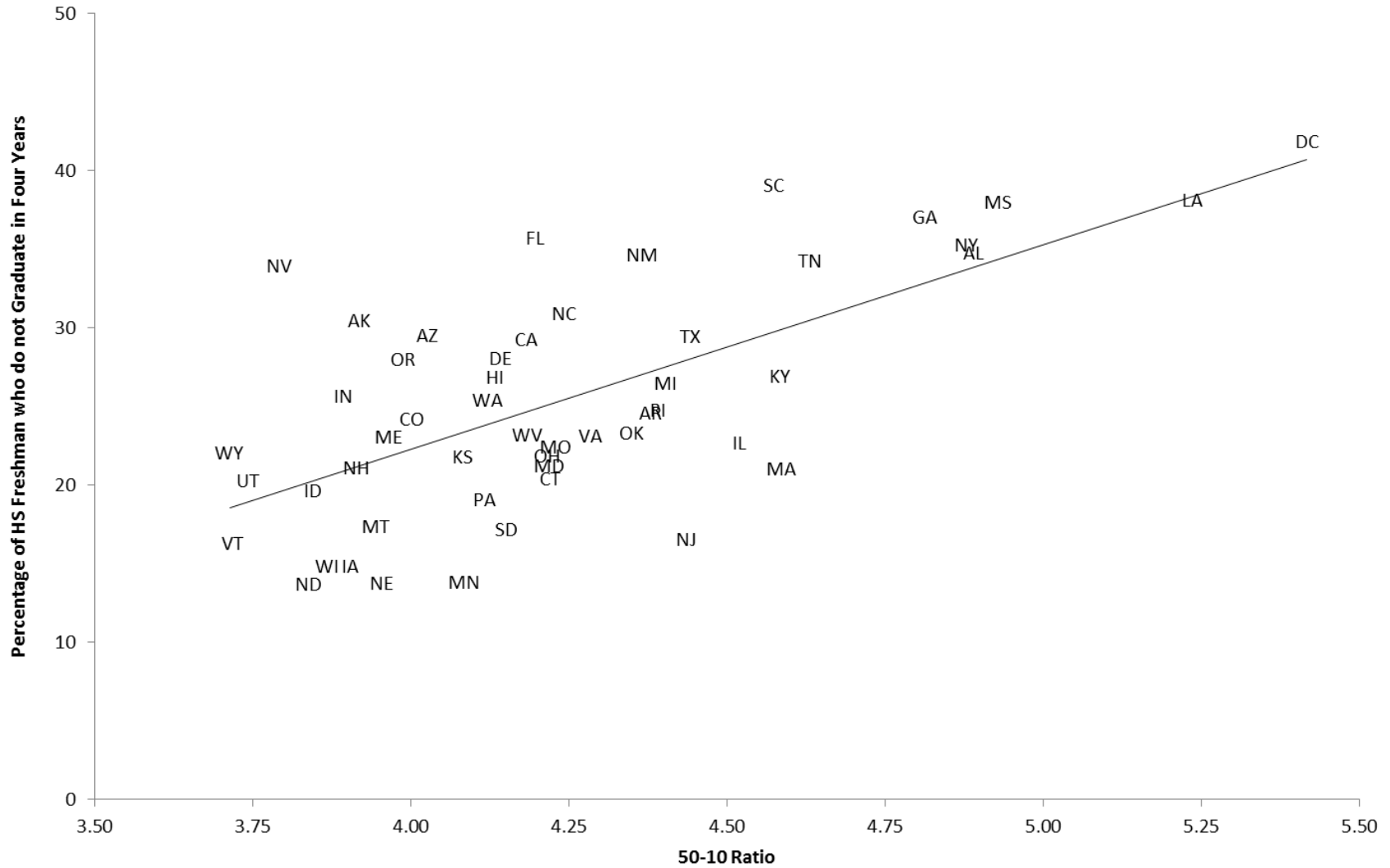


Figure 3: Relationship between Income Persistence and the Rate of High School Non-Completion

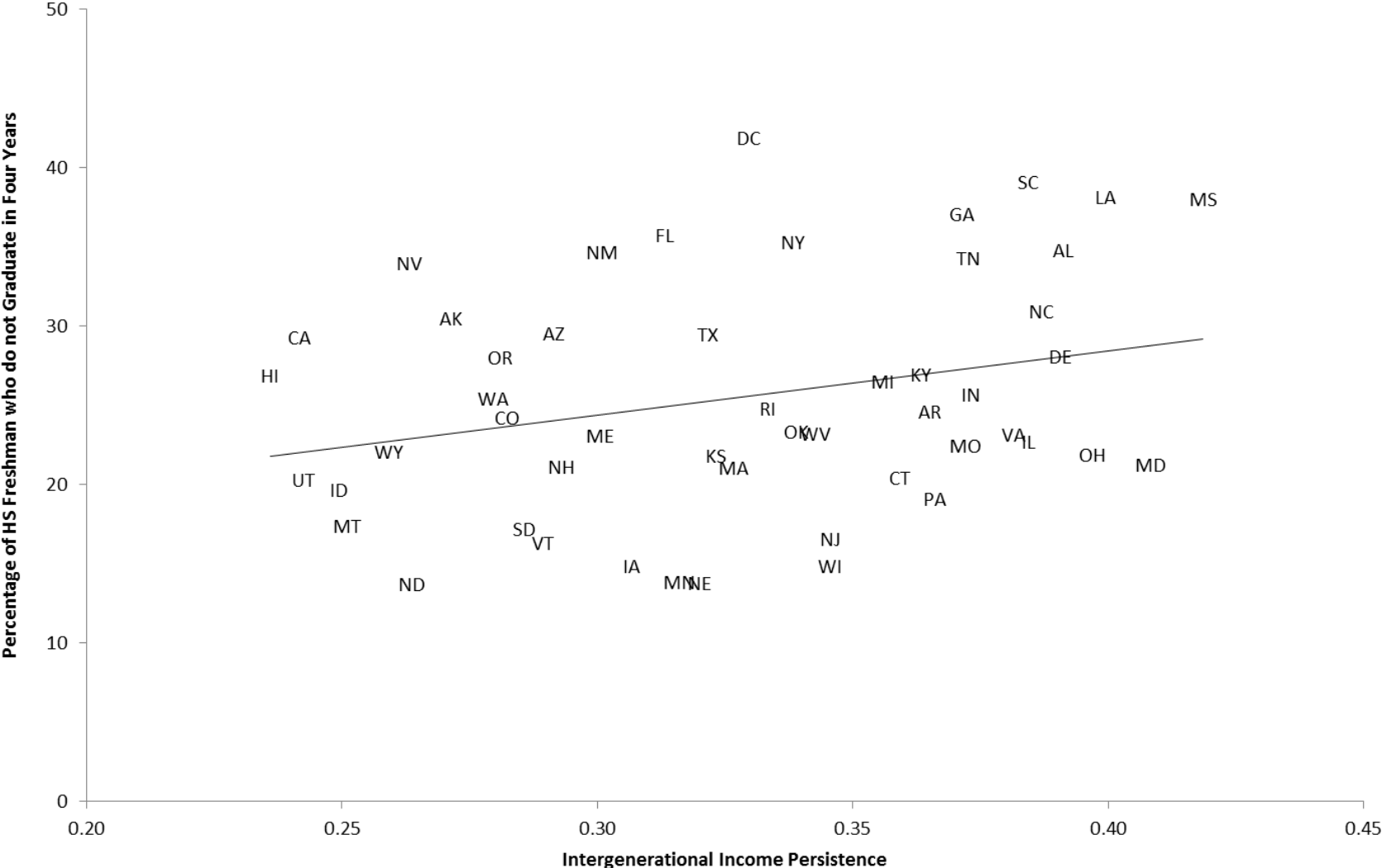


Figure 4: High School Dropout Rate for Boys by Mother's Level of Education and State Level of Income Inequality

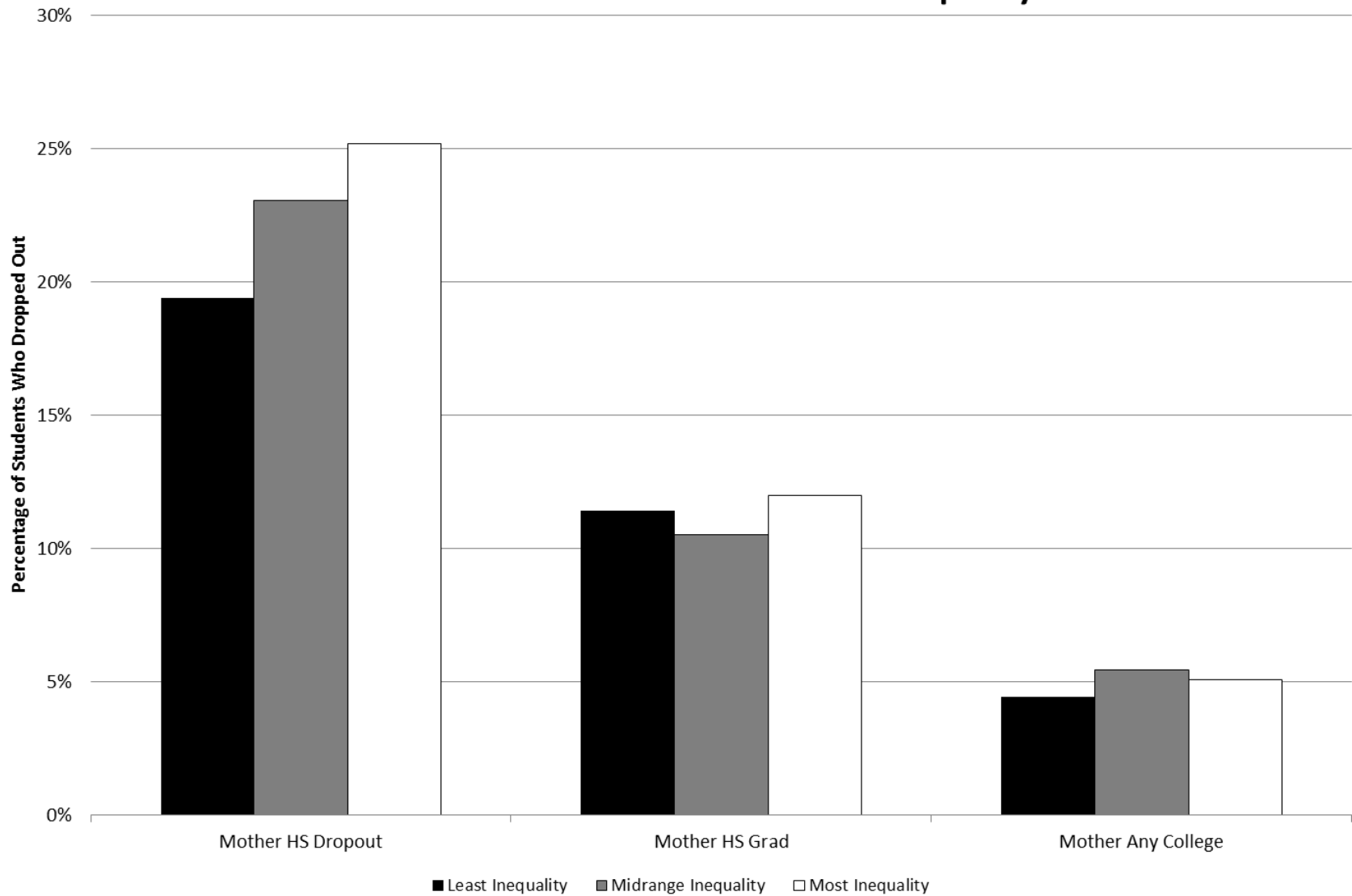
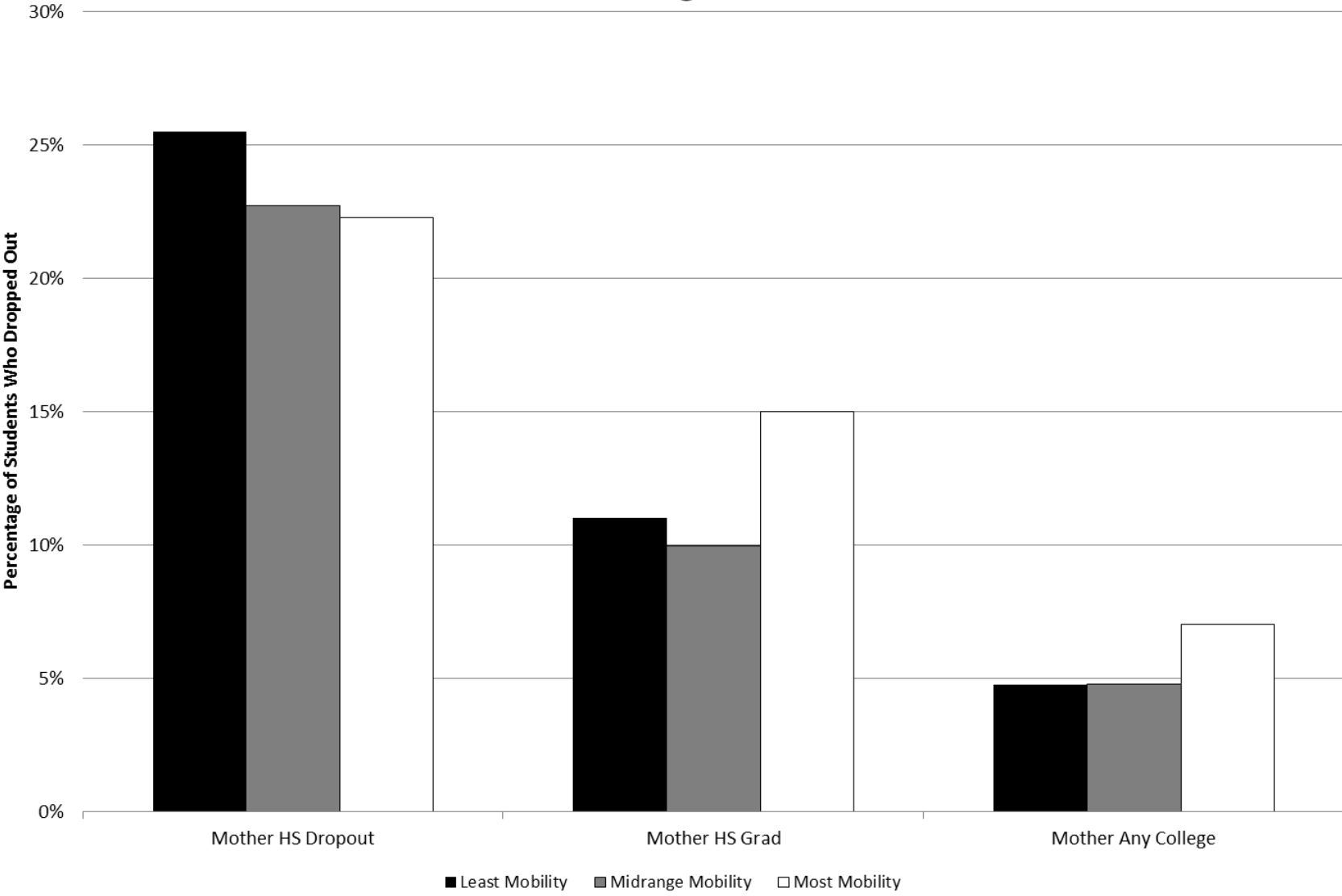


Figure 5: High School Dropout Rate for Boys by Mother's Level of Education and State Level of Intergenerational Income Persistence



**DATA APPENDIX:
MEASURING EDUCATIONAL ATTAINMENT IN NLSY AND NCES DATA**

This data appendix provides further details regarding the specific samples used in our analysis. All calculations performed include the sample restrictions described in the text, where we indicate that respondents whose educational attainment by age 20 is unknown and those whose mother's educational attainment is unknown are not included in the sample.

I. NATIONAL LONGITUDINAL SURVEYS OF YOUTH

A. 1979 Cohort

This data source originally surveyed 12,686 respondents born between 1957 and 1964, who were between the ages of 14 and 22 on the first survey date in 1979. The sample is not nationally representative, but sample weights are available to provide national representative estimates. Retention rates have been very high in these data, reducing the likelihood of attrition bias, particularly over relatively short periods. Respondents were re-interviewed every year through 1994 and then every other year after that. Because the NLSY is not a school-based survey, the universe of respondents is not restricted to those currently enrolled in a certain grade, as in the NCES data sources described below. On the other hand, some respondents are older than mandatory schooling ages on the initial survey and report their ultimate educational attainment and the timing of its completion retrospectively, introducing the possibility of recall bias.

B. 1997 Cohort

These data include information on 8,984 respondents who were born between 1980 and 1984, making them 12 to 18 on the first survey date. The sample is not nationally representative, but weights are available to provide nationally representative estimates.

Retention rates have been very high in these data, reducing the likelihood of attrition bias, particularly over relatively short periods. Respondents have been re-interviewed every other year since 1997 with the most recent available survey having been completed in 2011. Relative to the NLSY79, these data have the advantage that virtually all students are still in school at the time of the initial survey, so we can more reliably track their high school degree status as they age.

II. NATIONAL CENTER FOR EDUCATIONAL STATISTICS DATA

A. High School and Beyond (HSB)

HSB initially surveyed high school sophomores and seniors in the spring of 1980; we restrict our attention to the sophomores, most of whom were around 16 years old in that year. Respondents were re-interviewed every two years through 1986 and then again in 1992. HSB is a school-based survey; specific schools were selected to participate and the survey was administered to several students within the school. Over 30,000 sophomores in 1,015 high schools were surveyed in 1980. Of the original sample, half were selected to participate in the follow-up surveys and 79 percent responded to the follow-up survey. We restricted our analysis to those students who also participated in the base year survey.

We measure respondents' educational attainment in the second follow up, conducted in the spring and summer of 1984, when the respondents would have been around 20 years old. The second follow-up survey asks a direct question about whether respondents had graduated from high school. Respondents could have reported in response that they had graduated, had left school, were still enrolled in school, or whether they had earned a GED. For those still enrolled in school in the second follow-up, we code them as not having completed their degree by age 20 (i.e. as a "drop out").

B. National Education Longitudinal Survey (NELS)

NELS initially surveyed 8th graders in the spring of 1988, when most of them were 14 years old. They were re-interviewed in 1990, 1992, 1994, and 2000. In total, 14,915 respondents were interviewed initially in 1988 and again in the 1994 round, which represents the point at which we measure educational outcomes. Survey responses regarding educational attainment were recorded in each of these survey years and a subsample of these responses were checked against transcript records indicating their accuracy. The survey excluded 5.4 percent of selected students in the base year “because of physical or mental disabilities, or because of limited English language proficiency” (Ingels and Quinn, 1996). This restriction introduces sample selection bias since these students are more likely to drop out of high school subsequently.

The sample was “freshened” in subsequent surveys so that representative estimates could be drawn for the sophomore class in 1990 and the senior class in 1992. We focus on those respondents surveyed in the base year because using respondents from the refreshed sample would introduce an upward bias in measures of educational attainment in these data. Those students who have made it to their sophomore or senior years are a positively selected group of students, as we discuss below.

C. Education Longitudinal Survey of 2002 (ELS)

This survey included students who were in 10th grade in the spring of 2002. Students were re-surveyed in 2004 and 2006, so that they are around 20 years old in the latest year of available data. There were 15,300 students who responded to both the base year survey and the 2006 survey, when educational outcomes were measured.

III. DESCRIPTIVE STATISTICS

As Heckman and LaFontaine (2010) make clear, comparing educational attainment statistics from various micro datasets is a difficult task because of the idiosyncrasies of each. It is not our goal to track trends using these datasets, as they did, so we do not try to reconcile these differences. We do report statistics on educational attainment from each of them, though, for the purposes of detailing those differences and comparing the calculated statistics with outside sources for verification. In our econometric analysis, we control for these differences by including “dataset fixed effects.”

Appendix Table 1 indicates the percentage of survey respondents who graduate from high school, receive a GED, or drop out of high school in each dataset. Sample weights are used to adjust for the various sampling techniques used in each dataset. Discrepancies across datasets are extensive, yet they are consistent with past estimates (allowing for modest variation attributable to the sample restrictions we impose). For instance, Hill and Holzer (2007) examine data from the two NLSY surveys. We focus on educational attainment by age 20 and they focus on educational outcomes between 20 and 22. We find that 16.3 and 12.2 percent dropped out of high school and 5.1 percent and 6.9 percent have a GED in the 1979 and 1997 surveys, respectively. Their results are comparable: 16.8 percent and 12.8 percent dropped out and 4.3 percent and 5 percent earned a GED in the respective surveys. For HSB, we are able to replicate reported results (National Center for Educational Statistics, 1984), although our sample restrictions change the final values reported in this table somewhat. We match previous estimates because we are coding high school completion status directly for a single survey question. In the NELS, we estimate that 9.3 percent of students drop out and 5.0 percent of students obtain a GED by around age 20. Our estimates are comparable to those in Hurst, et al. (2004), who find that 12 percent of students drop out and 6 percent of students obtained a GED

by 1994 (when most respondents are age 20). In the ELS, we estimate that 7.5 percent of students drop out of high school and 4.3 percent obtain a GED by around age 20. This compares to 7.8 percent and 3.9 percent, respectively, reported in Bozick, et al. (2007).

It is difficult to determine the extent to which the differences in estimates across datasets are attributable to changes in outcomes over time or the differences in the nature of the surveys. An important difference is the sampling strategies used by the different surveys. The three NCES surveys are school-based and require students to be still enrolled in school to participate. This is particularly troublesome with the HSB and ELS surveys, in which youth need to “survive” to 10th grade to participate. In the NELS, students only need to “survive” to 8th grade, which is less likely to introduce bias. Nevertheless, these sampling strategies indicate that we should expect higher dropout rates in the two NLSY surveys, which is exactly what we see. It would also be reasonable that NELS had the next highest dropout rate and that hypothesis is confirmed in these data as well.

To better document this problem, we use data from the two NLSY surveys to examine the degree status of students and their highest grade completed at age 20. The results are presented in Appendix Table 2. In the NLSY79 and NLSY97, 5.1 percent and 6.9 percent of respondents, respectively, never make it to 10th grade by age 20. Of those who fail to reach that grade, most drop out. Omitting those students from the sample, as occurs in the HSB and ELS, imposes an upward bias in educational attainment. Indeed, this is a problem, albeit considerably smaller, even when starting a sample in 8th grade, as occurs in the NELS. Around half a percent of students fail to reach that grade in the two NLSY surveys.

Appendix Table 1: Educational Attainment Measured
in Alternative Longitudinal Data Sources.

| Educational Attainment by Age 20 | | | |
|----------------------------------|-----|------------------------|-------------------------|
| | GED | High School Dropout | High School Graduate |
| NLSY79 | 5.1 | 16.3 | 78.6 |
| HSB (1980) | 3.8 | 7.1 | 89.2 |
| NELS (1988) | 5.0 | 9.3 | 85.7 |
| NLSY97 | 6.9 | 12.2 | 81.0 |
| ELS (2002) | 4.3 | 7.5 | 88.3 |

Source: Authors' calculations.

Appendix Table 2: Degree Status by Highest Grade Completed at Age 20

| | Below 8 th Grade | 8 th Grade | 9 th Grade | 10 th Grade | 11 th Grade | 12 th Grade and Higher |
|-----------------------|--------------------------------|--------------------------|--------------------------|---------------------------|---------------------------|--------------------------------------|
| NLSY79 | | | | | | |
| Percent at Level | 0.7 | 1.6 | 2.8 | 4.3 | 5.0 | 85.6 |
| <u>Degree Status:</u> | | | | | | |
| HS Dropout | 99.4 | 97.4 | 95.7 | 95.5 | 89.5 | 3.3 |
| GED | 0.0 | 1.3 | 3.1 | 4.4 | 4.8 | 5.4 |
| HS Graduate | 0.6 | 1.3 | 1.3 | 0.1 | 5.8 | 91.4 |
| NLSY97 | | | | | | |
| Percent at Level | 0.4 | 2.6 | 3.9 | 5.1 | 6.0 | 82.0 |
| <u>Degree Status:</u> | | | | | | |
| HS Dropout | 93.9 | 73.7 | 63.6 | 58.1 | 59.5 | 1.0 |
| GED | 4.6 | 23.2 | 33.5 | 38.8 | 33.5 | 1.2 |
| HS Graduate | 1.6 | 3.1 | 2.9 | 3.2 | 7.0 | 97.8 |

Source: Authors' calculations.

SOURCES:

Bozick, Robert, Erich Lauff and John Wirt (2007). *Education Longitudinal Study of 2002 (ELS:2002): A First Look at the Initial Postsecondary Experiences of the High School Sophomore Class of 2002*. Washington, DC: National Center for Education Statistics.

Heckman, James J. and Paul A. LaFontaine (2010). "The American High School Graduation Rate: Trends and Levels," *Review of Economics and Statistics* 92 (2): 244-262.

Hill, Carolyn J. and Harry J. Holzer (2007). "Labor Market Experiences and the Transition to Adulthood," in Sheldon Danziger and Cecilia Rouse (eds.) *The Price of Independence: The Economics of Early Adulthood*. New York: Russell Sage Foundation.

Hurst, David, Dana Kelly, and Daniel Princiotta (2004). *Educational Attainment of High School Dropouts 8 Years Later*. Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Ingels, Steven J., and Peggy Quinn (1996). *Sample Exclusion in NELS:88: Characteristics of Base Year Ineligible Students; Changes in Eligibility Status after Four Years [National Education Longitudinal Study of 1988 Second Follow-Up]*. Washington, DC: U.S. Department of Education. National Center for Education Statistics.

National Center for Education Statistics (1984). *High School and Beyond, 1980: Sophomore Cohort Second Follow-up (1984). Volume I*. Washington, DC: Government Printing Office.