

‘YOU RAISE ME UP:’ INVESTIGATING MULTI-GENERATIONAL TRENDS AND
THE IMPACT OF ROLE MODEL EFFECTS IN WOMEN’S INTERGENERATIONAL
SOCIOECONOMIC MOBILITY IN THE UNITED STATES

A THESIS

Presented to

The Faculty of the Department of Economics and Business

The Colorado College

In Partial Fulfilment of the Requirements for the Degree

Bachelor of Arts

By:

Owen Raymond Rask

May 2024

‘YOU RAISE ME UP:’ INVESTIGATING MULTI-GENERATIONAL TRENDS AND
THE IMPACT OF ROLE MODEL EFFECTS IN WOMEN’S INTERGENERATIONAL
SOCIOECONOMIC MOBILITY IN THE UNITED STATES

Owen Raymond Rask

May 2024

Economics

Abstract

Studies of intergenerational mobility have often focused on single father-to-son generational relationships. In this study, we utilize the NLSY79 and NLSY97 datasets alongside a Heckman two-step correction to analyze multigenerational trends of U.S. socioeconomic mobility by sex and race. We measure mobility through intergenerational rank association (IRA), also investigating the multigenerational impact of role model effects on women’s mobility. First, we find that the role model effect has grown significantly for U.S. women born between 1957 to 1965 and 1981 to 1985, chiefly driven by white women. Second, like previous studies, we find minimal to no increase in U.S. mobility overall between these generations and negligible increases in U.S. women’s and men’s mobility when examined by sex. Thirdly, we find evidence supporting previous findings that a notable mobility gap exists between black and white American men. Additionally, our estimates indicate that this gap may increase over our investigated timeframe. Lastly, we show that with a significantly smaller sample not encompassing an entire population, through creating income percentiles from large micro datasets and utilizing a two-step Heckman correction, accurate intergenerational socioeconomic mobility measurements can be obtained compared to previous literature that utilizes an entire country cohort data.

KEYWORDS: Intergenerational Socioeconomic Mobility, Gender, Race
JEL CODES: J00, J62

ON MY HONOR, I HAVE NEITHER GIVEN NOR RECEIVED
UNAUTHORIZED AID ON THIS THESIS

A handwritten signature in black ink, reading "Owen Pasik". The signature is written in a cursive style with a large initial "O" and a stylized "P".

Signature

Table of Contents

Abstract.....	II
List of Tables	V
List of Figures.....	VI
Acknowledgments	VII
1. Introduction.....	1
1.1 Motivation.....	1
1.2 Background Information	2
2. Literature Review	6
3. Theory and Methodology	9
3.1 Theory and Underlying Assumptions	9
3.2 Measures of Intergenerational Socioeconomic Mobility	12
3.3 Methodology	13
4. Data	16
5. Results and Analysis	22
5.1 Findings Compared to Previous Literature	22
5.2 US Multigenerational Trends	26
5.3 US multigenerational Trends by Region	27
5.4 US Multigenerational Trends by Sex	31
5.5 US Multigenerational Trends by Race.....	37
5.6 US Multigenerational Trends by Race and Sex	40
5.7 Corrected Models and Role Model Effects	47
5.8 Education and Intermediate Outcomes	50
6. Conclusion	53
References.....	56
7. Appendix.....	59
7.1 Variable Definitions	59
7.2 Additional Tables	62

List of Tables

Table 1. NLSY 1979 Sample Summary Statistics.....	21
Table 2. NLSY 1997 Sample Summary Statistics.....	22
Table 3. NLSY 1979 and 1997 Overall OLS Regression Outputs.....	24
Table 4. NLSY 1979 Region OLS Regression Outputs.....	29
Table 5. NLSY 1997 Region OLS Regression Outputs.....	29
Table 6. NLSY 1979 and 1997 Women OLS Regression Outputs.....	34
Table 7. NLSY 1979 and 1997 Men OLS Regression Outputs.....	35
Table 8. NLSY 1979 and 1997 Race OLS Regression Outputs.....	38
Table 9. NLSY 1979 and 1997 Black Women and Men OLS Regression Outputs.....	42
Table 10. NLSY 1979 and 1997 Hispanic Women and Men OLS Regression Outputs.....	42
Table 11. NLSY 1979 and 1997 Other Women and Men OLS Regression Outputs.....	43
Table 12. NLSY 1979 and 1997 Women Probit and OLS Regression Outputs.....	48
Table 13. NLSY 1979 and 1997 Women by Race Probit Mean Marginal Effects.....	50
Table 14. NLSY 1979 Respondent Education Level OLS Regression Outputs.....	52
Table 15. NLSY 1997 Respondent Education Level OLS Regression Outputs.....	52

List of Figures

Figure 1 Linear Relationship of Parent to Respondent Income Percentiles.....	25
Figure 1.1 Linear Relationship of NSLY79 Sample.....	25
Figure 1.2 Linear Relationship of NSLY97 Sample.....	25
Figure 2 NLSY79 and NLSY97 Overall Cohort Mobility Regressions.....	26
Figure 3 Linear Relationships of Parent to Respondent Percentiles by Region.....	28
Figure 3.1 Linear Relationship for U.S. North East.....	28
Figure 3.2 Linear Relationship for U.S. North Central.....	28
Figure 3.3 Linear Relationship for U.S. West.....	28
Figure 3.4 Linear Relationship for U.S. South.....	29
Figure 4 NLSY79 and NLSY97 Cohort Mobility Regressions by Sex.....	35
Figure 5 NLSY79 and NLSY97 Cohort Mobility Regressions by Race.....	38
Figure 6 NLSY79 and NLSY97 Cohort Mobility Regressions by Race and Sex.....	43
Figure 6.1 Non-Corrected Models for U.S. Women Racial Groups.....	43
Figure 6.2 Corrected Models for U.S. Women Racial Groups.....	44
Figure 6.3 Models for U.S. Men Racial Groups.....	44
Figure 7 Linear Relationship of Parent to Respondent Percentiles by Sex.....	62
Figure 7.1 US Women NLSY79 and NLSY97 Linear Relationships.....	62
Figure 7.2 US Men NLSY79 and NLSY97 Linear Relationships.....	62
Figure 8 Linear Relationship of Parent to Respondent Percentiles by Race.....	62
Figure 8.1 NLSY79 and NLSY97 Linear Relationships Black Americans.....	62
Figure 8.2 NLSY79 and NLSY97 Linear Relationships Hispanic Americans.....	63
Figure 8.3 NLSY79 and NLSY97 Linear Relationships Other Americans.....	63
Figure 9 Linear Relationship of NLSY79 and NLSY97 Parent to Respondent Percentiles by Race and Sex.....	63
Figure 9.1 NLSY79 Linear Relationships Black American Men and Women.....	63
Figure 9.2 NLSY79 Linear Relationships Hispanic American Men and Women.....	63
Figure 9.3 NLSY79 Linear Relationships Other American Men and Women.....	64
Figure 9.4 NLSY97 Linear Relationships Black American Men and Women.....	64
Figure 9.5 NLSY97 Linear Relationships Hispanic American Men and Women.....	64
Figure 9.6 NLSY97 Linear Relationships Other American Men and Women.....	64
Figure 10. Linear Relationship of NLSY79 and NLSY97 Respondent Attained Education Level to Respondent Percentiles by Race and Sex.....	65
Figure 10.1 NLSY79 Black American Men and Women's Education.....	65
Figure 10.2 NLSY79 Hispanic American Men and Women's Education.....	65
Figure 10.3 NLSY79 Other American Men and Women's Education.....	65
Figure 10.4 NLSY97 Black American Men and Women's Education.....	66
Figure 10.5 NLSY97 Hispanic American Men and Women's Education.....	66
Figure 10.6 NLSY97 Other American Men and Women's Education.....	66

Acknowledgments

I would like first to thank Dr. Minho Kim, who, alongside being a close advisor and beloved teacher throughout my undergraduate career, took a chance on a college junior and sponsored the genesis of what turned into a research project and, eventually, this thesis. Secondly, thank Dr. Daniel Johnson for being my thesis advisor and advising me along the finalization process in understanding how to communicate the work you are currently reading. Both men have been instrumental in the creation and completion of this project. Thank you to the CC SCoRe grant program and Lisa Schwartz for providing the initial funding to begin this project during my Junior summer. Thanks to Samuel Frykholm, Walt Jones, and Gretchen Wardell for often indulging me and letting me talk about this project at them as I work through my thoughts. Last (but certainly not least), I would like to thank my parents for conversing with me consistently regarding this research project and offering their thoughts and advice—especially my father.

Introduction

Motivation

In the twentieth century, the United States was touted as the land where one's ability could bring them from rags to riches. With the economic prosperity that occurred during WWII and continued after its conclusion, the economic makeup of the United States society was radically altered. Accompanying these substantial economic shifts were numerous social justice movements as those who had long been excluded from the economic world of the U.S. fought for access to its growing opportunities and wealth.

However, for US women, a group who arguably made the most considerable social and economic strides during this time, little is known about their economic mobility throughout these years. What is known is that throughout the twentieth century, women's involvement in the United States workforce increased dramatically. The initial surges in women's employment were in the 1930s to 1950s, primarily induced by the Great Depression and WWII, which created labor supply shocks, leading to a significant influx of women in the US labor force (Bellou & Cardia; 2021Goldin & Olivetti 2013).

These initial surges translated into persistent trends, as post-WWII women's workforce participation increased due to cultural changes, technological improvements, economic growth, and habit formation throughout the twentieth century. By the mid to late 1900s, the succeeding cohorts of US women entering the workforce flattened their lifetime labor force participation cycle as they began to view their employment with long-term horizons, investing more in education, having higher likelihoods of re-entering the workforce after having children, and working later into their old age (Goldin & Mitchell 2017; Goldin 2006). While the evidence

indicating the persistent growth in US women's labor force participation is apparent and documented, determining the magnitude of economic advancement of US women during this time is less clear. This successive generational employment was accompanied by a reduction in their real wages, and the benefits of increased employment were primarily reaped by educated white women (Bellou & Cardia 2021). Thus, research on the trends of non-white US women's historical socioeconomic movement is sparse.

Interestingly, the academic and political interest in intergenerational economic mobility has recently surged in popularity. These studies typically aim to quantify country-wide mobility in terms of the relation of the independence of a parent's lifetime socioeconomic status to their child's socioeconomic status (SES) when in their 30s-40s. This tactic has led researchers to investigate human capital predictors on economic status, primarily education and ability measurements (Chetty et al. 2019; Greg et al. 2018; Jajtner 2020). With the increased attention on intergenerational socioeconomic mobility and the small number of studies on the differences in mobility within country populations (especially those historically oppressed or excluded from economic activity), we aim to add to the small literature that addresses those research gaps.

Background Information

Studies of intergenerational mobility have historically focused on father-to-son relationships and singular gender data to obtain mobility measurements (Gregg et al., 2018.; Luke, 2021). These studies have traditionally relied on specific sample populations within countries—usually, the majority in the country studied, often white men. This precedent is due to challenges associated with modeling women's participation in the labor force—especially for older datasets—and finding reliable

longitudinal data for income and labor force participation for an entire country's population. For these reasons, there are few studies of intergenerational mobility utilizing father-to-daughter SES, family-to-daughter SES, or mother-to-daughter SES. This study addresses these gaps in the literature by focusing on the intergenerational socioeconomic mobility of marginalized groups. We investigate U.S. populations traditionally excluded from these studies, such as Black, Hispanic, and white women.

Chetty et al. have led the way in addressing these gaps in the last decade. Their 2014a and 2014b studies present findings for trends in U.S. women's mobility and teenage birth rates, while in their 2019 study, they provide a comprehensive breakdown of U.S. mobility by race for a U.S. cohort born in the 1980s. Chetty et al. have surmounted the issues that limited previous researchers by obtaining data representative of the entire U.S. birth cohorts for their years under study. However, this data is hard to obtain and only permits the collection of specific variables. We hope to recreate the results of Chetty et al.'s pioneering studies on U.S. intergenerational socioeconomic mobility trends utilizing much smaller samples (the NLSY79 and NLSY97 datasets), which, while not encompassing the entirety of the U.S. population, provide extensive individual characteristics with which a researcher could study alongside being free and widely available. If our results are similarly conclusive and robust as previous literature with our smaller samples, it could open the door for future research that investigates more specific questions regarding intergenerational mobility than presented in this study or previous literature.

Studies of socioeconomic mobility often measure mobility in one or two ways: intergenerational elasticity (IGE) and intergenerational rank association (IRA). IGE is a relative measure of mobility, with high IGE's signifying low mobility while low IGE's signify relatively high mobility for the group or society in question. However,

many studies in the past decade have shown IGE to be incredibly sensitive; thus, IRA has emerged as the alternative and more robust method to measure intergenerational mobility (Chetty et al., 2014a; Chetty et al., 2014b; Dahl & Delaire, 2008). Following recent trends in the literature, this study utilizes the traditional method of calculating IRA for both our cohorts. The decision to focus exclusively on IRA arises from its well-documented robustness compared to IGE in the literature, especially when comparing between generations and different datasets—as is necessary for this study.

With our samples not encompassing entire birth cohorts, we institute specific corrective measures to address issues that arise when including populations traditionally excluded in models of intergenerational socioeconomic mobility. We employ a Heckman two-step correction to check for potential selection bias in our women cohorts and source income census income microdata from IPUMS CPS to create our percentile ranks. To our knowledge, we are the first to utilize this type of correction in a two-stage least squares regression model of intergenerational socioeconomic mobility. This correction addresses the challenge of modeling women's income through accounting for self-selection bias in women leaving the workforce for childcare (Heckman, 1979; Mulligan & Rubenstein, 2008).

Additionally, we look multigenerational, comparing a generation of women and men born in the U.S. in the late 1950s to early 1960s and a generation of women and men born in the 1980s. We utilize the NLSY79 and NLSY97 to keep our data as consistent and similar as possible across the generations. To our knowledge, this is the first U.S. multigenerational study of socioeconomic mobility measuring IRA that compasses a cohort born in the United States prior to 1970.

When looking multi-generationally, multiple other questions arise which could be investigated. One of those questions in this study's scope—due to the inclusion of

the Heckman correction—was to investigate the influence of role model effects on women’s employment and, subsequently, income over the generations. Thus, a last line of inquiry this study considers is that accompanying the growth of US women’s prevalence within the workforce, to what degree does a mother working outside of the household encourage their daughter to do the same in adulthood, and does that social effect have a sizeable enough return on women’s income prospects to influence their mobility. Recent evidence suggests that role model effects are critical in individuals’ choices and experiences in life—especially for underrepresented groups—from educational experience to career choice decisions (Kofoed & McGovney, 2017; Humlum et al., 2012). Here, we add to this literature by analyzing the impact of a female respondent’s mother working outside of the house during their teenage years and how this factor impacts their employment likelihood and adulthood income percentile rank across generations.

This study investigates three primary questions. First, has U.S. intergenerational socioeconomic mobility changed for those born in the U.S. between 1957 and 1965 and those born between 1981 and 1985, and if so, specifically for whom? Second, what has been the impact of role model effects on the historical increase in women’s employment, and did it contribute to any changes in their intergenerational mobility? Third, to what degree of accuracy can U.S. population mobility measurements (in the form of IRA calculations) be obtained using certain corrections on sample cohort datasets?

The following section reviews the literature on intergenerational mobility, the Heckman two-step correction, and the impact of role model effects. Section 3 explains our methodology, while Section 4 outlines all the data used in this study and how

variables were defined. Our results and a discussion of those findings are in section 5. We provide our overall conclusions in section 6.

Literature Review

Studies of intergenerational socioeconomic mobility typically utilize the relationship between a father and son's logged income to calculate estimates of intergenerational elasticity (IGE). These measurements of IGE were interpreted as estimations of the degree of relative mobility in the studied country. However, IGE has fallen out of favor in the last decade, as recent studies have shown how it can provide drastically different measurements based on regression variable specification—such as to include or exclude observations that report zero income. This issue was exhaustively explored in Chetty et al.'s 2014a paper.

Intergenerational rank association (IRA) has come to supersede IGE as the preferred statistic to calculate when investigating intergenerational socioeconomic mobility. Dahl and DeLeire first proposed intergenerational rank association in their 2008 study. IRA follows the same process as IGE; however, instead of regressing a child's proxied lifetime logged income on their parent's proxied lifetime logged income, income percentiles are utilized. This change allows for more robust, accurate, and consistent measurements (Chetty et al., 2014a; Chetty et al., 2014b; Mitnik et al., 2015).

This change in measurement has been accompanied by a shift in intergenerational socioeconomic mobility questions researchers have pursued. There has been a movement from attempting to solely quantify the entire mobility of a country to identifying group differences within countries and institutions that promote mobility—primarily regarding differing types of education (Chetty et al., 2017; Gregg et al., 2019). However, studies of intergenerational mobility have only recently begun

to focus on within-society differences.

In their 2017 paper, Piotrowska and Kośny investigate Poland's political transformation in the 1990s and its impact on intergenerational mobility, specifically that of women. They analyze mother-to-daughter relationships from the 1990s to the 2010s, finding that the three cohorts of mother-to-daughter relationships had increasing mobility measures over time, moving from 0.36 to 0.52 to 0.56¹. Jajtner, in her 2020 study, shows how children of work-limited parents experience less absolute intergenerational mobility relative to their peers without work-limited parents, and this effect is magnified for daughters.

Chetty et al.'s 2014a and 2014b papers find that mobility differences within the U.S. are primarily a local issue. One's commuter zone greatly influences their upward and relative mobility. Also, intermediate outcomes (such as education and teen birth rate) are not significantly different across zones or generations; thus, mobility differences arise before U.S. citizens enter the workforce.

Chetty et al.'s 2019 paper investigates differences in intergenerational mobility based on race in the US. They find that Black Americans and First Nation Americans have significantly lower rates of upward mobility and higher rates of downward mobility when compared to White Americans. These racial disparities are even prevalent at the highest socioeconomic level for Black and First Nation Americans. These recent studies show that focusing on groups within populations paints a drastically different picture of country-specific mobility.

Another intra-country investigation of mobility is by examining mobility at different income brackets. These studies are motivated by recent research showing

¹ It should be noted that in this study, transition matrices were used. Thus, their mobility calculation ranged from 0 (which measured complete immobility) to 1 (which measured perfect mobility).

that the mobility for those starting at the lower end of the socioeconomic distribution is significantly different from that of those starting at a higher socioeconomic status. These studies typically employ transition matrices or subdivide their samples by income percentile brackets (Chetty et al., 2014a; Chetty et al., 2014b; Chetty et al., 2019). Additionally, other authors have branched out from this foundation, such as Gregg et al. in their 2018 study, where they utilize unconditional quantile regression to explore nonlinearities within the distributions of those studied. They find that education is dwarfed in significance for mobility compared to parental income at the top of the earnings distribution.

With this study utilizing historical women's income data, potential sample selection issues quickly arise. As mentioned, to address this issue, we follow an extensive and tested body of literature that employs a Heckman two-step correction least squared regressions procedure (Woolridge, 2010). When utilizing this correction, most previous literature includes measurements or proxies for marital status and children in the first step probit model, which are often interacted with each other (Heckman, 1976; Mulligan & Rubinstein, 2008). There has been some discussion about the endogeneity of these predictors, as they potentially influence both the choice of and outcomes from employment (Craigie, 2021). We follow this established literature and include marital status, number of kids, and their interaction alongside the respondent's region of residence during their upbringing in our first step probit equations when modeling our women cohorts.

Lastly, with this study's focus on multigenerational trends in women's socioeconomic mobility, we examine the impact of role model effects on US women's mobility. While extensive literature surrounds role-model effects in education (Carrell et al., 2010), its impact on employment is less studied. Kofoed and

McGovney (2017) show through a random assortment of gender and race-specific mentorship for West Point cadets that for women and black cadets, having a role model with similar gender or race can influence their choice of occupation post-training by 5.2-16.6% and 6.2% respectively.

Role model effects have also been shown to have relevant impacts when investigating generational trends. Hilmer and Hilmer's 2007 study shows that the increasing number of women economists has potentially reversed the findings of Neumark and Gardecki (1998) that there are differences between the impact of male and female advisors on female Ph.D. economic students. There is also theoretical grounding for this phenomenon, as economic models that account for personal and societal identity, when tested against data, show that these social factors influence an individual's educational plans and explain specific gender differences in career choice (Humlum et al., 2012).

Theory and Methodology

Theory and Underlying Assumptions

In general, the investigation of intergenerational socioeconomic mobility is based upon the notion that within a society, upward socioeconomic mobility is predicated primarily upon one's abilities, luck, and factors of birth. In this tradition, this paper follows the model laid out by Becker and Tomes (1979), where human and non-human capital investments are said to increase the quality of one's child, and the returns to those investments can then be measured by that child's income in adulthood. Alongside investments, each child receives an endowment that includes inherited genetic traits (such as race and gender), family connections and reputation, and pre-disposed skills. Thus, this model purports that a child's success is associated with the investments their parents make in them during their upbringing, the societal

benefit of the inherited endowment they receive by being born into a specific familial unit, and the current conditions of the market conditions, all accompanied by the factor of luck.

However, in discussing specific endowment traits, there is also an important difference between universally advantageous endowment factors (such as higher parent socioeconomic status) and socially specific endowment factors, such as sex and race. These society-specific factors have an inherited historical context that cannot be ignored and most greatly impact those groups in societies that have historically been marginalized. This distinction is crucial when looking inter and multi-generationally, as those who successfully climbed the socioeconomic ladder early had more resources to invest in their children over the generations, while those who could not climb—due to either ability or societal oppression—often got stuck. Over multiple generations, this creates a vicious generational feedback loop that increasingly benefits those already succeeding. In this paper, we continue to illuminate some of these specific multigenerational societal endowment trends within the United States.

Any investigation of intergenerational relationships is constrained largely by the availability of data relevant to the question under study. In the case of this study, with the desire to investigate multigenerational socioeconomic trends by race and gender and other upbringing qualifiers such as role-model effects, procuring a free and accessible dataset of sufficient panel data was a monumental task. A complete assessment of our questions would require a dataset that accurately demographically represents the U.S. by gender and race, includes relevant variables about respondents' household influences and upbringing, is representative of the socioeconomic diversity and wealth of individuals within the country, and can compile this information consistently from respondents and eventually their children for multiple decades.

These desired requirements are beyond challenging to meet, as most panel datasets have only been running for at most three generations—among the other needed factors. Thus, in the construction of this study, the (public) dataset that allowed for the best fit regarding these desired categories was chosen as the starting point. Then, the theory and methodology were molded to fit what was available. The NLSY79 and NLSY97 datasets were chosen because they were consistent, publicly available studies that could be analyzed multi-generationally, were largely representative of the US populous, and provided comprehensive individual respondent statistics.

When using these datasets, certain assumptions must be made. First, both NLSY surveys do not encompass the entire U.S. population and thus are not wholly representative of the U.S. population. This study does its best to correct these issues, but certain groups within the U.S. (such as First Nation and Asian Americans) cannot be modeled. Second is the continuity between the generations. Due to the immense similarities between how the NLSY datasets were collected, we believe there is sufficient continuity to compare both datasets. During the creation of this study, many potential family makeups and alternative questions were considered (for example, single-parent versus parents living together). This study acknowledges there is still a large amount of simplification of respondent's upbringing that was either excluded or not available in these datasets.

A last common assumption in intergenerational relationship studies concerns proxying for one's economic life cycle. If one views the socioeconomic status of a respondent or their parent in only one year, many factors could affect that single number. Thus, as we will discuss later, we average our income data across specific age years for both datasets as a proxy for their lifetime income.

Measures of Intergenerational Socioeconomic Mobility

There are two primary established methods of measuring intergenerational socioeconomic mobility. First, and what has been done in most of the literature, is utilizing intergenerational elasticity (IGE) (Chetty et al., 2014a; Gregg et al., 2018). IGE is calculated by regressing the log of a child's lifetime income on the log of their lifetime parents—where lifetime income is proxied. However, as Chetty et al. explain in their 2014a paper, IGE estimates vary widely depending on the sample definitions of an individual's income—for example, including those who report 0 income by converting them to one or dropping them. With the problems identified when calculating IGE and this study's data challenges, which necessitate using three different measurements of income (respondent income in wage/salaries, gross household income in the NLSY97 dataset, and net family income in the NLSY79 dataset), we feel it is unwise to analyze or investigate calculations of IGE.

Chetty et al.'s same 2014a paper builds on Dahl and Deleire's 2008 work to explain an alternative, more robust method of calculating socioeconomic mobility: intergenerational rank association (IRA). IRA is like IGE in that it still utilizes an OLS regression; however, for IRA, one regresses the parent's lifetime income percentile (measured against their representative cohort) on their child's lifetime income percentile (also measured against their representative cohort).

The coefficient on the parent's rank percentile is known as the rank-rank slope, which captures the association between the parent's income position in their respective distribution and their child's income position in their respective distribution. This study exclusively investigates intergenerational socioeconomic mobility utilizing IRA estimates, as discussed in Dahl and Deleire (2008) and Chetty et al. (2014a).

Methodology

In calculating IRA, we follow the precedent established by Chetty et al. in their 2014a and 2014b studies and expanded upon in their 2019 study. However, since we look multigenerational with new data, we differ slightly from their assumptions. We assume a discrete-time setting where t denotes the generation and i represents an individual in each generation. Later in the paper, we show that the relationship between the mean rank of the parent's income percentile rank to their child's income percentile rank is linear. Within this framework, the intergenerational rank association (IRA) is estimated using a least ordinary squares (OLS) regression of the respondent's income percentile ($y_{i,t}$) on their family income percentile rank during their childhood ($y_{i,t-1}$).

$$y_{i,t} = \beta + \lambda y_{i,t-1} + e_i \quad (3.1)$$

This equation produces two measurements of mobility. Following the terminology established by Chetty et al. in their 2014 studies, this regression produces a relative and absolute measurement for each sample.

The relative measurement is captured by λ , the coefficient of a child's parent income percentile rank. As a coefficient, λ measures the relationship between a child's respective income percentile rank among their cohort and their parent's respective household income percentile rank by their cohort. Due to it being relative, this measurement investigates the outcomes of children from low-income families compared to children from high-income families.

In its interpretation, let us say, for example, that $\lambda = 0.10$. This value means that for every percent increase in a parent's lifetime household income percentage, we expect their children's lifetime income percentile to increase by .1%. However, a more straightforward interpretation (since we are working in percentages) is to

discuss our coefficients for a 10% increase in a parent's lifetime income percentile. Using the previous example, we expect a child's lifetime income percentile to increase by 1%. Thus, our relative mobility measure (λ) can be thought of as measuring how much of one's parents' socioeconomic status (measured in household income percentiles) is directly translated to their children simply because their parents started at a certain socioeconomic level. In essence, it captures how one's parent's socioeconomic status raises their child's socioeconomic status, keeping everything else fixed.

The absolute measurement of mobility is captured by β in the equation. β measures the percentile income rank of a child whose parents have no income. When aggregated with many observations, it captures the mean child rank of those from the lowest income percentile families. Thus, β captures an absolute baseline, theoretically implying that its value measures the income percentile rank of children born in their cohort with no previous family socioeconomic status transferred to them.

When viewing these measurements together, absolute mobility can be viewed as more socially important than relative mobility. For instance, an increase in relative mobility (measured by a mathematical decrease in λ) alone shows that parents' household income plays a less significant role but does not tell us much about those on the lower end of the socioeconomic landscape—and also poorer outcomes for the rich is not always socially advantageous (Chetty et al., 2014a).

Absolute mobility provides the other half of the picture, as it captures the movement in the socioeconomic status starting line (measured in income) that every child in a generation receives simply by being born during that time. This measurement is even more essential when comparing across generations. For example, if absolute mobility (β) increases between two generations, and the Pareto

principle holds, it clearly indicates an overall increase in the welfare of society—assuming welfare can be accurately captured exclusively by income. Invoking a typical economic saying, absolute mobility is like the tide, and a rising tide lifts all boats while a falling tide lowers all boats.

Absolute mobility can and has been defined in many different ways. We follow Chetty et al.'s 2014a method of reporting absolute upward mobility, which is the mean rank of a respondent whose parents are at the 25th household income percentile—primarily for ease of comparison. In addition, we also report absolute mobility in terms of the constant, the mean rank of children with parents at the 0th percentile of the income distribution. We term this definition of absolute mobility the baseline mobility for a cohort.

Because we are interested in positing the impact of long-term social and cultural movements on U.S. demographic groups when comparing across generations, most individuals would expect these movements to raise the standard for the entire demographic group, and thus, the baseline should move. Thus, baseline mobility provides another straightforward interpretation yet an important comparison between the timeframes.

Moving forward in this paper, one can interpret the interaction between changing values of λ and β in the following ways. If λ increases and β decreases, that implies an overall decrease in intergenerational socioeconomic mobility between the generations. If λ decreases and β increases, that implies an overall increase in mobility. When the two terms diverge in their signs, it depends on the magnitude. For example, a $-\lambda$ and a $-\beta$ indicate that the influence of parental income has decreased (leading to more mobility) while, at the same time, the baseline has lowered

(everyone starts at a lower socioeconomic level)—and vice versa when the signs are flipped.

A last deviation from Chetty et al. (2019) is that we assume that relative and absolute mobility (λ and β) can vary across our generations. We make this assumption due to analyzing data almost a generation earlier than analyzed by Chetty et al. 2014a and 2014b, which was a time of rapid economic changes. This study believes that U.S. mobility could vary across generations prior to its apparent stabilization in the 1970s, as shown by Chetty et al. (2014b; 2019).

Finally, studies of intergenerational socioeconomic mobility can run into attenuation and life cycle bias issues. As Chetty et al. (2014a) explain, the literature has shown that measuring a respondent's income solely during their initial years in the labor market can understate intergenerational persistence in lifetime income. This fact is due to a body of literature that shows how children with high lifetime incomes have steeper earnings profiles when young (Haider & Solon, 2006). If this issue is not addressed, it can result in life-cycle bias issues. To avoid life cycle bias within our calculations, we follow previous precedent to the greatest degree allowed by our data limitations and measure our respondents' income from their ages of their mid-twenties to their mid-thirties (24 to 34 in the NLSY79 and 23 to 33 for the NLSY97). Attenuation bias arises when a study measures income for either the parent or child in a single year. Thus, we average the income percentile ranks of our respondents over ten years and our respondent's parents over seven years.

Data

This study utilizes the NLSY79 and NLSY97 datasets. These longitudinal surveys followed the cohort members beginning in their teenage or early adult years through adulthood and are continuing today. Both surveys ask questions to the parents

of the cohort members and the cohort members themselves as they age. For the NLSY79 cohort, members' families were 'screened' with questions answered by their families in 1978, and then the cohort was interviewed annually from 1979 to 1994 and has been interviewed biannually since then. For the NLSY97, members' families were again screened in 1998, with the cohort being interviewed annually from 1997 to 2011 and biannually since then. For this analysis, with the desire to focus on those groups traditionally unrepresented in the study of intergenerational socioeconomic mobility, we begin by including all 12,686 respondents in the NLSY79 and 8984 respondents in the NLSY97 prior to observation removal due to missing values.

The NLSY79 follows those born between 1957 to 1965. Thus, in their initial survey date in 1979, the initial survey age ranged from 14 to 22, meaning respondents were in their teenage to young adulthood stages of life. Meanwhile, the NLSY97 follows those born between 1981 and 1985, with this cohort being in their teenage years (aged 12-16) when initially surveyed in 1997. With the two NSLY studies containing many variables, detailed descriptions of our variable definitions and their parameters are included in the Appendix. In the rest of this section, we will touch upon the most significant variables in our study.

Percentile parental income rank for each dataset is calculated in a similar method. For the NLSY79 cohort, the respondent's total previous net family income (NFI) is the base measure for their parent's total income. In the NLSY79, NFI provides a composite income figure from various income values for the household members related to the respondent by blood or marriage in the past calendar year. With it being the total income from the last calendar year, our study includes the NFIs of the respondents for seven years, from 1978 to 1984. We chose these years as they provide the most accurate proxy for lifetime parent income while the youngest

respondents are in their teenage years up and the oldest respondents are closest to their teenage years (the youngest age cohort during this time ages from 13 to 19, while the oldest ages from 21 to 27), and there is comparable data in the NLSY97 survey. We understand that for the oldest NLSY79 cohort, issues arise if there is some potential inclusion of spousal income. However, to include as many observations as possible, this potential source of error remains in our data and could affect the magnitude of our NLSY79 sample mobility estimates.

Similarly, for the NLSY97 cohort, the gross household income (GHHI) in the past year is utilized as the family income variable. In the NLSY97, GHHI is compiled from several questions regarding income sources, both of the respondent and their parent. However, in the earlier rounds of the survey, this variable was primarily parent's income if the respondent resided with them, which was true of most respondents until they reached their mid to late twenties. Following the desire to proxy parents' lifetime income during the participants' teenage years, we take GHHI for the participants' parents from 1997 to 2003. During this timeframe, the youngest of the cohort aged 12 to 18, while the oldest aged 16 to 22. We believe that the GHHI variable accurately captures parent household income within this timeframe, which is supported by this paper's results.

Respondent income was taken over their mid-twenties to mid-thirties to proxy for their lifetime income. To keep our sample data consistent, the youngest in the cohort was utilized as the base in determining the years observed. For the NLSY79 cohort, respondent income was taken from the variable of the total income from wages and salaries received in the past year. With it also recounting the past year's income, we analyzed these variables from 1989 to 1999, when the youngest in the cohort aged 24 to 34 while the oldest aged 32 to 42. Similarly, for the NLSY97

cohort, respondent income was also taken from the total income from wages and salaries received in the past year variable. Again, the years 2008 to 2018 were utilized, with the youngest in the cohort aged from 23 to 33 while the oldest aged from 27 to 37 to keep the age range similar between the two studies. Additionally, our study purposefully does not include the impact of COVID-19 on incomes, as that question is beyond the scope of this study.

Once all the parent and corresponding respondent incomes had been amassed, we had to standardize them into percentiles to calculate IRA, which required more work and data. Typical, more extensive intergenerational socioeconomic mobility studies utilize datasets covering entire, country-wide cohorts. Thus, they can calculate their percentage rank within their sample birth cohort alone (Chetty et al., 2014a; Chetty et al., 2014b; Chetty et al., 2019). With the limitation of our dataset's observations, we drew a heavy amount of income data from the IPUMS Current Population Survey (CPS) to construct representable IRA percentile estimates.

IPUMS CPS is a free microdata source that has harmonized monthly U.S. labor force surveys since 1962. For this study's purposes, we draw on its database of Total Household Income (THHI) and Wage and Salary Income (WSI) by year, including Annual Social and Economic Supplement Weights for each variable. This dataset contains roughly 150,000 to 200,000 household income and 50,000 to 100,000 wage and salary income observations per the years this survey considers for its respondent and respondent parent income data. This survey recognizes the differences between the definitions and subsequent reporting of NFI in the NLSY79, GHHI in the NLSY97, and THHI in the IPUMS CPS. These differences in quantifying measurements of yearly income are a limitation of the datasets used in this study. Even with this limitation, due to the nature of IRA estimates turning specific numbers

into more inclusive bins that minimize some of these specific dollar income variation differences, we believe that our intergenerational rank association measurements are accurate and can be compared between generations—which are further supported by our results.

The percentile ranks for each year are calculated from the corresponding IPSUMS CPS THHI or WSI data year, where we utilize those ten to hundreds of thousands of observations to create income percentile brackets of (1% to 100%). From there, we run the respective NLSY79 and NLSY97 data by year through these already defined income percentile bins and assign each year's respondent income and parent income a respective percent of their cohort's income percentile. Lastly, to proxy for the lifetime income percentile, we average those percentiles across the already indicated years to create the average respondent and parent income percentile ranks.

Next, with this study's focus on investigating the impact of role model effects on women's employment, clearly defining this proxy is important. With our datasets, we decide to proxy the role model effect for female respondents by including an indicator variable as to whether the respondent's mother was employed when they were a teenager—which is present in both surveys. In the NLSY79, we utilize a qualitative question in the initial 1979 survey that indicates if, when the respondent was 14, their mother (including stepmother or their mother figure) worked for pay. In the NLSY97, through a combination of specific household members' characteristics (age, sex, and employment status in 1997), we code a variable that indicates when the respondent was either 12 to 16 if their mother was employed.

A final point about our data samples is that, in the NLSY97 sample, an additional race option was provided: mixed race non-Hispanic. With this study

comparing U.S. group mobility across generations, this category only has 83 total observations. Thus, we decided to omit these observations and racial category from our analysis.

Included in Appendix section 7.1 is a list of the rest of the variables, alongside their definitions. With these definitions, variables, and parameters, our sample sizes decreased due to missing values—a common occurrence in long panel-data studies as respondents drop out. Later in our study, when calculating our corrected models, we lose some observations reported in the summary tables below due to NA for female respondents’ variables in our probit equation. When omitting NAs in parent or respondent income percentile rank, in our NLSY79 sample, our initial 12,686 observations decreased to 10,814 observations; demographically, this sample encompasses 1,405 black women, 935 Hispanic women, 3,140 non-black or Hispanic (white) women, 1,429 black men, 921 Hispanic men, and 2,984 non-black or Hispanic (white) men. The summary statistics of this sample are provided in Table 1 below.

Table 1: NLSY 1979 Sample Summary Statistics					
	Mean	St. Dev.	Min	Median	Max
Average Net Family Income Percentile	40.54	22.9	1.0	37.46	100.0
Average Respondent Income Percentile	44.11	28.3	1.0	44.11	100.0

Total Number of Observations (N) = 10,814

Source: own calculations based on NLSY79 sample.

The initial 8,984 observations in our NLSY97 sample lessened to 7,863 observations with a demographic makeup of 992 black women, 832 Hispanic women, 1,964 non-black or Hispanic (white) women, 992 black men, 856 Hispanic men, and 2,098 non-black or Hispanic (white) men. The 1997 sample’s summary statistics are presented in Table 2 below.

Table 2: NLSY 1997 Sample Summary Statistics					
	Mean	St. Dev.	Min	Median	Max
Average Gross Household Income Percentile	43.19	26.1	1.0	41.0	100.0
Average Respondent Income Percentile	44.0	19.2	1.0	44.0	100.0
<i>Total Number of Observations (N) = 7,863</i>					

Source: own calculations based on NLSY97 sample.

Results and Analysis

Our results section is split into eight primary sections. Section 5.1 compares our overall multigenerational trends to findings in previous literature to establish the representativeness of our samples. Our subsequent sections investigate U.S. trends in intergenerational socioeconomic mobility utilizing our samples, starting broad and ending as specific as our dataset permits. Section 5.2 considers overall mobility trends between the 1957 to 1965 and 1981 to 1985 cohorts. Section 5.3 investigates trends in U.S. mobility by region, section 5.4 looks at trends by sex, and section 5.5 looks at these trends by race. Section 5.6 analyzes trends by sex and race groupings, section 5.7 discusses the corrected models and potential influence of the role model effect, and section 5.8 considers a possible explanatory variable explaining group mobility differences: education.

Findings Compared to Previous Literature

We first compare our results to the previous literature to determine the representativeness of our samples in measuring mobility. With this study's focus on the U.S. population, its usage of parent income measured in terms of household, looking multigenerational, and its emphasis on modeling women and all racial groups, we believe that comparing our results to Chetty et al.'s previous findings in their

2014a, 2014b, and 2019 studies is most appropriate. When viewing the entire U.S. population in their 2014a and 2014b papers (1980 to 1991 U.S. birth cohort), they find a relative mobility of 0.35 when utilizing parent and child-family income definitions and a relative mobility of 0.282 when utilizing parent-family and child earnings income definitions. Their 2019 study found a relative mobility of 0.34 when pooling all their U.S. racial mobility data (1978 - 1983 U.S. birth cohort) using family income definitions. Additionally, in their 2014b study, Chetty et al. found minimal to negligible increases in mobility between the cohorts born between 1971 and 1993.

As Table 3 shows, using our NLSY survey data, we estimate a relative mobility rank of 0.443 for the NLSY79 sample and 0.355 for the NLSY97 sample with our income definitions. With our NLSY97 cohort being born between 1981 and 1985, we feel that comparing Chetty et al.'s findings (0.34, 0.28, 0.35) to our 97 model results (0.355) is most appropriate. As we can see, our relative mobility estimate is almost identical to Chetty et al.'s when defined as both parent and child family income. However, our estimate is slightly larger when defined by parent family to child earnings income—larger by 0.073.

Additionally, our samples' multigenerational trends indicate a decrease in relative mobility and a slight increase in baseline mobility. However, as Figure 2 shows, we find both generations have an almost identical absolute upward mobility of around 37, indicating no change between the generations. When multigenerational trends are compared to Chetty et al.'s 2014b results, our study reports, identically, no change if absolute upward mobility is the primary measure. However, when viewing absolute mobility in terms of the baseline combined with relative mobility, we report a decrease in relative mobility and a slight increase in baseline mobility, indicating a minimal increase in U.S. intergeneration mobility between the generations.

Table 3: NLSY 1979 and 1997 Overall OLS Regression Outputs

	79 Sample	97 Sample
Average Percentile Household Income	0.443*** p = 0.000	0.355*** p = 0.000
Constant	26.134*** p = 0.000	28.679*** p = 0.000
Total Observations	10,814	7,863
R ²	0.129	0.139
Adjusted R ²	0.129	0.139
Residual Std. Error	26.036(df = 6059)	23.077(df = 3059)
F Statistic	794.001*** (df=1;6059)	522.657*** (df=1;3059)
Note:	+ p<0.1; * p<0.05; ** p<0.01; ***p<0.001	

Source: own calculations based on NLSY79 and NLSY97 samples.

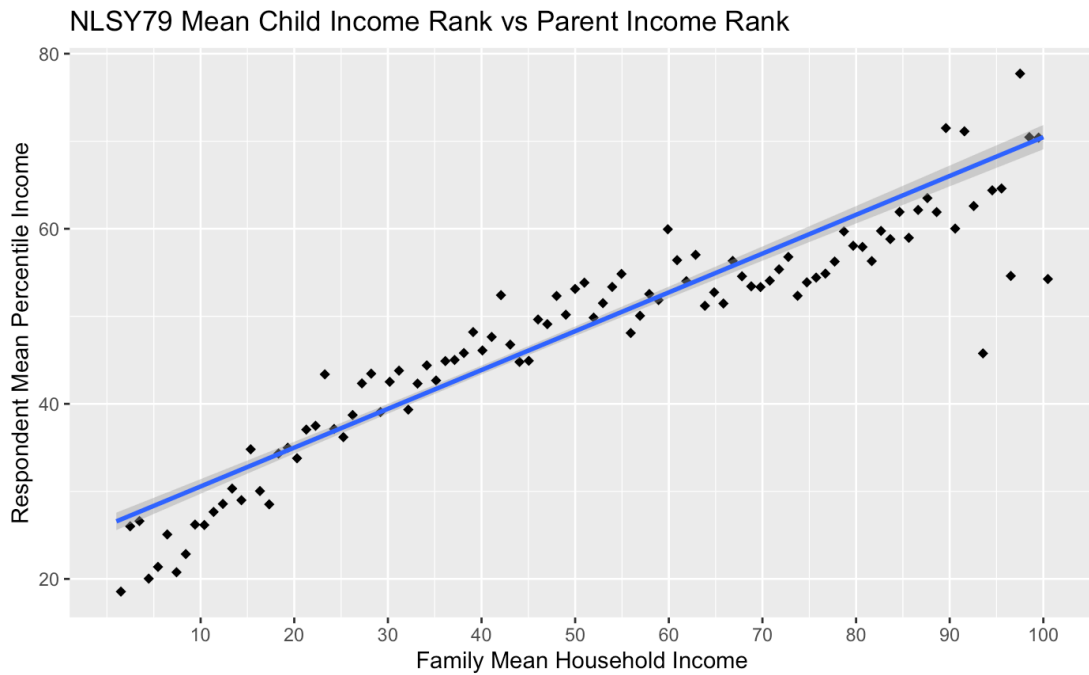
Lastly, we follow Chetty et al.'s precedent of checking our data and variable definitions for linearity. Figures 1.1 and 1.2 below show the binned scatter plot of the mean percentile rank of a respondent's parents (P_i) on the x-axis and the mean percentile rank of the respondent as an adult (R_i) on the y-axis for our NLSY samples. These graphs show $E[R_i | P_i = p]$: the conditional expectation of the respondent's income percentile rank given their parents' household income percentile rank p . The graphs show strong linearity of these conditional expectations for our data. However, our data possibly falters from past compared to past literature when looking at the top income percentiles, primarily due to our sample size and the relative lack of high-income percentile individuals in our samples.

With these assumptions checked, we believe our samples still capture accurate estimates of intergenerational socioeconomic mobility compared to the previous literature. Interestingly, when we use parent-family-to-child earnings income definitions, our overall results are closer to previous literature utilizing family income definitions for the overall U.S. population. However, as later results sections will show, when splitting our samples by sex and race, our estimates are closer to previous

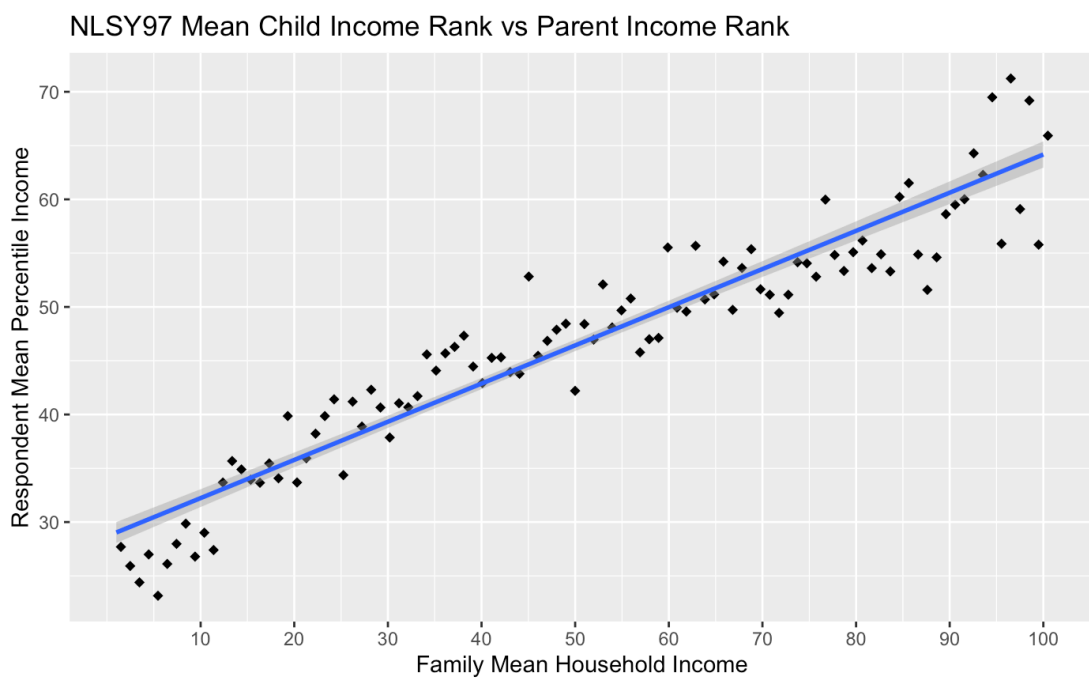
literature estimates of the same groupings when using child earnings and parent-family income definitions.

Figure 1: Sample Linear Relationship of Parent to Respondent Income Percentiles

1.1: *Linear Relationship of Parent to Respondent Percentiles NSLY79 Sample*



1.2: *Linear Relationship of Parent to Respondent Percentiles NSLY97 Sample*

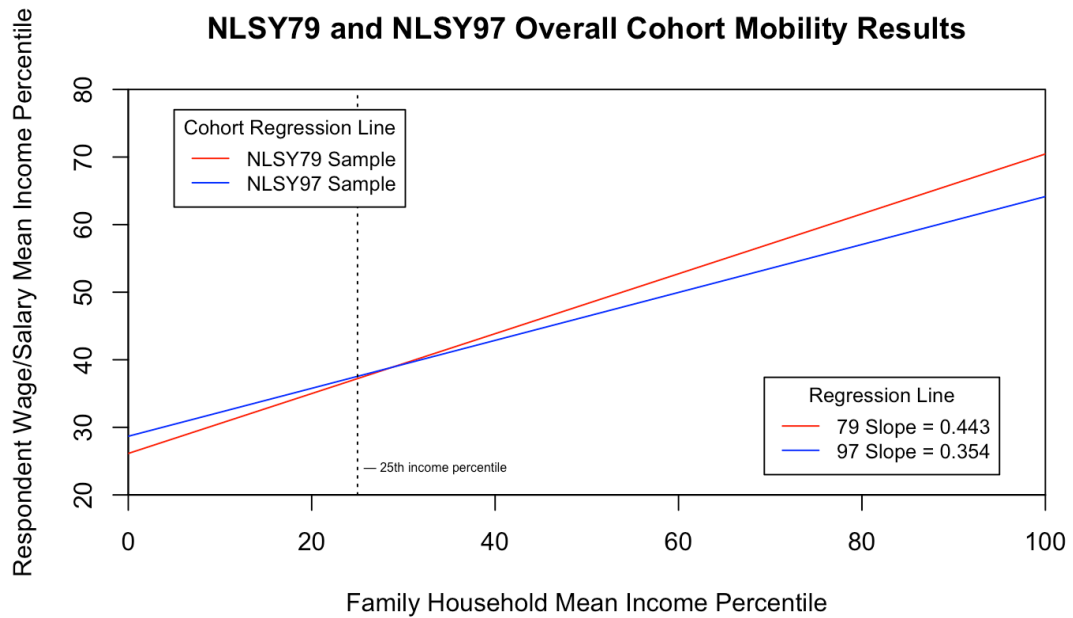


Source: Made in R based on NLSY79 and NLSY97 samples.

US Multigenerational Trends

Comparing the mobility of those born in the late 1950s to early 1960s with those born in the early to mid-1980s, as reported in Table 3 and shown in Figure 2, we report a minimal increase in U.S. socioeconomic mobility depending on the definition. Between the two generations, relative mobility decreased from 0.443 to 0.355, baseline mobility increased from 26.134 to 28.679, and absolute upward mobility stayed the same at about 37.5.

Figure 2: NLSY79 and NLSY97 Overall Cohort Mobility Regressions



Source: Made in R based on NLSY79 and NLSY97 samples.

The decrease in relative mobility is most substantial, as our estimates indicate that the importance of one's parent's household income percentile has decreased by about 20%. Our estimates of baseline mobility show an increase of 26.134 to 28.679. Thus, between this timeframe, the baseline or socioeconomic level of those born into families with the least resources increased by around 10%. Meanwhile, our measure of absolute upward mobility changed by a factor of 0.34, indicating no change. Thus, it depends on the income percentile one is interested in when discussing mobility

changes. However, when taken all together, the decrease in relative mobility alongside the increase in baseline mobility and stagnation of absolute upward mobility indicate a slight increase in overall U.S. mobility at the ends of the U.S. socioeconomic distribution, with no real changes for most Americans between these two generations.

US multigenerational Trends by Region

We now strive to investigate how this overall slight increase in mobility between the generations has been distributed within the U.S. and, more importantly, how these overall trends differ within the United States. We begin, again, by following Chetty et al.'s example and code each respondent in our sample by region. Due to the limits of our data, we can only investigate this topic by the four regions recorded in each study: North East, North Central, West, and South. We calculate measures of absolute (β_r) and relative (λ_r) mobility described earlier through the slope and intercept modeled in Equation 5.1 below, where $R_{i,r}$ indicate the respondent's representative income percentile rank by the region they lived in when a teenager and $P_{i,r}$ similarly indicates the respondent's respective parent's representative income percentile rank.

$$R_{i,r} = \beta_r + \lambda_r P_{i,r} + e_{i,r} \quad (5.1)$$

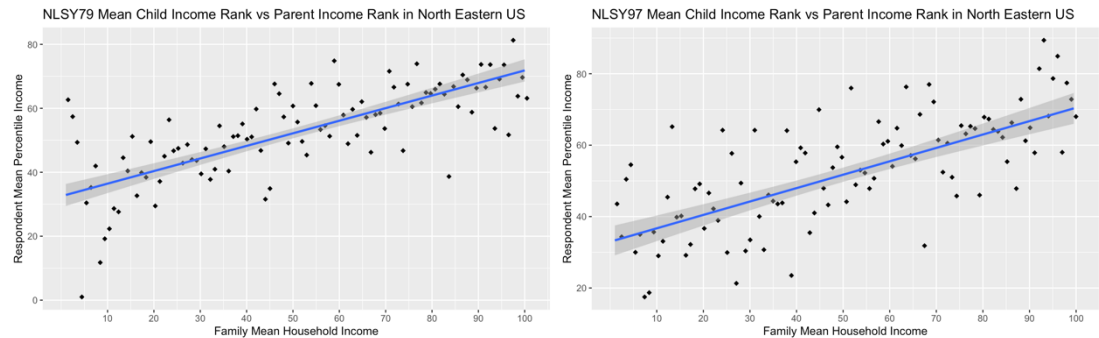
Table 4 below shows the outputs of each regression for the 79 sample, while Table 5 records the outputs for the 97 sample. Due to the abundance of models, we did not graph region differences alongside each other, as the graph becomes extremely cluttered and does not aid in visually interpreting these results.

Compared to our national estimates, it is important to note that when coded by region, the strength of our linear relationships begins to wane. Below in Figure 3, we have provided the binned scatter plots of the mean respondent's income percentile on

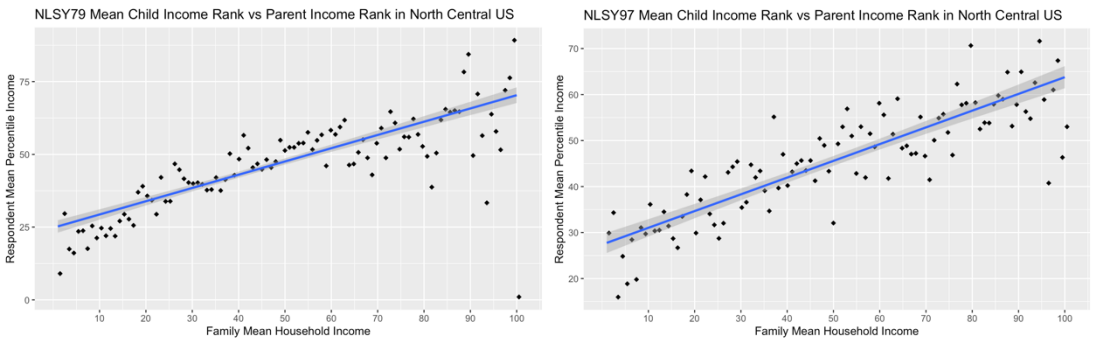
the x-axis and the mean respondent's parents' income percentile on the y-axis for each region group. While all graphs show a linear relationship, they differ in the strength of their linear relationships—specifically, 97 North East and 97 West. These issues are due to the lower number of observations when we split our data into four subsamples, as shown in Tables 4 and 5 below.

Figure 3: Linear Relationships of Parent to Respondent Percentiles by Region

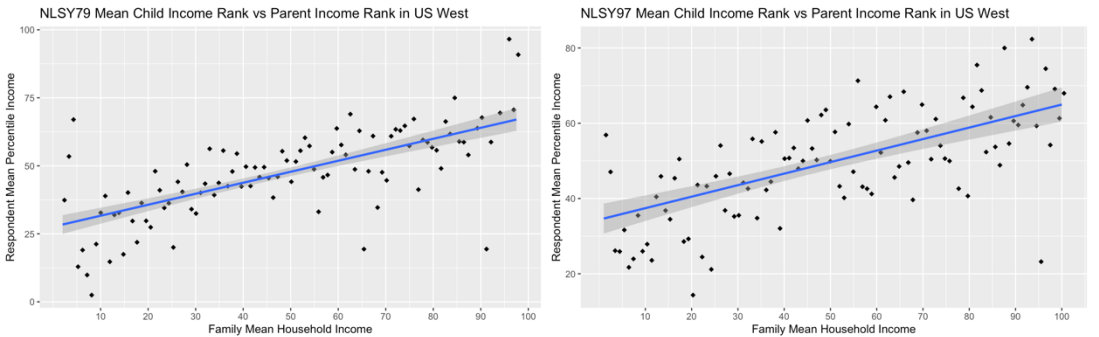
3.1. Linear Relationships for U.S. North East



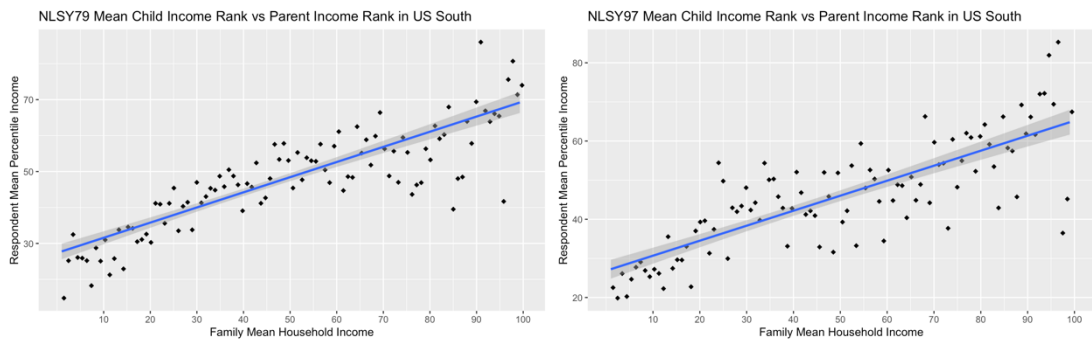
3.2. Linear Relationships for North Central U.S.



3.3. Linear Relationships for U.S. West



3.4. Linear Relationships for U.S. South



Source: Made in R based on NLSY79 and NLSY97 samples.

Table 4: NLSY 1979 Region OLS Regression Outputs

	79 N. East	79 N. Central	79 West	79 South
Average Percentile Household Income	0.393*** p = 0.000	0.455*** p = 0.000	0.404*** p = 0.000	0.421*** p = 0.000
Constant	32.504*** p = 0.000	24.785*** p = 0.000	27.613*** p = 0.000	27.356*** p = 0.000
Total Observations	1,103	2,668	1,137	2,200
R ²	0.122	0.134	0.094	0.123
Adjusted R ²	0.122	0.134	0.093	0.122
Residual Std. Error	26.036(df = 6059)	23.077(df = 3059)	17.045(df=6005)	14.826(df=2,995)
F Statistic	794.001*** (df=1;6059)	522.657*** (df=1;3059)	65.745*** (df=65;6005)	33.111*** (df=65;2,995)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 sample

Table 5: NLSY 1997 Region OLS Regression Outputs

	97 N. East	97 N. Central	97 West	97 South
Average Percentile Household Income	0.375*** p = 0.000	0.364*** p = 0.000	0.306*** p = 0.000	0.383*** p = 0.000
Constant	32.959*** p = 0.000	27.392*** p = 0.000	34.372*** p = 0.000	26.850*** p = 0.000
Total Observations	510	1,791	614	1,153
R ²	0.156	0.139	0.090	0.164
Adjusted R ²	0.154	0.139	0.088	0.163
Inverse Mills Ratio			-65.713*** (7.930)	-39.238*** (6.453)
Residual Std. Error	26.036(df = 6059)	23.077(df = 3059)	17.045(df=6005)	14.826(df=2,995)
F Statistic	794.001*** (df=1;6059)	522.657*** (df=1;3059)	65.745*** (df=65;6005)	33.111*** (df=65;2,995)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY97 sample

Tables 4 and 5 indicate that regions within the U.S. have experienced different trends in mobility. First, The North Eastern U.S. has barely changed, as relative

mobility decreased by .018, baseline mobility for the region increased by 0.455, and absolute upward mobility effectively did not change. Thus, those within the North Eastern region of the U.S. appear to have experienced no changes to their mobility between the two generations. For the North Central region of the U.S., there appears to have been a slight increase in mobility. Relative mobility decreased by 0.091, baseline mobility increased by 2.607, and absolute upward mobility increased by around 0.33. The American West experienced a more significant increase in mobility, with a relative mobility decrease of 0.098, a baseline mobility increase of 6.759, and an absolute upward mobility increase of 4.309.

Lastly, the American South experienced the least mobility increase and potentially a decrease between the two generations. While relative mobility did decrease by 0.038, the south experienced the only decrease in baseline mobility (by a factor of 0.506) and absolute upward mobility (by a factor of 1.5). Thus, when looking regionally, the gains in overall U.S. mobility were experienced almost entirely by the Western U.S. Meanwhile, the North Eastern and North Central regions experienced no significant changes, and the Southern U.S. experienced minimal to negligible decreases in their mobility.

When comparing our NLSY97 sample estimates to Chetty et al. s.' 2014a and 2014b findings, although our estimates are much more generalized, they follow the same trends and estimates as Chetty's. Regarding absolute upward mobility², Chetty et al. find the South to have the lowest³ while the West has the highest, with the North

² Overall, our absolute upward mobility measurements compared to Chetty et al. 2014a for the North East, North Central, and West regions are slightly below when averaged by careful eyeballing for the identified region's overall range of motilities. However, the ordering/significance of our region estimates are similar to Chetty et al., and our relative motilities are consistent when eyeballing and averaging Chetty et al.'s results by region.

³ As an example, in Chetty et al. 2014a, their range for the Southern U.S. absolute upward mobility is often between 26 and 45 for most areas. When averaged, one gets 35.5, which is approximately our absolute upward mobility estimate for the South of 36.

East being in the middle and the North Central region being the most diverse. Our relative mobility estimates are similar in ranking, trend, and when averaged, with the Southern U.S. having the lowest relative mobility and the Western U.S. having the lowest relative mobility. Regarding multigenerational mobility trends by region, Chetty et al. 2014b find no significant changes over time for any region. Our results for the North East, North Central, and Southern U.S. regions agree with these previous results. At the same time, our estimates for the American West differ, reporting a slight increase in overall mobility.

US Multigenerational Trends by Sex

We continue our investigation of U.S. multigenerational trends by looking at the mobility of our samples based on sex. Similar to the previous sections, we split our NLSY79 and NLSY97 samples by sex and estimate measures of absolute (β_s) and relative (λ_s) mobility modeled in Equation 5.2 below, where $R_{i,s}$ indicates the respondent's representative income percentile rank by their assigned sex at birth and P_i indicates the respondent's respective parent's income percentile rank.

$$R_{i,s} = \beta_s + \lambda_s P_i + e_{i,s} \quad (5.2)$$

Similar to our previous analysis, we report the degree of the linear relationship between these sample's parent and child mean income percentiles in Figure 3. Like the overall samples, each sample shows a high degree of a linear relationship. These binned scatter plots are included in this study's Appendix, section 7.2.

However, when investigating differences by sex, we deviate slightly from the previous literature. In addition to calculating baseline, relative, and absolute upward mobility traditionally as described above in Equation 5.2, since we utilize historical data that is not wholly representative of the entire U.S. birth cohort, we also construct probit equations for each female respondent's likelihood of employment to institute a

two-step Heckman correction. This correction is chosen to potentially account for the issues that arise (such as sample selection bias and reverse causality) when modeling women's employment and income data. For this correction, we estimate a probit model for women's employment ($E_{i,w}$) for each NLSY79 and NLSY97 dataset of the form

$$E_{i,w} = \beta_w + \lambda_w P_i + e_{i,w} \quad (5.3)$$

We add to this probit equation vector N . Vector N contains all the variables unique to our probit model, which we include only in the first step of our 2SLS. Following the established literature when modeling women's employment, the primary variables unique to our probit model include the number of children the respondents have and their marital status during their viewed year of employment, and those two interacted with each other (Heckman, 1979; Mulligan & Rubenstein, 2008). Additionally, since we are modeling women's employment as far back as the 1970s, we include if the respondent grew up Catholic to proxy for attitudes towards birth control and contraceptives and their region of residence during their teenage years. Lastly, with our interest in the impact of role model effects on promoting upward women's mobility, our final variable within vector N identifies whether the respondent's mother worked during their teenage years—which serves as our role model effect proxy.

Thus,

$$\text{Vector } N = (\theta_i X_i \text{ moth_emp} + \gamma_i X_i \text{ num_child} + \tau_i X_i \text{ married} + \omega_i X_i \text{ catholic} + \phi_i X_i \text{ num_child*married} + \delta_i X_i \text{ Region}). \quad (5.4)$$

And our first step probit equation can be written as,

$$E_{i,w} = \beta_w + \lambda_w P_i + \alpha_w N_i + e_{i,w} \quad (5.5)$$

After running Equation 5.5 above, we can use the estimated positive values and estimated errors to calculate the inverse mills ratio for every respondent. In technical

terms, since we believe our data might be distributed in a truncated version of the normal distribution due to women self-selecting out of the workforce, we calculate the inverse mills ratio traditionally laid out by Heckman in his 1979 paper as the ratio of our samples PDF to its complementary CDF—assuming our data is distributed normally. This explanation is represented mathematically as

$$E[X | X > \alpha] = \mu + \sigma \frac{\theta(\frac{\alpha - \mu}{\sigma})}{1 - \Phi(\frac{\alpha - \mu}{\sigma})}, \quad E[X | X < \alpha] = \mu - \sigma \frac{\theta(\frac{\alpha - \mu}{\sigma})}{\Phi(\frac{\alpha - \mu}{\sigma})} \quad (5.6)$$

Explaining this math, we propose that there are censored observations in our data.

Due to modeling women's income, which heavily relies on their employment (especially women's employment in the 1970s), we might expect an issue where women are self-selecting to leave the workforce. The most common reason for this self-selection out of the workforce is due to childcare costs; if a family is not making enough income to afford external childcare, it will pressure one of the parents to drop out of the workforce to care for the children. In American society, this choice has traditionally been the burden of the woman of the household. Thus, women would purposefully leave the workforce to care for their children and report no income.

This choice creates a significant issue when attempting to model our women respondent's income from our samples. Our data is missing information and is not truly randomly distributed; our dependent variable (the women's respondent's income) is censored. This self-selection causes more of our modeled respondents' women's income to be observed at zero or if they have no income. If we run our model without this data correction, it will produce biased estimates in our regression skewed toward zero, which is not true for the actual population—invalidating our results.

We first estimate this parameter using a probit model to correct this issue, which assumes that our data and error terms must follow a normal, non-censored distribution. From these estimated results, which include the corrected expected

positive values of women's employment that tend toward zero in our data, we calculate the inverse mills ratio (IMR). The IMR is an estimator we can use to predict the likelihood that positive women respondents' incomes show up as zeros in our sample data. Once the IMR is calculated for each respondent, it is added as a final independent variable into our second step OLS, which now means we have corrected this sample selection bias in our model—interpreting the significance of the IMR in our OLS output tells us if this correction is necessary for our data. Lastly, the second step of the two-step Heckman correction OLS model is the same as Equation 5.2 with the inverse mills ratio included.

With methodology established, Table 6 below shows both regression outputs for the models utilizing women's specific data—with and without the correction from Equation 5.5—while Table 7 is the model for men solely utilizing Equation 5.2.

Figure 4 shows the regression lines of all three models plotted on the same graph.

Table 6: NLSY 1979 and 1997 Women OLS Regression Outputs

	79 Sample	97 Sample	79 (w/ correction)	97 (w/correction)
Average Percentile Household Income	0.366*** p = 0.000	0.346*** p = 0.000	0.204*** p = 0.000	0.273*** p = 0.000
Constant	21.366*** p = 0.000	24.506*** p = 0.000	60.012*** p = 0.000	38.730*** p = 0.000
Total Observations	5,480	3,882	3,957	2,679
R ²	0.100	0.145	0.170	0.144
Adjusted R ²	0.100	0.145	0.169	0.143
Inverse Mills Ratio			-57.619*** (6.168)	-12.365*** (5.116)
Residual Std. Error	25.045(df = 5478)	21.802(df = 3880)		
F Statistic	607.484*** (df=1;5478)	658.395*** (df=1;3880)		

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

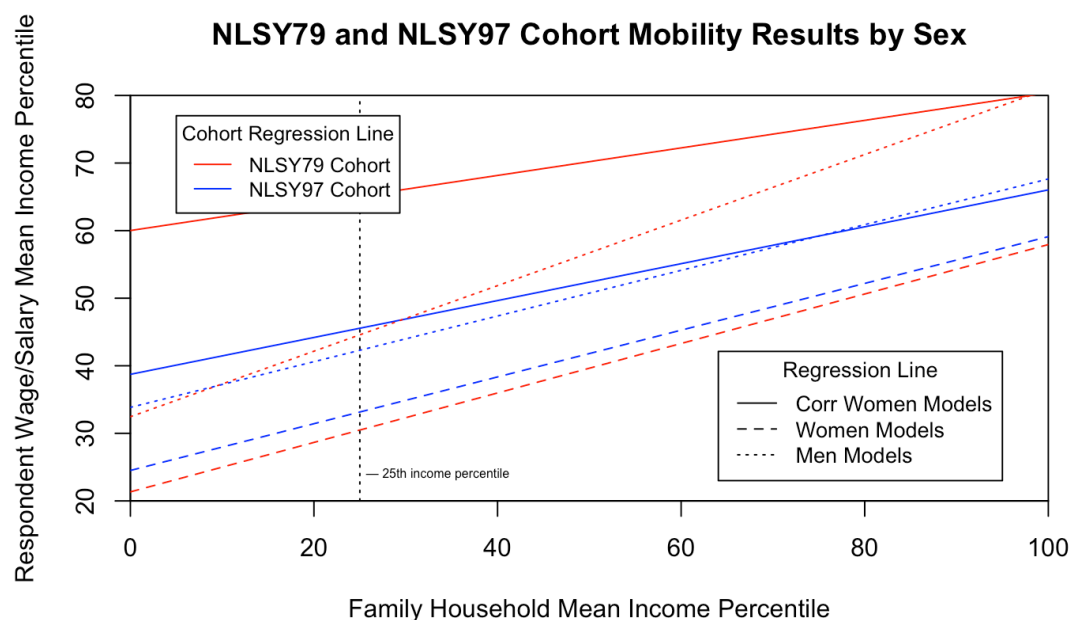
Table 7: NLSY 1979 and 1997 Men OLS Regression Outputs

	79 Sample	97 Sample
Average Percentile Household Income	0.485*** p = 0.000	0.338*** p = 0.000
Constant	32.456*** p = 0.000	33.843*** p = 0.000
Total Observations	5,334	7,863
R ²	0.164	0.125
Adjusted R ²	0.164	0.125
Residual Std. Error	25.192(df = 5332)	23.445(df = 3979)
F Statistic	1,043.924*** (df=1;5332)	569.003*** (df=1;3979)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Figure 4: NLSY79 and NLSY97 Cohort Mobility Regressions by Sex



Source: Made in R based on NLSY79 and NLSY97 samples.

Tables 6 and 7 and Figure 4 again paint the picture of a negligible increase in U.S. mobility over these two generations. When looking at the non-corrected models, U.S. women's relative mobility minimally decreased by 0.02, their baseline mobility increased by 3.14, and their absolute upward mobility increased by 2.64. These estimates suggest that their baseline and absolute upward mobility minimally rose but that their parent's socioeconomic status has played the same role in their income

percentile distribution over both generations. Compared to the previous literature, our relative mobility estimate for the non-corrected U.S. women's model (0.346) is above Chetty et al.'s finding when utilizing parent family and female child individual earnings definitions, around 0.1 more than their estimated 0.249.

When interpreting the corrected model, we can see that the inverse mill's ratios are incredibly significant, indicating that our samples suffer from a truncated distribution and are essential to include. Interestingly, when the correction is added, we see a decrease in U.S. women's relative mobility—though extremely minimal. However, comparing our corrected model for the NLSY97 women sample to the previous literature, we observe that it is more accurate. Chetty et al. (2014a) find a relative mobility of 0.249 when female children's earnings income rank with parent-family income rank definitions are utilized, only around 0.024 below our corrected model estimate (0.273).

This finding, alongside the significance of the IMR in our corrected models, suggests that our models with the correction are more accurate than those without the correction—in terms of estimating relative mobility. Thus, this study proceeds with the finding that the relative mobility of U.S. women increased by 0.07 between these generations, which is a negligible effect. Thus, when taking both models together, we find that (like previous studies) U.S. women's mobility likely did not change, except for a possible slight increase in their baseline mobility, although to a degree, we cannot accurately report. We cannot accurately interpret any corrected model's baseline or absolute upward mobility measurement due to, as stated in the methodology section, adding IMR as an independent variable alongside family household income percentile rank greatly affects the constant (β)—from which both baseline and absolute upward mobility are derived.

Table 7 indicates that U.S. men also experienced a negligible increase in mobility. Their relative mobility measures decreased by 0.147 (from 0.485 to 0.338), their baseline mobility increased by 1.387, and their absolute upward mobility decreased by around 2.3. Previous literature (Chetty et al., 2014a) found a relative mobility for U.S. men of 0.313 when utilizing individual earnings percentile rank and parent family income percentile rank, only 0.025 below our estimate of 0.338. Thus, our sample estimates suggest that men’s intergenerational socioeconomic mobility has negligibly increased between these U.S. generations—supporting past findings.

US Multigenerational Trends by Race

Next, we investigate the overall increase in U.S. mobility between the generations by U.S. races. Our methodology is the same as our previous queries. We split our samples by their race and estimate relative, baseline, and absolute upward mobility. Again, due to our data restrictions, we can only investigate this topic by three races recorded in both studies: Black, Hispanic, and Other. For both NLSY surveys, although the final race category is ‘other,’ most of that sample is perceived as racially white. Thus, we view the ‘other’ race category as essentially white Americans. While the NLSY97 data also included mixed race as an option, it is tiny and was not present in the NLSY79. Thus, we decided to omit those observations. Again, we measure absolute mobility (β_{ra}) and relative mobility (λ_{ra}) identically to our previous sections modeled by Equation 5.7 below.

$$R_{i,ra} = \beta_{ra} + \lambda_{ra} P_i + e_i \quad (5.7)$$

Our checks for the strength of the linear relationship between a respondent’s income percentile rank and their parent’s income percentile rank for our subdivided race samples are included in the Appendix, section 7.2. Table 8 below records these regression outputs, while Figure 5 shows all their regression lines on a single graph.

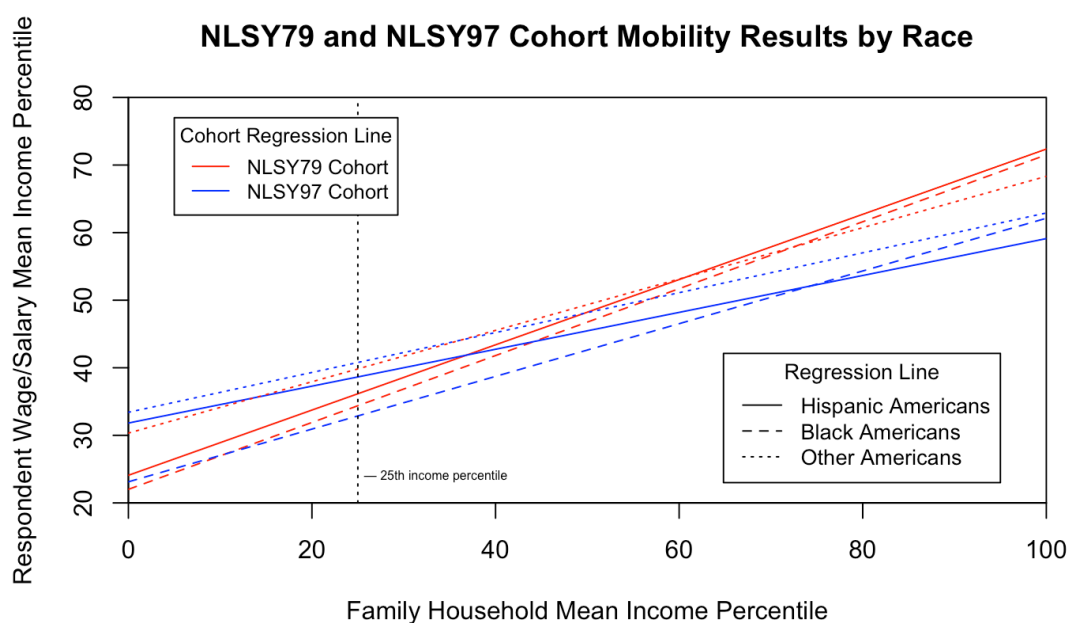
Table 8: NLSY 1979 and 1997 Race OLS Regression Outputs

	79 Black	97 Black	79 Hispanic	97 Hispanic	79 Other	97 Other
Average Percentile Household Income	0.495*** p = 0.000	0.390*** p = 0.000	0.483*** p = 0.000	0.310*** p = 0.000	0.380*** p = 0.000	0.295*** p = 0.000
Constant	22.007*** p = 0.000	23.123*** p = 0.000	24.086*** p = 0.000	31.819*** p = 0.000	30.332*** p = 0.000	33.408*** p = 0.000
Total Observations	2,834	2,041	1,856	1,688	6,124	4,062
R ²	0.147	0.163	0.147	0.095	0.092	0.087
Adjusted R ²	0.146	0.162	0.146	0.095	0.092	0.087
Residual Std. Error	24.374(df = 2832)	21.191(df = 2039)	25.147(df = 1854)	22.196(df = 1686)	27.491(df = 6122)	24.000(df = 4060)
F Statistic	486.232*** (df=1;2832)	396.226*** (df=1;2039)	319.122*** (df=1;1854)	177.485*** (df=1;1686)	618.617*** (df=1;6122)	386.307*** (df=1;4060)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Figure 5: NLSY79 and NLSY97 Cohort Mobility Regressions by Race



Source: Made in R based on NLSY79 and NLSY97 samples.

Table 8 shows that for all three races considered in this study, relative mobility decreased and absolute mobility increased between the two generations to varying degrees. Black Americans experienced the smallest increases in mobility, with a relative mobility decrease of 0.105, an increase in baseline mobility of 1.116, and a decrease in absolute mobility of about 1.50. Hispanic Americans saw the most significant increase in mobility between the generations, experiencing a relative

mobility decrease of 0.173, a baseline mobility increase of 7.733, and an absolute upward mobility increase of around 4.10. Lastly, other (white) Americans were in the middle, as they saw a decrease in their relative mobility of 0.085, an increase in their baseline mobility of 3.076, and an increase in their absolute upward mobility of about 0.90.

To quickly compare our estimates to previous literature, Chetty et al. 2019 find a relative mobility of 0.32 for white Americans, 0.28 for black Americans, and 0.26 for Hispanic Americans. Our estimates of relative mobility (NLSY97 sample) for white and Hispanic Americans are almost identical (differing by 0.025 and 0.05, respectively), while our relative mobility estimate for black Americans is slightly larger (0.11 greater). Additionally, our baseline estimates for the three races are slightly below Chetty et al.'s 2019 estimates, differing by 3.5 for white Americans, 4.321 for Hispanic Americans, and 2.307 for black Americans.

Returning to our estimates, taken together, these estimates suggest that Hispanic Americans saw overall increases in their mobility, other (white) Americans potentially experienced a minimal increase in mobility, and black Americans experienced a negligible increase or decrease. A more concerning picture emerges when interpreting these changes between races in the context of multigenerational movement. While Hispanic Americans have greatly moved toward similar socioeconomic levels as white Americans during these decades, there is still immense disparity between White Americans and Black Americans. Specifically, between the generations, the gap between White Americans and black Americans has increased. The differences between their relative mobilities have negligibly changed (0.115 in 1979 and 0.095 in 1997), but the difference in change between Black and White Americans' absolute mobility measurements is more pressing.

In the U.S. generation born in the late 1950s to early 1960s (1979 sample), the difference between Black and White Americans' socioeconomic baseline was 8.325 percentile ranks. Chetty et al. (2019) also find a disparity between black and white Americans regarding absolute income measurements. In their 2019 study, Chetty et al. found a consistent gap of around 5 percentile ranks regardless of parent income percentile rank—when the child's individual income definition is used. We estimate a larger difference of 10.3 percentile income ranks at the baseline. However, we do not find it to be consistent, as it slowly converges as one moves up the socioeconomic ladder (ending at a difference of almost 1 percentile income rank at the 100% parent household income bracket). Thus, there is a similar 5% income percentile rank gap for the middle of our sample socioeconomic distribution, but it diverges near the extremes of the distribution. This difference is again attributed to the smaller number of observations in our samples at the extremes of income percentile distribution.

However, for those born in the early to mid-1980s, this same racial, socioeconomic baseline gap had increased to 10.285 by almost two percentile income ranks. For absolute upward mobility, a similar pattern emerges as the gap between the 25th percentile of white and black Americans in the NLSY79 sample was 5.52, while it increased to 7.95 in the NLSY97 sample. This finding is a—albeit a minimal—divergence from Chetty et al.'s 2019 findings on the consistency of the black-white American absolute income disparity.

Thus, our estimates suggest that while Hispanic Americans have made significant strides in socioeconomic mobility, closing the gap between themselves and other (white) Americans, the gap in socioeconomic mobility between black and other (white) Americans have slightly increased between these two generations. If our estimates are accurate, they suggest that the intergenerational mobility gap between

black-white Americans increased between the late 1950s to early 1960s and the early 1980s before stabilizing to a difference around the early teens in the 1970s and 80s (based on Chetty et al.'s 2019 findings). Our estimates suggest that, beyond remaining, the black-white absolute mobility divide could narrowly increase over the coming generations.

U.S. Multigenerational Trends by Race and Sex

Our final attempt to break down the previous understandings of the overall U.S. mobility trends is to investigate the varying levels of mobility by racial sex group. Like our past analyses, we split our samples by race (black, Hispanic, and other) and sex, again estimating relative, baseline, and absolute upward mobility. As before, we measure absolute mobility ($\beta_{ra,s}$) and relative mobility ($\lambda_{ra,s}$) modeled by Equation 5.8 below.

$$R_{i,ra,s} = \beta_{ra,s} + \lambda_{ra,s} P_i + e_i \quad (5.8)$$

Additionally, for the racial women subsamples, we also estimated a corrected model identical to the one employed and explained in section 5.5; however, this time, the probit estimation is also by race and sex sample subgroups. The second step of these corrected 2SLS models is identical to Equation 5.8 above.

The strength of the linear relationship between a respondent's income percentile rank and their parent's income percentile rank for our racial sex samples is in the Appendix: section 7.2. To briefly note, as this is subdividing our samples into six groups, we are again approaching smaller numbers of observations for each regression. This fact lessens the strength of our linearity—although figure 8 still shows linear relationships for each sex/race group. Table 9 includes the results for Black Americans by sex for each sample, table 10 houses the results for Hispanic Americans, and Table 11 pertains to other (white) Americans. Figure 9 compares their

regression lines by race and sex, including both models. Figure 9.1 shows all the non-corrected women's regression models by race on the same graphs; Figure 9.2 does the same for the corrected women's models by race, and lastly, Figure 9.3 shows all the men's models by race.

Table 9: NLSY 1979 and 1997 Black Women and Men OLS Regression Outputs

	79 BW	97 BW	79 BM	97 BM	79 BW (Corr)	97 BW (Corr)
Average Percentile Household Income	0.531*** p = 0.000	0.406*** p = 0.000	0.433*** p = 0.000	0.371*** p = 0.000	0.286*** p = 0.000	0.366*** p = 0.000
Constant	17.235*** p = 0.000	21.958*** p = 0.000	27.647*** p = 0.000	24.461*** p = 0.000	54.446*** p = 0.000	26.433*** p = 0.000
Total Observations	1,405	1,049	1,429	992	1,208	763
R ²	0.185	0.196	0.107	0.134	0.245	0.190
Adjusted R ²	0.185	0.196	0.107	0.133	0.244	0.187
Inverse Mills Ratio					-50.957*** (10.525)	0.757 (6.642)
Residual Std. Error	22.915(df = 1403)	19.516(df = 1047)	25.186(df = 1427)	22.818(df = 990)		
F Statistic	318.932*** (df=1;1403)	254.770*** (df=1;1047)	171.788*** (df=1;1427)	152.914*** (df=1;990)		

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Table 10: NLSY 1979 and 1997 Hispanic Women and Men OLS Regression Outputs

	79 HW	97 HW	79 HM	97 HM	79 HW (Corr)	97 HW (Corr)
Average Percentile Household Income	0.458*** p = 0.000	0.319*** p = 0.000	0.471*** p = 0.000	0.260*** p = 0.000	0.266*** p = 0.000	0.126*** p = 0.000
Constant	17.611*** p = 0.000	26.286*** p = 0.000	32.090*** p = 0.000	38.835*** p = 0.000	51.450*** p = 0.000	51.352*** p = 0.000
Total Observations	935	832	921	856	750	585
R ²	0.146	0.103	0.150	0.073	0.187	0.085
Adjusted R ²	0.145	0.102	0.149	0.072	0.184	0.080
Inverse Mills Ratio					-41.995*** (9.425)	-25.107* (12.327)
Residual Std. Error	23.823(df = 933)	20.768(df = 830)	24.221(df = 919)	22.347(df = 854)		
F Statistic	159.201*** (df=1;933)	96.153*** (df=1;830)	162.665*** (df=1;919)	67.351*** (df=1;854)		

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Table 11: NLSY 1979 and 1997 Other Women and Men OLS Regression Outputs

	79 OW	97 OW	79 OM	97 OM	79 OW (Corr)	97 OW (Corr)
Average Percentile Household Income	0.290*** p = 0.000	0.322*** p = 0.000	0.422*** p = 0.000	0.238*** p = 0.000	0.220*** p = 0.000	0.217*** p = 0.000
Constant	24.405*** p = 0.000	25.792*** p = 0.000	38.847*** p = 0.000	42.340*** p = 0.000	50.099*** p = 0.000	46.941*** p = 0.000
Total Observations	3,140	1,964	2,984	2,098	1,999	1,307
R ²	0.060	0.109	0.134	0.061	0.093	0.104
Adjusted R ²	0.060	0.108	0.134	0.061	0.092	0.102
Inverse Mills Ratio					-38.866*** (6.211)	-19.822** (6.937)
Residual Std. Error	26.108(df = 3138)	23.339(df = 1962)	24.804(df = 2982)	23.065(df = 2096)		
F Statistic	199.638*** (df=1;3138)	238.870*** (df=1;1962)	463.068*** (df=1;2982)	137.119*** (df=1;2096)		

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Figure 6: NLSY79 and NLSY97 Cohort Mobility Regressions by Race and Sex

Figure 6.1: Non-Corrected Models for U.S. Women Racial Groups

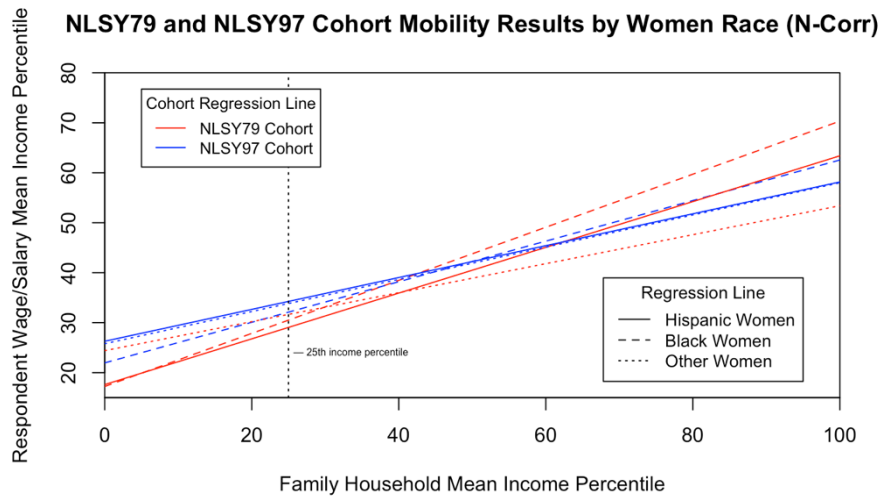


Figure 6.2: Corrected Models for U.S. Women Racial Groups

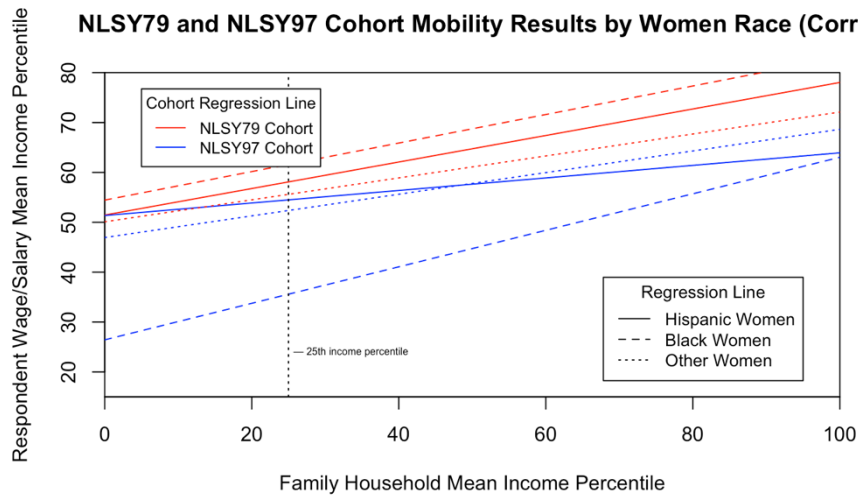
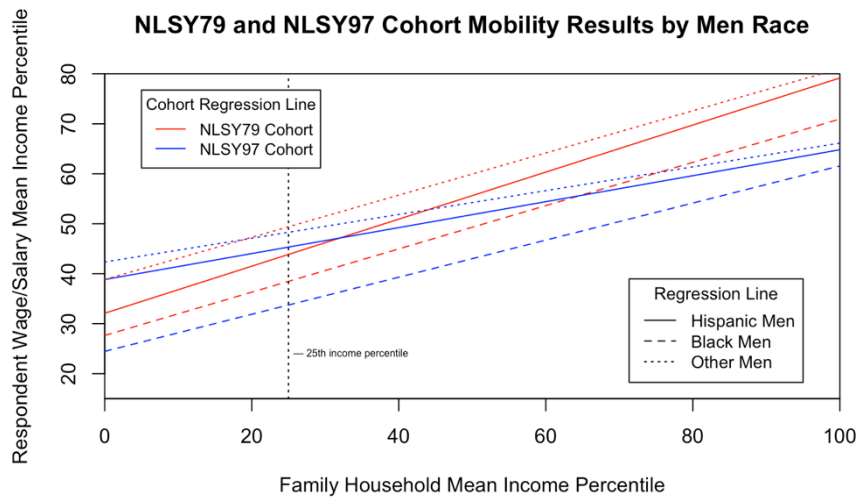


Figure 6.3: Models for U.S. Men Racial Groups



Source: Made in R based on NLSY79 and NLSY97 samples.

When we look at multigenerational mobility by U.S. racial and sex groups, differences emerge. First, as seen when looking at simply racial multigenerational mobility in section 5.6, Hispanic women and Hispanic men have made significant strides in relative and absolute mobility. In the non-corrected model, Hispanic women's relative mobility decreased by 0.139, their baseline mobility increased by 8.675, and their absolute upward mobility increased by 5.2. Hispanic men's relative mobility decreased by 0.211, their baseline mobility increased by 6.745, and their absolute upward mobility by 1.5. The Heckman corrected models for Hispanic

women also reflect a decrease of 0.135 relative mobility, supporting the trends of the non-corrected models. These estimates show that both Hispanic men and women experienced increased mobility between generations.

Other (white) American men and women are less consistent in their mobility patterns. Other women's relative mobility increased negligibly (by 0.032), their baseline mobility barely increased by 1.387, and their absolute upward mobility increased by 2.2. The corrected models for other women also show a slight increase in relative mobility—by a factor of 0.013. During this same time, other (white) men made slightly larger strides in mobility, as their relative mobility decreased by a factor of 0.184, their baseline mobility increased by 3.493, and their absolute upward mobility decreased by 1.1. Thus, these estimates suggest that other (white) women and men's mobility negligibly changed between these two generations.

Lastly, table 8 suggests that black women's mobility seems to have barely changed while black men's mobility has decreased over the two generations. Black women's relative mobility decreased by 0.125 in their non-corrected models while increasing by 0.051 in the corrected models. Their baseline mobility increased by 4.723, and their absolute upward mobility increased by 1.6. Suppose the corrected models are used as the primary interpretation of relative mobility. In that case, their mobility appears to have not changed, and the two models taken together potentially suggest a slight increase from the absolute mobility measures. However, black men's relative mobility decreased by 0.062 alongside a decrease in their baseline mobility of 3.186 and absolute upward mobility of 4.7. Thus, when these results are compared to our previous individual investigations by race and sex separately, they suggest a few things.

1. There were no substantial changes in mobility for other (white) men or women.
2. The increase in mobility for Hispanic Americans was experienced by both Hispanic men and women.
3. The stagnation of overall black American mobility seems to be explained by an equalization of the drop in black male mobility and a slight increase in black female mobility.
4. By sex, the possible minimal increase for American women is almost wholly driven by Hispanics. In contrast, black men's mobility decreases have primarily offset the increases for Hispanic men.

When our findings for the differences between black and white Americans by sex (for the NLSY97 sample) are compared to Chetty et al.'s 2019 study, we find similar patterns, although different magnitudes. Chetty et al. (2019) find that black men are consistently 10 income percentile ranks below white men, while black women are consistently 1 income percentile rank higher than white women. These estimates are found when the child's individual income and parent's household income definitions are utilized. Our estimates show an absolute mobility difference between black and other (white) men of almost 18 at the 0th parent income percentile rank, 14.5 at the 25th parent income percentile rank, 11.23 at the 50th parent income percentile rank, 7.9 at the 75th parent income percentile rank, and 4.6 at the 100th parent income percentile rank. Thus, we report a greater degree of the male black-to-white mobility gap near the lower socioeconomic distribution and variation as one climbs the socioeconomic ladder.

When comparing black and white women by individual income and parent household income percentile definitions, we report a gap of 3.8 in the baseline

mobility (0th parent income percentile rank)—where white women are above black women—a difference of 1.73 at the 25th parent income percentile rank, a difference of 0.37 at the 50th parent income percentile rank—where black women have overtaken white women—a difference of 2.5 at the 75th parent income percentile rank, and a difference of 4.5 at the 100th parent income percentile rank.

Thus, although our magnitudes between U.S. black and white respondents by sex differ in our NLSY97 sample, they include Chetty et al.'s 2019 findings around the 50th to 75th parent income percentile bracket for each group—the middle of our samples. Additionally, our sample follows similar trends and numbers (the large gap between white-black men and the slight increase black women have over white women) but misses the linearity and consistency of these estimates they found in their results. As stated in previous sections, these differences result from our limited observations, especially in its representativeness of the U.S. socioeconomic spread at the extremes (low and high-income families and respondents).

Corrected Models and Role Model Effects

This section investigates the intergenerational impact of role model effects and other predictor variables within our first step correction probit equations. We begin by investigating the impact of role-model effects on all U.S. women. Table 12 below shows the first and second steps for our overall U.S. women's two-step corrected models initially discussed in section 5.4. The respondent's region is all measured against the North Central Region of the United States, and marriage status is measured against not being married.

Before interpreting Table 12 below, it is crucial to understand that the numbers under the probit model columns are not the coefficients of the probit equations. Instead, They are the at mean marginal effects of those variables—except

for the constants, whose coefficients of the probit equation reported from the Heckman selection model, and are thus meaningless. When interpreting these results, they are all percentages (not multiplied by 100) of how much that variable increases the likelihood that the female respondent will be employed. For example, if a number within the probit model were 0.15 and were statistically significant, we would say that if a female respondent had that factor, they were 15% more likely to be employed in adulthood.

Table 12: NLSY 1979 and 1997 Women Probit and OLS Regression Outputs

	79 Probit	79 OLS	97 Probit	97 OLS
Average Percentile Household Income	0.001 ^{***} p = 0.000	0.204 ^{***} p = 0.000	0.002 ^{***} p = 0.000	0.273 ^{***} p = 0.000
Mother Worked For Pay	0.017 p = 0.247		0.077 ^{***} p = 0.000	
Respondent Married	-0.013 p = 0.631		0.129 ^{***} p = 0.000	
Respondent is Catholic	-0.015 p = 0.360		0.026 p = 0.210	
Respondent Number Of Children	-0.058 ^{***} p = 0.000		0.012 p = 0.282	
Children/Married Interacted	0.007 p = 0.440		-0.068 ^{***} p = 0.000	
Respondent Region North East	-0.046 ⁺ p = 0.067		-0.019 p = 0.570	
Respondent Region South	-0.023 p = 0.237		-0.003 p = 0.897	
Respondent Region West	-0.061 ^{**} p = 0.009		-0.076 ^{**} p = 0.009	
Constant	0.839 ^{***} p = 0.000	60.012 ^{***} p = 0.000	0.133 p = 0.166	38.730 ^{***} p = 0.000
Total Observations	3,967	3,957	2,679	2,679
R ²	0.170	0.170	0.144	0.144
Adjusted R ²	0.169	0.169	0.143	0.143
Rho	-1.259	-1.259	-0.570	-0.570
Inverse Mills Ratio	-57.619 ^{***} (6.168)	-57.619 ^{***} (6.168)	-12.365 ^{***} (5.116)	-12.365 ^{***} (5.116)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

As we can see, between the two generations, the impact of the role model effect (a female respondent's mother working outside the house when they were a teenager) on their employment likelihood became statistically significant. For U.S. women born in the 1957 to 1965 cohort, if their mother worked, it was not a significant factor in increasing their likelihood of employment. However, for the 1981 to 1985 cohort, if a female respondent's mother worked, it became statistically significant and increased their chances of employment by almost 8%. Thus, these estimates indicate that the role-model effect for U.S. women grew greatly between these generations, leading more women to become employed. However, when this fact is analyzed alongside the second step, OLS outcomes for women discussed in section 5.4, we can see that even though the role-model effect grew significantly, it did not impact U.S. women's mobility to any significant degree.

Our next line of inquiry is to investigate if a specific U.S. women's racial group has driven this growth of the role-model effect. Table 12 below shows the 1979 and 1997 sample probit model marginal effects for black, Hispanic, and other (white) women taken from the corrected models initially presented in Tables 9, 10, and 11 in section 5.6.

Table 12 suggests that the overall increase in the impact of role models on U.S. women was primarily driven by other (white) women and helped slightly by black women. For other (white) U.S. women, the impact of the role model effect on their employment became significant. It accounted for about an 8% increase in the likelihood of employment in the second generation. For Black U.S. women, a female respondent's mother working was significant in both generations and increased slightly between the generations by 0.031. Meanwhile, for Hispanic women, the role model effect has never been significant in predicting their adulthood employment.

Table 13: NLSY 1979 and 1997 Women by Race Probit Mean Marginal Effects

	79 BW	97 BW	79 HW	97 HW	79 OW	97 OW
Average Percentile Household Income	0.003*** p = 0.000	0.005*** p = 0.000	0.002* p = 0.013	0.003** p = 0.004	0.001+ p = 0.052	0.002*** p = 0.003
Mother Worked For Pay	0.046+ p = 0.095	0.077* p = 0.031	-0.060 p = 0.091+	0.056 p = 0.149	0.018 p = 0.352	0.079** p = 0.006
Respondent Married	0.043 p = 0.397	0.056 p = 0.421	-0.001 p = 0.985	0.084 p = 0.282	-0.053 p = 0.142	0.233*** p = 0.000
Respondent is Catholic	-0.033 p = 0.540	-0.084 p = 0.280	-0.062 p = 0.208	0.040 p = 0.347	-0.040+ p = 0.083	0.041 p = 0.193
Respondent Number Of Children	-0.041*** p = 0.000	0.006*** p = 0.682	-0.069*** p = 0.000	0.017 p = 0.491	0.062*** p = 0.000	0.013 p = 0.491
Children/Married Interacted	0.013 p = 0.510	-0.035*** p = 0.251	0.011 p = 0.619	-0.052 p = 0.121	0.001 p = 0.974	-0.098*** p = 0.000
Respondent Region North East	0.003 p = 0.954	-0.020 p = 0.752	-0.001 p = 0.993	-0.004 p = 0.969	-0.066* p = 0.044	-0.029 p = 0.503
Respondent Region South	0.042 p = 0.231	0.032 p = 0.473	0.036 p = 0.588	-0.079 p = 0.306	-0.082** p = 0.002	-0.041 p = 0.222
Respondent Region West	-0.047 p = 0.424	-0.242** p = 0.007	-0.006 p = 0.921	-0.106 p = 0.143	-0.114*** p = 0.001	-0.064 p = 0.109
Constant	54.446*** p = 0.000	26.433*** p = 0.000	51.450*** p = 0.000	51.352*** p = 0.000	50.099*** p = 0.000	46.941*** p = 0.000
Total Observations	1,208	763	750	585	1,999	1,307
R ²	0.245	0.190	0.187	0.085	0.093	0.104
Adjusted R ²	0.244	0.187	0.184	0.080	0.092	0.102
Rho	-1.230	0.043	-1.153	-0.977	-1.092	-0.774
Inverse Mills Rat	-50.957*** (10.525)	0.757 (6.642)	-41.995*** (9.425)	-25.107* (12.327)	-38.866*** (6.211)	-19.822** (6.937)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 and NLSY97 samples.

Education and Intermediate Outcomes

Lastly, we complement our investigation of multigenerational trends in income mobility by analyzing the impact of parent household income percentile and an intermediate outcome. A benefit of utilizing the NLSY79 and NLSY97 datasets is that they record many variables pertinent to the respondent's childhood environment, family relations, etc. We hope this paper's foundation encourages other researchers to explore the relationship between parent household income percentile and the

multitude of intermediate outcomes recorded in the NLSY datasets. However, due to the scope of this, it is most appropriate to investigate the relationship between parent household income percentile rank and their child's educational attainment, as education is commonly described in U.S. culture as the great equalizer and its role in intergenerational socioeconomic mobility has been studied extensively within the literature.

We utilize our samples' smallest demographic split (by race and sex, outlined in section 5.7) to investigate this relationship. If we let $R_{i,ed}$ represent our respondent's highest level of education completed by the year 2000 (NLSY79 sample) or the year 2019 (NLSY97 sample), then identical to our past investigations, we estimate absolute mobility ($\beta_{ra, s}$) and relative mobility ($\lambda_{ra, s}$) modeled by Equation 5.9 below,

$$R_{i,ed,ra,s} = \beta_{ed,ra,s} + \lambda_{ed,ra,s} P_i + e_i \quad (5.9)$$

In Appendix 7.2, we provide Figure 10, which presents the binned scatter plots of a respondent's highest level of education completed against their parent's household income percentile rank split by race and sex for each generation sample. As it can be observed, these relationships are again linear. However, their strength in linearity again varies by subgroup, with the weakest linear relationships being near the highest parent income percentiles for the NLSY79 samples and black and Hispanic men in the NLSY97 sample. Again, we attribute these issues to the smaller number of observations. Table 14 shows the relationships for each subgroup in the NLSY79 samples, while Table 15 does the same for the NLSY97 sample.

Table 14: NLSY 1979 Respondent Education Level OLS Regression Outputs

	79 BW	79 BM	79 HW	79 HM	79 OW	79 OM
Average Percentile Household Income	0.037*** p = 0.000	0.026*** p = 0.000	0.040*** p = 0.000	0.034*** p = 0.000	0.032*** p = 0.000	0.026*** p = 0.000
Constant	11.922*** p = 0.000	11.813*** p = 0.000	11.134*** p = 0.000	10.997*** p = 0.000	12.098*** p = 0.000	12.158*** p = 0.000
Total Observations	1,253	1,170	783	745	2,053	1,978
R ²	0.126	0.064	0.103	0.073	0.082	0.051
Adjusted R ²	0.125	0.064	0.102	0.072	0.082	0.050
Residual Std. Error	2.206(df = 1251)	1.992(df = 1168)	2.572(df = 781)	2.593(df = 743)	2.342(df = 2051)	2.532(df = 1976)
F Statistic	179.605*** (df=1;1251)	89.363*** (df=1;1168)	89.669*** (df=1;781)	58.505*** (df=1;743)	183.285*** (df=1;2051)	105.890*** (df=1; 1976)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY79 sample.

Table 15: NLSY 1997 Respondent Education Level OLS Regression Outputs

	97 BW	97 BM	97 HW	97 HM	97 OW	97 OM
Average Percentile Household Income	0.051*** p = 0.000	0.044*** p = 0.000	0.041*** p = 0.000	0.031*** p = 0.000	0.043*** p = 0.000	0.037*** p = 0.000
Constant	12.487*** p = 0.000	11.277*** p = 0.000	12.049*** p = 0.000	11.928*** p = 0.000	12.766*** p = 0.000	12.187*** p = 0.000
Total Observations	956	853	734	677	1,702	1,738
R ²	0.156	0.137	0.094	0.068	0.123	0.092
Adjusted R ²	0.155	0.136	0.093	0.067	0.123	0.091
Residual Std. Error	2.834(df = 954)	2.637(df = 851)	2.822(df = 732)	2.724(df = 675)	2.915(df = 1700)	2.889(df = 1736)
F Statistic	175.702*** (df=1;954)	135.273*** (df=1;851)	75.867*** (df=1;732)	49.269*** (df=1;675)	239.252*** (df=1;1700)	175.693*** (df=1; 1736)

Note: + p<0.1; * p<0.05; ** p<0.01; ***p<0.001

Source: own calculations based on NLSY97 sample.

When interpreting these coefficients, we can see slight yet negligible variation in the relationship between parent household income percentile ranks and their child's level of educational achievement. First, for those at the lowest end of the socioeconomic distribution, where one's parents theoretically have no income, we can see that across all racial and sex groups, these children are between the 11 and 12th income percentile rank in the 1979 and 1997 cohorts. When we look at the impacts of additional education grades completed, we can see some variation between subgroups, although incredibly small. For example, the differences between the highest and

lowest relative mobility in the 1979 sample (between black or other men and Hispanic women), at the highest education level (grade 20), there is a difference of 0.8 - 0.52 or only 0.28 of an income percentile rank. In the 1997 sample, this difference is identical, 0.102 - 0.74, a difference of 0.28 of an income percentile rank—obtained by comparing Hispanic men and black women.

Thus, although there are slight changes in the impact of one's parent's household income percentile rank on educational attainment between the generations, they are negligible and show that educational attainment (which has increased overall between the two samples) does not explain any differences or changes in intergenerational socioeconomic mobility observed between the generations. Thus, these estimates suggest that a significant amount of the variation in income percentile rank between demographic and sex groups within the United States must arise before our respondent enters the labor market. This finding is consistent with prior literature (Chetty et al., 2014a; Chetty et al., 2014b).

Conclusion

This study utilized NLSY79 and NLSY97 data to present measurements of intergenerational socioeconomic mobility trends across multiple generations within the United States. Between the generations of those born in the late 1950s to early 1960s and those born in the early 1980s, overall U.S. intergenerational mobility did not change or minimally increase depending on the definition of mobility utilized.

In the United States, we find evidence to support previous claims that intergenerational mobility varies by region, sex, and race. Regarding region, the intergenerational mobility over our studied timeframe has remained the same in the North Eastern or North Central United States. We find slight increases in

intergenerational mobility in the American West and slight to negligible decreases in the American South.

When viewing our multigenerational trends by sex, we estimate that U.S. women's intergenerational mobility negligibly increased, while American men's mobility did not change. Intergenerational mobility trends by race paint a slightly different picture. The three races considered in this study, black, Hispanic, and other (white) Americans, all saw increased socioeconomic mobility but to varying degrees. Hispanic Americans saw the most significant increases during this timeframe, black Americans saw either a negligible increase or decrease in their mobility, and other (white) Americans experienced a minimal increase. Most concerning, though, is that our multigenerational estimate trends suggest a small but growing increase in the gap between the socioeconomic mobility of white and black Americans during our studied timeframe.

This study also investigated if there are significant differences in mobility within racial groups by sex. Consistent with previous literature, we find that Hispanic women and men have experienced mobility increases. For other (white) Americans, the gains in mobility were almost entirely experienced by other (white) men. For black Americans, the stagnation in mobility was caused by the positive mobility gains of black women against the mobility losses by black men.

We also find that even though our data is not entirely representative of the timeframe's birth cohort, our employed corrections models for modeling U.S. women and using census microdata to create income percentile brackets result in incredibly similar IRA estimates by region, race, and sex when compared to previous studies. Lastly, we find that while role model effects became significantly more important in encouraging U.S. women to enter the workforce between the two generations, it was

almost wholly driven by other (white) women and had no impact on U.S. women's intergenerational socioeconomic mobility.

This study analyzed U.S. multigenerational trends in intergenerational socioeconomic mobility during significant economic, social, and cultural advancement for U.S. women and black Americans. However, these groups experienced the least improvements in intergenerational mobility compared to others in the U.S., specifically Hispanic Americans. More concerning, the mobility advancements of Black Americans are less than that of other Americans, leading to the gap in mobility between black Americans and everyone else to widen. This point is especially pertinent to U.S. Black men.

Our results indicate that the generational impact of massive cultural movements such as the civil rights movement and the advancement of American women in the workforce has impacted those groups' upward socioeconomic advancement little to none in the long run. Additionally, if an intermediate outcome (such as education) and an environmental upbringing influence (such as the role model effect for women) both fail to affect or change intergenerational mobility significantly, how can we begin to address these gaps? This question is especially pertinent if these gaps are increasing, as our estimates between U.S. white and black men suggest. As previous studies advocate, lessening these disparities is daunting and will require targeted interventionist policies to support those socioeconomically disadvantaged for generations.

References

- Alon, S., & Tienda, M. (2005). Job mobility and early career wage growth of White, African-American, and Hispanic women*. *Social Science Quarterly*, 86(s1), 1196–1217. <https://doi.org/10.1111/j.0038-4941.2005.00342.x>
- Amuedo-Dorantes, C., & Kimmel, J. (2005). ?the motherhood wage gap for women in the United States: The importance of college and fertility delay? *Review of Economics of the Household*, 3(1), 17–48. <https://doi.org/10.1007/s11150-004-0978-9>
- Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6), 1153–1189. <https://doi.org/10.1086/260831>
- Bellou, A., & Cardia, E. (2021). The Great Depression and the rise of female employment: A new hypothesis. *Explorations in Economic History*, 80, 101383. <https://doi.org/10.1016/j.eeh.2020.101383>
- Carrell, S.E., Page, M. E., & West, J. E. (2010). Sex and Science: How Professor Gender Perpetuates the Gender Gap*. *Quarterly Journal of Economics*, 125(3), 1101–1144. <https://doi.org/10.1162/qjec.2010.125.3.1101>
- Chetty, R., Friedman, J. N., Saez, E., Turner, N., & Yagan, D. (2017). Mobility Report Cards: The Role of Colleges in Intergenerational Mobility*. *National Bureau of Economic Research*.
- Chetty, R., Hendren, N., Jones, M. R., & Porter, S. R. (2019). Race and Economic Opportunity in the United States: An intergenerational perspective*. *The Quarterly Journal of Economics*, 135(2), 711–783. <https://doi.org/10.1093/qje/qjz042>
- Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014a). Where is the land of opportunity? the geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, 129(4), 1553–1624. <https://doi.org/10.3386/w19843>
- Chetty, R., Hendren, N., Kline, P., Saez, E., & Turner, N. (2014b). Is the United States still a land of opportunity? recent trends in intergenerational mobility. *American Economic Review*, 104(5), 141–147. <https://doi.org/10.1257/aer.104.5.141>
- Craigie, T.-A. (2021). Men’s incarceration and women’s labor market outcomes. *Feminist Economics*, 27(4), 1–28. <https://doi.org/10.1080/13545701.2021.1942510>
- Dahl, Molly W., and Thomas DeLeire (2008). The Association between Children’s Earnings and Fathers’ Lifetime Earnings: Estimates Using Administrative Data. Institute for Research on Poverty, University of Wisconsin–Madison.

- Dale, S. B., & Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, 117(4), 1491–1527. <https://doi.org/10.1162/003355302320935089>
- Fernandez, R., Fogli, A., & Olivetti, C. (2004). Preference formation and the rise of Women's labor force participation: Evidence from WWII. *National Bureau of Economic Research*. <https://doi.org/10.3386/w10589>
- Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American Economic Review*, 96(2), 1–21. <https://doi.org/10.1257/000282806777212350>
- Goldin, C., & Mitchell, J. (2017). The new life cycle of Women's employment: Disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives*, 31(1), 161–182. <https://doi.org/10.1257/jep.31.1.161>
- Goldin, C., & Olivetti, C. (2013). Shocking labor supply: A reassessment of the role of World War II on women's labor supply. *American Economic Review*, 103(3), 257–262. <https://doi.org/10.1257/aer.103.3.257>
- Gregg, P., Macmillan, L., & Vittori, C. (2018). Intergenerational income mobility: Access to top jobs, the low-pay no-pay cycle and the role of education in a common framework. *Journal of Population Economics*, 32(2), 501–528. <https://doi.org/10.1007/s00148-018-0722-z>
- Haider, S., & Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. *American Economic Review*, 96(4), 1308–1320. <https://doi.org/10.1257/aer.96.4.1308>
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153. <https://doi.org/10.2307/1912352>
- Heckman, J. J., Lyons, T. M., & Todd, P. E. (2000). Understanding black–white wage differentials: 1960–1990. *American Economic Review*, 90(2), 344–349. <https://doi.org/10.1257/aer.90.2.344>
- Hilmer, C., & Hilmer, M. (2007). Women helping women, men helping women? same-gender mentoring, initial job placements, and Early Career Publishing success for economics phds. *American Economic Review*, 97(2), 422–426. <https://doi.org/10.1257/aer.97.2.422>
- Humlum, M. K., Kleinjans, K. J., & Nielsen, H. S. (2012). An economic analysis of identity and career choice*. *Economic Inquiry*, 50(1), 39–61. <https://doi.org/10.1111/j.1465-7295.2009.00234.x>
- Jajtner, K. M. (2020). Work-limiting disability and intergenerational economic mobility. *Social Science Quarterly*, 101(5), 2001–2016. <https://doi.org/10.1111/ssqu.12836>

- Kofoed, M. S., & McGovney, E. (2017). The effect of same-gender or same-race role models on Occupation choice. *Journal of Human Resources*, 54(2), 430–467. <https://doi.org/10.3368/jhr.54.2.0416.7838r1>
- Luke, N. (2021). Gender and social mobility. *Social Mobility in Developing Countries*, 374–397. <https://doi.org/10.1093/oso/9780192896858.003.0015>
- Mitnik, P., Victoria B., David G., & Michael W. (2014). New Estimates of Intergenerational Income Mobility Using Administrative Data, ' Mimeo, Statistics of Income, Internal Revenue Service.
- Mulligan, C. B., & Rubinstein, Y. (2008). Selection, investment, and women's relative wages over time*. *Quarterly Journal of Economics*, 123(3), 1061–1110. <https://doi.org/10.1162/qjec.2008.123.3.1061>
- Neumark, D., & Gardecki, R. (1998). Women helping women? role model and mentoring effects on female ph.D.. students in economics. *The Journal of Human Resources*, 33(1), 220. <https://doi.org/10.2307/146320>
- Piotrowska, M., & Kośny, M. (2017). Economic transition and intergenerational mobility in Poland. *Economics & Sociology*, 10(3), 59–71. <https://doi.org/10.14254/2071-789x.2017/10-3/4>
- Restuccia, D., & Urrutia, C. (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review*, 94(5), 1354–1378. <https://doi.org/10.1257/0002828043052213>
- Wooldridge, J. M. (2010) *Econometric Analysis of Cross Section and Panel Data, Second Edition*. MIT Press.

Appendix

Variable Definitions

Average Respondent income percentile rank. The Original data for the NLSY79 sample was obtained from the NLSY79 “Total Income From Wages and Salary in Past Calendar Year” variable for 1990 to 1994, 1996, 1998, and 2000. For the NLSY97 sample, the initial data was from the NLSY97 “Total Income From Wages and Salary in Past Year” variable for the years 2009 to 2011, 2013, 2015, 2017, and 2019. Then, IPUMS CPS Wage and Salary Income microdata were collected for these respective years. Using the IPUMS data, percentile ranks were calculated for each respective year. Then, the respective NLSY respondent income variables were run through these percentile brackets for each respective year. Finally, the years for each respondent’s income percentiles were averaged. Note that zeros were included in the calculations and were not counted as NAs.

Average family income percentile rank. The Original data was obtained from the NLSY79 “Net Family Income” variable from 1979 to 1985 and the NLSY97 “Gross Household Income” variable from 1997 to 2003. Then, IPUMS CPS Total Household Income microdata were collected for these respective years. Using the IPUMS data, percentile ranks were calculated for each respective year. Then, the respective NLSY family income variables were run through these percentile brackets for each respective year. Finally, the seven years of household income percentiles were averaged for each respondent’s parents. Note that zeros were included in the calculations and were not counted as NAs.

Respondent’s Mother Employed (Role Model Effect). For the NLSY79 sample, this data was obtained from the binary variable, “Did Adult Female Present in Household at Age 14 Work for Pay?” For the NLSY97 sample, it had to be coded;

thus, in the “Household Member Specific Variables” category for 1997, “Sex,” “Age,” “HHID,” and “Employed” were obtained for the first six household members were obtained. Then, based on age and sex, they were identified as the respondent’s mother or not, and lastly, they were coded based on whether they were employed or not. Thus, these measurements only proxy for a female respondent’s mother working at least one year during their teenage years.

Respondent Employment. For the NLSY79 sample, this data was obtained from the NLSY79 “Number of Weeks Worked in Past Calendar Year *Key*” variable for the year 2000. For the NLSY97 sample, this data was obtained from the “Weeks R Worked Any Job Year 18” variable. From here, a respondent is coded as employed (1) if they worked 40 or more weeks out of the year and unemployed (0) if they worked less than 40 weeks out of the year.

Respondent Married. For the NLSY79 sample, this data was obtained from the NLSY79 “Marital Status” variable in the year 2000. For the NLSY97 sample, it was obtained from the “Marital Status” variable in 2019. For each variable, if the respondent was currently married, they were coded as 1. Otherwise, they were coded as 0.

Respondent’s Number of Children. For the NLSY79 sample, this data was obtained from the NLSY79 “Number of Bio/Step/Adpt Children in Household” variable for 2000. For the NLSY97 sample, it was obtained from the “Bio Children R Has in Household (Under Age 18 Only)” variable for 2019.

Respondent Region. For the NLSY79 sample, the data was obtained from the NLSY79 “Region of Current Residence” variable in 1979. For the NLSY97 sample, it was obtained from the “Census Region of Residence” variable in 1997. The breakdown of the U.S. states that compose each region is below (same region definitions for each NLSY sample):

Northeast. CT, ME, MA, NH, NJ, NY, PA, RI, VT

North Central. IL, IN, IA, KS, MI, MN, MO, NE, OH, ND, SD, WI

South. AL, AR, DE, DC, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV

West. AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY

Respondent Catholic. For the NLSY79 sample, the data was obtained from the NLSY79 “In What Religion Was R Raised?” variable and coded 1 if they were Catholic and 0 if they were raised in any other religion. For the NLSY97 sample, it was obtained from the “What Religion was PR Raised in?” variable and again coded 1 if Catholic and 0 for all other religions.

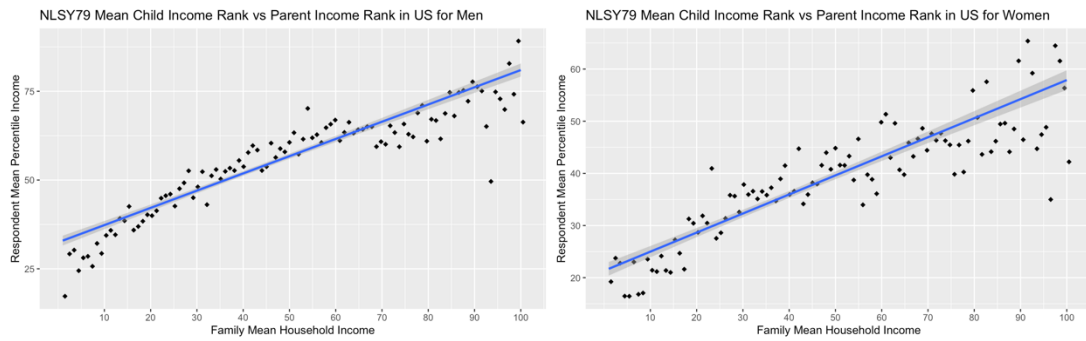
Respondent Education Level. For the NLSY79 sample, the data was obtained from the NLSY79 “Highest Grade Completed as of May 1 Survey Year (Revised)” variable for the year 2000. For the NLSY97 sample, it was obtained from the “R’s Highest Grade Completed” variable for 2019. For both datasets, each education level is signified below.

- 0 means no education
- 1 through 12 signify, respectively, grades 1st through 12th
- 13 through 19 translate to 1st through 7th year at college
- 20 translates to 8th year at college or more

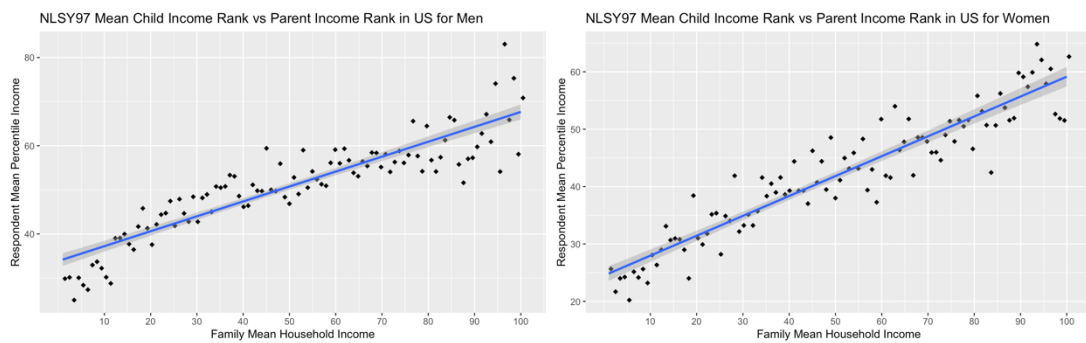
Additional Tables

Figure 7: Linear Relationship of Parent to Respondent Percentiles by Sex

7.1. US Women NLSY79 and NLSY97 Linear Relationships



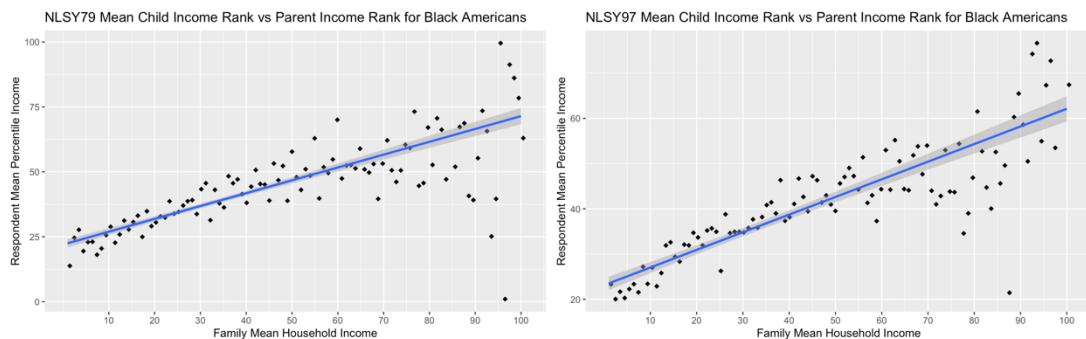
7.2. US Men NLSY79 and NLSY97 Linear Relationships



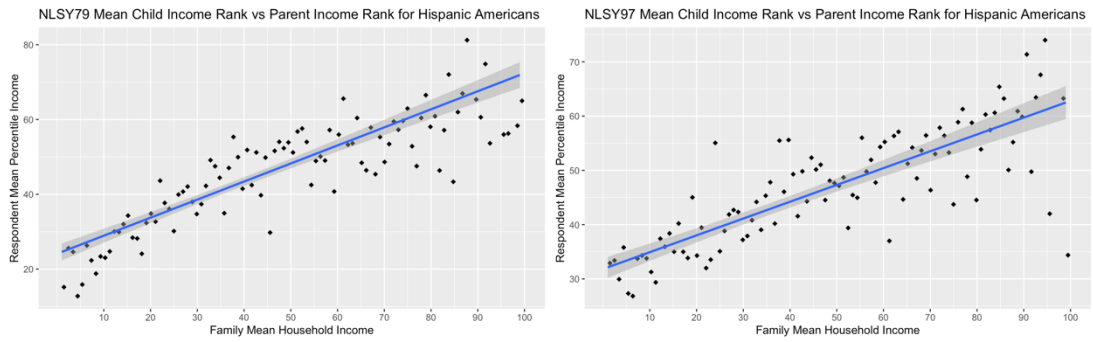
Source: Made in R based on NLSY79 and NLSY97 samples.

Figure 8: Linear Relationship of Parent to Respondent Percentiles by Race

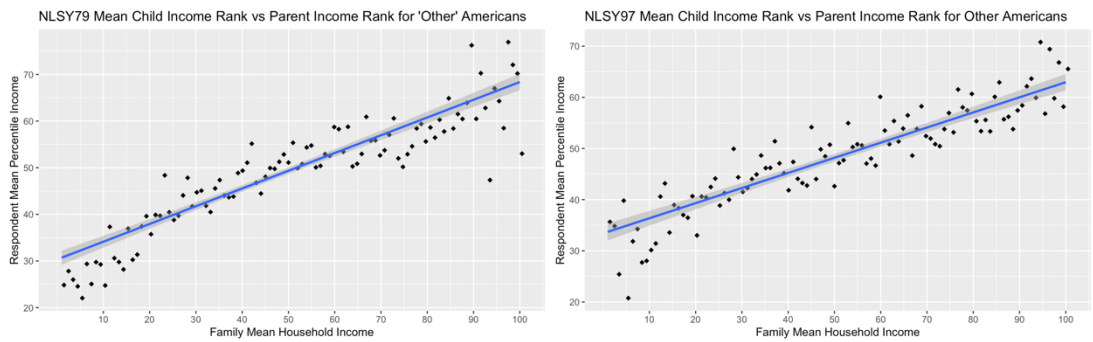
8.1. NLSY79 and NLSY97 Linear Relationships Black Americans



8.2. NLSY79 and NLSY97 Linear Relationships Hispanic Americans



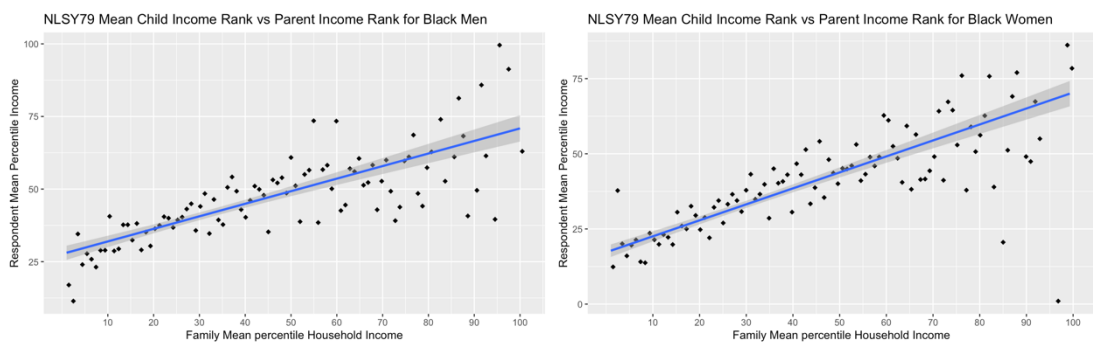
8.3. NLSY79 and NLSY97 Linear Relationships Other Americans



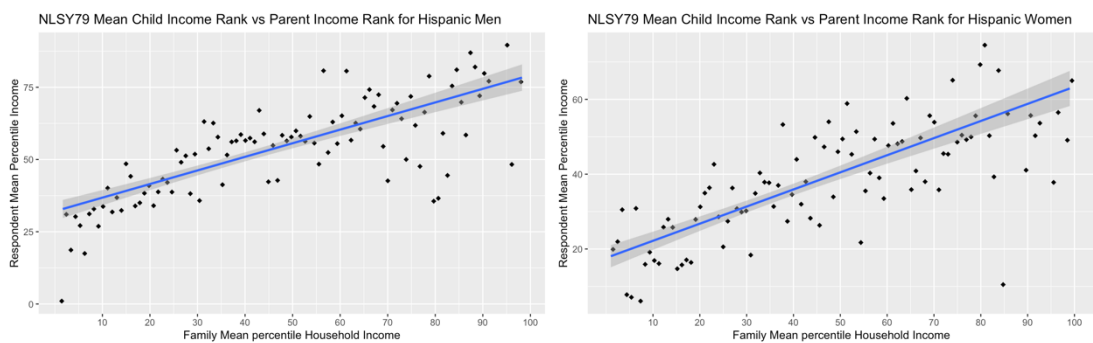
Source: Made in R based on NLSY79 and NLSY97 samples.

Figure 9: Linear Relationship of NLSY79 and NLSY97 Parent to Respondent Percentiles by Race and Sex

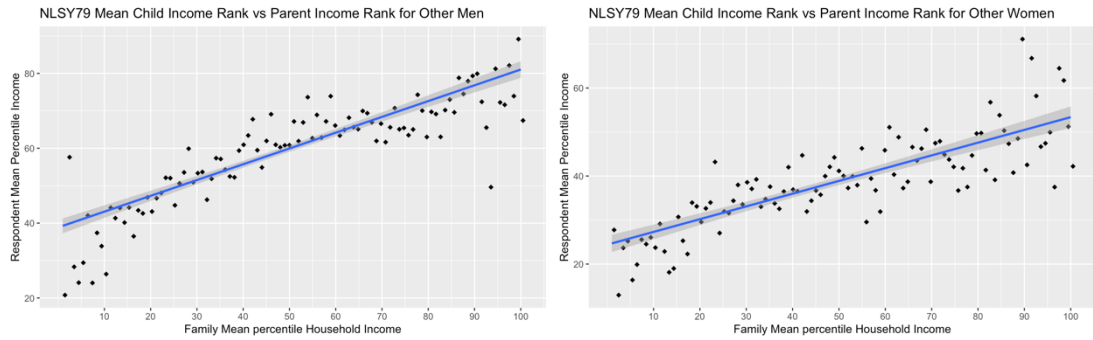
9.1. NLSY79 Linear Relationships Black American Men and Women



9.2. NLSY79 Linear Relationships Hispanic American Men and Women

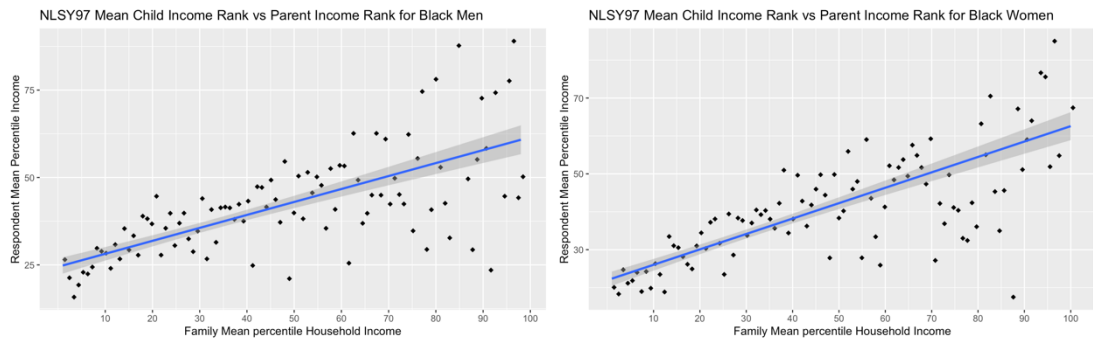


9.3. NLSY79 Linear Relationships Other American Men and Women

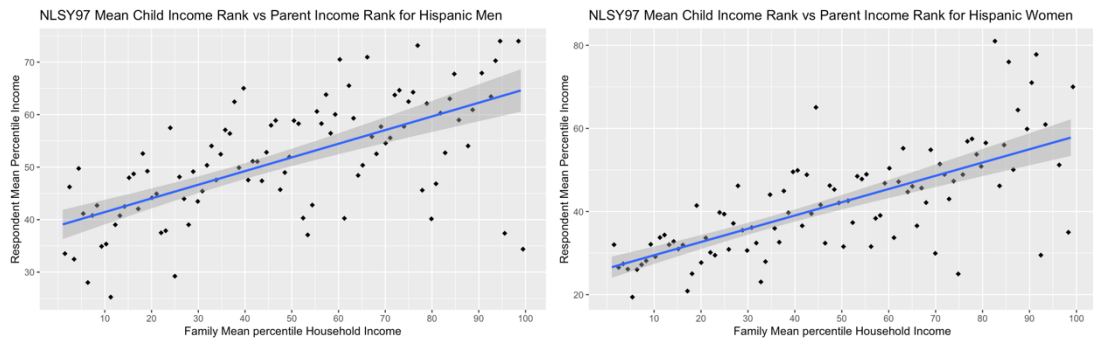


Source: own calculations based on NLSY79 samples.

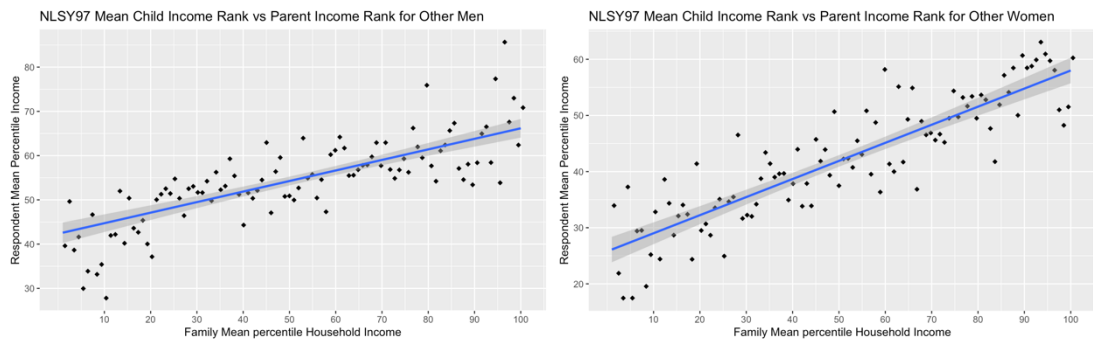
9.4. NLSY97 Linear Relationships Black American Men and Women



9.5. NLSY97 Linear Relationships Hispanic American Men and Women



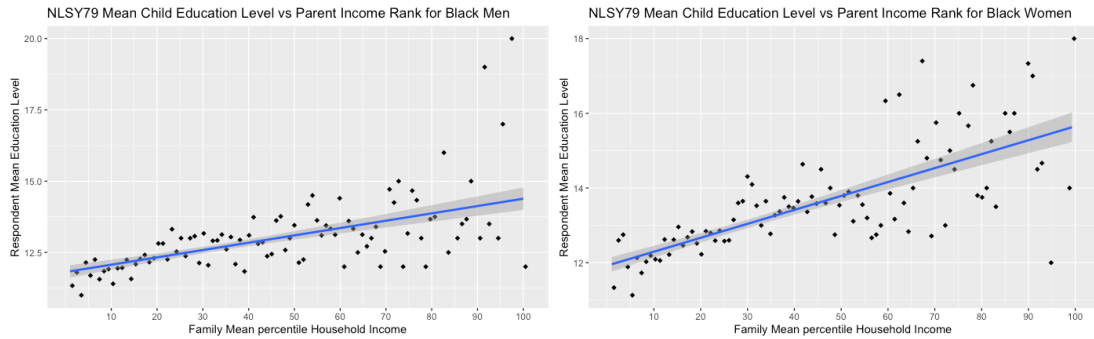
9.6. NLSY97 Linear Relationships Other American Men and Women



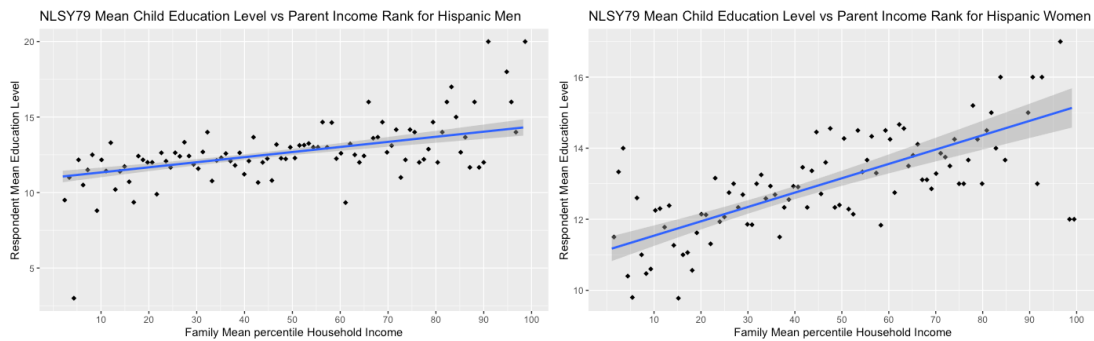
Source: Made in R based on NLSY97 samples.

Figure 10: Linear Relationship of NLSY79 and NLSY97 Respondent Attained Education Level to Respondent Percentiles by Race and Sex

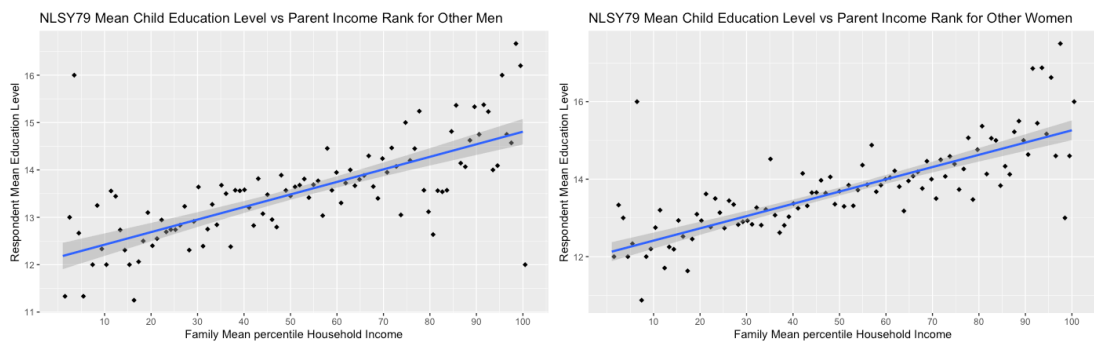
10.1. NLSY79 Black American Men and Women's Education



10.2. NLSY79 Hispanic American Men and Women's Education

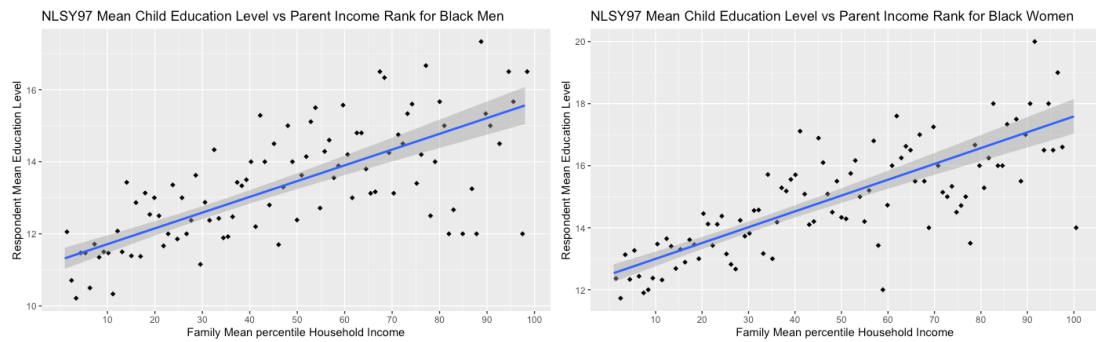


10.3. NLSY79 Other American Men and Women's Education

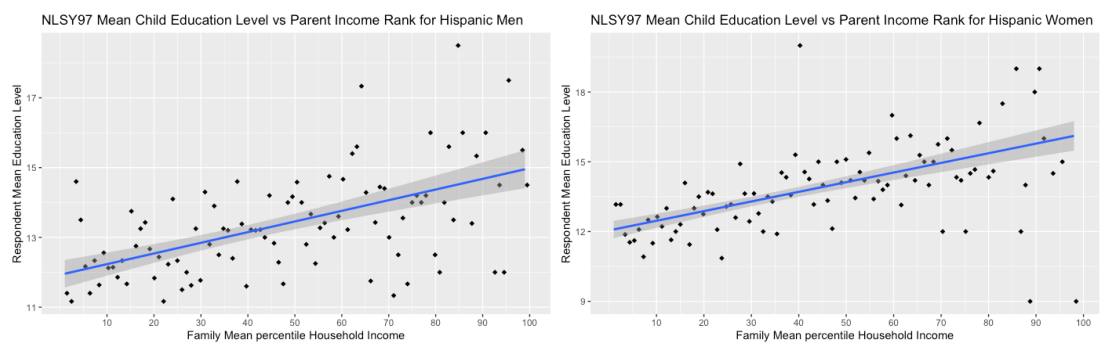


Source: Made in R based on NLSY79 samples.

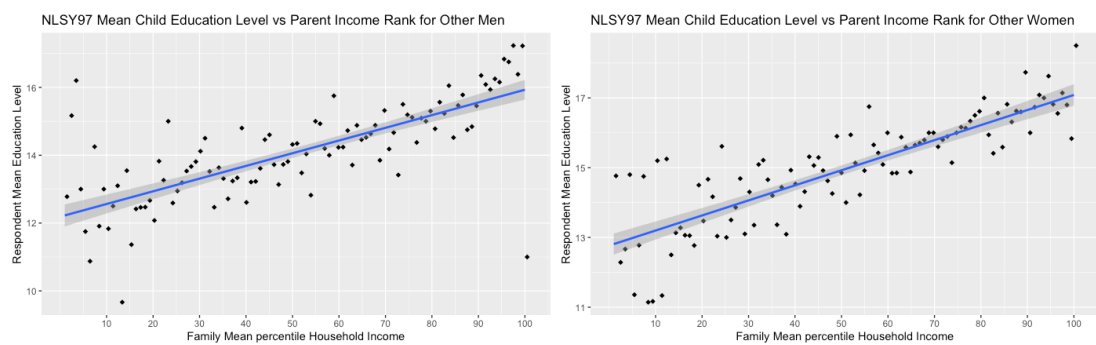
10.4. NLSY97 Black American Men and Women's Education



10.5. NLSY97 Hispanic American Men and Women's Education



10.6. NLSY97 Other American Men and Women's Education



Source: Made in R based on NLSY97 samples.