From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Sociology

2024

THREE ESSAYS ON INTERGENERATIONAL MOBILITY

Wei, Lai

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01bc386n59f

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

THREE ESSAYS ON INTERGENERATIONAL MOBILITY

LAI WEI

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY

THE DEPARTMENT OF SOCIOLOGY

Adviser: Yu Xie

May 2024

© Copyright by Lai Wei, 2024. All rights reserved.

ABSTRACT

This dissertation develops and applies innovative causal inference methods to investigate the causes, mechanisms, and consequences of intergenerational mobility within the United States. Chapter 1 redefines the concept of mobility effects, a traditional focus of social stratification research, by shifting attention from the individual to the societal level. It establishes a methodological framework for assessing the impact of macro-level social mobility, specifically examining how changes in social mobility influence societal outcome distributions. This approach is demonstrated through two analyses exploring the effects of social mobility on average fertility rates and income inequality in the United States. Chapter 2 explores the influence of labor market transformations on the mechanisms underlying intergenerational income transmission. Utilizing data from diverse sources and employing a novel causal decomposition technique, this chapter reveals a shift in intergenerational income transmission away from cognitive factors towards non-cognitive ones in recent decades in the United States. Chapter 3 evaluates the impact of parental college education on intergenerational mobility. While previous discussions have primarily centered on the college attainment of children, this study redirects attention to the educational backgrounds of parents. The research demonstrates that children of college-educated parents experience significantly greater mobility compared to those with non-college-educated parents. I assess the robustness of this pattern to confounding factors by evaluating a novel causal estimand. Collectively, this dissertation addresses longstanding puzzles in social mobility studies through the application of cutting-edge causal inference methodologies. These methodological advancements offer new insights on traditional research questions and theoretical perspectives.

ACKNOWLDGEMENTS

I am deeply grateful to my advisor, Yu Xie, whose mentorship transformed me from an undergraduate novice in research to a social observer and scientist. Under his guidance I learned how to be a scholar and a colleague. My dissertation would not be possible without the contributions of the other distinguished committee members. I extend my sincere appreciation to Brandon Stewart for his encouraging support and insightful critiques, which significantly enriched my research. I am grateful to Matthew Salganik for his readiness to challenge my ideas and offer novel perspectives, and to Xiang Zhou for his invaluable assistance in theoretical development and methodology. It has been a privilege to work with such distinguished scholars.

I am also indebted to the faculty members within the department who generously shared their time and expertise. I want to thank Miguel Centeno, Dalton Conley, Adam Goldstein, Janet Vertesi, Tod Hamilton, and Mitchel Duneier for their guidance and support throughout my graduate journey. I am grateful to James Raymo for graciously agreeing to serve as my external reader.

The exceptional students in the program have been a constant source of inspiration and support. I am particularly grateful to Han Zhang and Ian Lundberg for their invaluable insights and guidance during the early stages of my program. I am also thankful to my other peers, including Lemeng Liang, Ziyao Tian, Ha-Joon Chung, Max Fineman, Marie-Lou Laprise, Angela Li, Hunter York, Alejandro Schugurensky, Fumiya Uchikoshi, Lauren Clingan, Gavin Cook, Henry Gomory, Chris Felton, Shay O'Brien, Alex Kindel, Larry Liu, Hannah Waight, Robin Lee, Grace Tien, and Cheng Cheng, for their collegiality and encouragement.

In addition to my committee members and peers, I am grateful for the constructive feedback and support from colleagues such as Xi Song, Siwei Cheng, Fabian Pfeffer, Liying

Luo, Adam Goldstein, and Boyan Zheng, whose insights greatly improved different parts of my dissertation.

I am fortunate to have received support from various institutions, including the Department of Sociology and Center on Contemporary China, whose assistance facilitated my study and research. I am also grateful to the Center for Social Research at Peking University for hosting me during my time in China. I also want to thank the regular research workshops I attend throughout my graduate study. Those platforms provide me with opportunities to communicate with colleagues on issues I am interested in. I thank the Stewart Lab, Strat Workshop, Social Statistics Reading Groups at Princeton, Harvard, and Penn, China Workshop, and READI Meeting. Their main organizers are Brandon Stewart, Yu Xie, Xiang Zhou, Xi Song, Fumiya Uchikoshi, and Joeun Kim.

I owe a debt of gratitude to my mentors during my college years—Fei Yan, Andrew Walder, Yuan Shen, Hao Luo, Lu Zheng, and Jingming Liu—for igniting my passion for academia and guiding me onto the path of a scholar. Among them I am utmost grateful to Fei Yan, my undergraduate advisor, whose support and encouragement motivated me to pursue a Ph.D. in the U.S.

Finally, none of this would have been possible without the unwavering support of my family. I am deeply grateful to my partner, Ruifan Olive Ma, for her unconditional love and support. I extend my thanks to my parents, Guangwei Wei and Dongyun Zhang, especially my mother, Dongyun Zhang, whose dedication to knowledge, integrity, and diligence has been a constant source of inspiration ever since I was a child.

TABLE OF CONTENTS

ABSTRACT ***3***

ACKNOWLDGEMENTS ***4***

TABLE OF CONTENTS ***6***

INTRODUCTION ***8***

CHAPTER 1. SOCIAL MOBILITY AS CAUSAL INTERVENTION ***11***

INTRODUCTION 12

UNIDENTIFIABILITY OF INDIVIDUAL MOBILITY EFFECTS 13

THEORETICAL MOTIVATION OF SME 17

SME: DEFINITION, IDENTIFICATION, AND ESTIMATION 20

APPLICATIONS: SME ON AVERAGE FERTILITY AND INCOME INEQUALITY 40

DISCUSSION AND CONCLUSION 49

APPENDICES 53

FIGURES 58

CHAPTER 2. LABOR MARKET TRANSFORMATION AND EFFECTIVELY MAINTAINED IMMOBILITY IN THE UNITED STATES ***66***

INTRODUCTION 67

MOTIVATION 71

THEORETICAL AND ANALYTICAL FRAMEWORK 84

DATA 98

RESULTS 102

DISCUSSION 107

APPENDICES 111

FIGURES AND TABLES 133

CHAPTER 3 THE IMPACT OF PARENTAL EDUCATION ON INTERGENERATIONAL INCOME TRANSMISSION ***141***

INTRODUCTION 142

DOES COLLEGE IMPROVE SOCIAL MOBILITY? 144

PARENTAL INCOME AND EDUCATION: OVERLAPPING, ADDITIVE OR INTERACTIVE EFFECTS? 146

RESEARCH HYPOTHESES: COMPLEMENTARITY OR SUBSTITUTIVITY? 150

DATA 155

DESCRIPTIVE RESULTS 157

ADDRESSING SELECTION: A NEW ESTIMAND 160

CAUSAL ANALYSES 168
CONCLUSION 172
APPENDICES 176
FIGURES AND TABLES 177
CONCLUSION ***186***
BIBLIOGRAPHY ***190***

INTRODUCTION

The question of what drives social stratification, characterized by the uneven distribution of positions and resources based on socioeconomic or demographic factors, has long been central to sociology. While we have developed a comprehensive understanding of the descriptive landscape of social mobility—knowing the extent of mobility across different periods, countries, and measures—our grasp of the causes of intergenerational immobility and its consequences remains less clear.

Existing research faces methodological limitations that hinder a satisfactory explanation. Primarily, most quantitative studies fail to directly address social stratification, which entails the association between multiple variables. Instead, they often utilize regression models to explain individual outcomes, focusing on a single dependent variable. For instance, a regression model might analyze income as the dependent variable, with parental income and education as independent variables, revealing how parental income, given children's education levels, influences children's income. However, such models do not fully capture the factors contributing to income persistence—the association between parent and children's incomes. Consequently, contemporary social stratification research offers limited insight into the potential impacts of policy interventions. Questions regarding effective policies to reduce racial income disparities or enhance intergenerational mobility remain largely unanswered. What data, assumptions, and methodologies are necessary to credibly assess these policy effects? These critical inquiries remain largely unaddressed.

This dissertation is largely motivated by these methodological shortcomings. Drawing on tools from causal inference, each of its three articles aims to provide a clear definition of the estimand, elucidate identification assumptions, and employ flexible estimators to derive

estimates. This approach aligns with recent calls to enhance research practices in sociology (Lundberg et al., 2021). In contrast to earlier studies focusing on individual outcomes, these papers prioritize social mobility itself—the relationship between parental and children status—as the primary statistical objective of explanation. The three articles respectively tackle the consequences, mechanisms, and causes of intergenerational mobility.

Chapter 1 studies the consequence of intergenerational mobility. While researchers have long been intrigued by the implications of mobility, commonly referred to as the "mobility effect," earlier formulations of this problem encountered significant identification challenges. This arises from the difficulty in envisioning a counterfactual scenario at the individual level where one's mobility status changes while the origin and destination remain constant. In collaboration with Yu Xie, we redefine the mobility effect at the societal level, focusing on the potential consequences of altering the shape of social mobility—the association between parent and child status. By doing so, we establish a clear causal understanding of the issue. Leveraging this methodology, we examine the potential impact of enhancing social mobility on average fertility and income inequality in the United States. Our analysis reveals that while improving social mobility is unlikely to affect fertility rates significantly, it holds considerable potential for mitigating income inequality within society.

Chapter 2 studies the mechanism of intergenerational mobility. In recent decades, intergenerational income mobility in the United States has exhibited a stable trend. The US labor market, on the other hand, has undergone a notable transformation in skill requirements over recent decades, with cognitive skills diminishing in importance while non-cognitive skills have gained prominence. This shift suggests a corresponding alteration in the channel through which income is transmitted between generations. To assess the evolving mechanisms of

intergenerational income persistence, I developed a novel sequential decomposition method that extends the conventional Kitagawa-Blinder-Oaxaca decomposition and has clear causal interpretations as interventional effects. Utilizing data from diverse sources, including the National Longitudinal Survey of Youth '79 and '97, multiple waves of the US census, the American Community Surveys, the Occupational Information Network (O*NET), and information extracted from newspaper job advertisements, I discovered that the stability in the income mobility rate results from the interplay of two opposing forces. Specifically, intergenerational persistence through the cognitive channel has diminished, while persistence through the non-cognitive channel has increased.

Chapter 3 studies the causes of intergenerational mobility. Extensive research has explored the relationship between college education and intergenerational mobility in the United States. One prominent theory that has emerged from this literature is the "great equalizer" hypothesis, which suggests that college education serves as a pathway to higher social mobility for individuals. However, a gap in the literature exists regarding whether the children of college-educated parents experience higher or lower social mobility compared to those with non-college-educated parents. This study addresses this gap by evaluating the intergenerational mobility of children based on their parents' educational attainment. Specifically, it focuses on a novel causal estimand that approximates the counterfactual mobility if college-educated parents had not completed college and if non-college-educated parents had completed college. The findings indicate that children of college-educated parents indeed experience higher mobility, and this result remains robust to potential confounding factors. Further analyses reveal that factors such as educational expectations, parenting style, and educational expenditure all play a role in the "cross-generation" equalizing power of college education.

CHAPTER 1. SOCIAL MOBILITY AS CAUSAL INTERVENTION

Abstract[1]

The study of mobility effects is an important subject of study for sociologists. Empirical investigations of individual mobility effects, however, have been hindered by one fundamental limitation, the unidentifiability of mobility effects when origin and destination are held constant. Given this fundamental limitation, we propose to reconceptualize mobility effects from the micro to the macro level. Instead of micro-level mobility effects, the primary focus of the past literature, we ask alternative research questions about macro-level mobility effects: what happens to the population distribution of an outcome if we manipulate the mobility regime, that is, if we alter the observed association between social origin and social destination? The proposed method bridges the macro and micro agendas in social stratification research and has wider applications in social stratification beyond the study of mobility effects. We illustrate the method with two analyses that evaluate the impact of social mobility on average fertility and income inequality in the United States respectively.

[1] This research paper is coauthored with Yu Xie (Department of Sociology, Princeton University).

INTRODUCTION

Intergenerational mobility has been a core subject of sociological and demographic research since the middle of the last century (Lipset and Bendix, 1962; Blau and Duncan, 1967). One longstanding strand of this research is concerned with mobility effects. Tracing back to Sorokin (1927), social scientists have long hypothesized that individuals of different social mobility statuses display different social behaviors. For instance, the dissociative hypothesis states that social mobility incurs a penalty on the subjective well-being of socially mobile individuals because they are dissociated from their original families (Ellis and Lane, 1967; Houle and Martin, 2011; Hadjar and Samuel, 2015; Chan, 2018).

Despite their theoretical appeal, the empirical evaluations of hypotheses about individual-level mobility effects are hindered by one fundamental limitation relating to identification: mathematical dependence among family origin, individual destination, and mobility. Since mobility is defined by origin and destination, any model that tries to incorporate all three variables is inherently indeterminate, because there is no variation in mobility holding origin and destination constant. With no logical way of manipulating mobility with origin and destination fixed, there is no logical identification of causal effects (Holland, 1986). Earlier methods attempted to solve the unidentifiability issue (Duncan, 1966; Hope, 1975; Sobel, 1981, 1985; Luo, 2021), but they all require statistical and substantive assumptions which are sometimes implicit and untestable.

In this research, we reconceptualize the study of mobility effects by changing the level of analysis from the individual to the society. At the individual level, there is no mobility effect with origin and destination fixed. At the society level, however, mobility is still manipulable with the marginal distributions of origin and destination fixed. Using a counterfactual mobility table as a

causal intervention, and the population distribution of the dependent variable as an outcome, we ask an alternative research question about mobility effects: what happens to the distribution of an outcome if we manipulate the mobility regime of a society? One example of this type of question would be: if the social fluidity of the US were the same as that of Sweden, how would the average subjective well-being of US citizens change?

To answer research questions conceptualized in this way, we develop in this paper a new quantity of interest, the social mobility effect (SME). In the past decades, the dominant direction of social mobility research has been comparative descriptive studies of social mobility tables. SME not only asks whether mobility patterns differ, but also whether and to what extent such differences matter. As a conceptual tool, SME serves to bridge two disjointed agendas in social stratification research, the macro approach and the micro approach.

This article proceeds as follows. The second section describes the identification issue of the individual-level mobility effect. The third section reviews and situates SME in the literature of social stratification and describes why SME has the potential to advance the field of social stratification research. The fourth section describes the definition, identification, and estimation of SME. The fifth section illustrates the use of SME by studying the mobility effect on fertility and income inequality in the US. The sixth section concludes.

UNIDENTIFIABILITY OF INDIVIDUAL MOBILITY EFFECTS

Prior research has focused exclusively on the effect of mobility at the individual level. It asks the research question: net of the influence of social origin and social destination, what is the influence of individual mobility status on the individual's outcomes. However, it is well-known that the individual mobility effect has identifiability issues, as there is a functional dependence

among origin, destination, and mobility, akin to the Age-Period- Cohort (APC) problem (for a recent review of the APC problem see Fosse and Winship (2019)). Additional assumptions are required to make the statistical model identifiable. Let O, D, and Y denote respectively social origin or parental characteristics, social destination or own characteristics, and an outcome[2]. O and D can be continuous or categorical in this setting[3]. M is a mobility indicator that can be either the interaction of O and D when they are categorical or the difference between O and D when they are continuous. The earlier literature estimated the mobility effect with a generalized linear model where g is a link function, by the following model:

$$E[Y|O,D,M] = g(\alpha + \beta_1 O + \beta_2 D + \beta_3 M)$$

One apparent problem with this model is that not all coefficients are estimable due to the functional dependence of O, D, and M. For example, if O and D are continuous so that mobility is specified as a difference, $M = O - D$, then the model would have perfectly collinear independent variables and only two of the three β s would be identifiable. If O and D are categorical, so that M is the interaction of the two, M is identifiable only with a scale normalization.

Three approaches have been adopted to address the issue of statistical unidentifiability: identification via constraint/assumption, identification via normalization/coding scheme, and partial identification. To make the model identifiable, some prior researchers impose constraints on the coefficients, essentially bringing in extra information. The most straightforward constraint

[2] In the case of the Diagonal Mobility Model, origin and destination are specified as mean outcomes of the corresponding diagonal cells in the mobility table, which are essentially scaled categorical variables.
[3] In the illustration here O and D can be continuous but our methodological proposal discussed in this article only deals with categorical origin or destination.

is to assume that one of the three factors, $O, D,$ or M, has no influence on the outcome so that its β s are zero; on this basis the model becomes identifiable. An alternative is to estimate a non-linear model where the independent variables are trans- formed to avoid functional dependence. The substantive assumption under this constraint is that the independent variables affect the outcome in a specified non-linear, instead of linear, way.

The second approach to identifying the mobility effect, especially when origin and destination are specified as categorical variables, is to rely on normalizations that do not contain inherent meanings (Duncan, 1966; Hope, 1975; Luo, 2021). They use coefficients derived from a model where social origin and social destination are interacted to measure mobility effects. The coefficients in these regressions, however, are intrinsically undetermined and are subject to normalization. Because the coefficient is undetermined, the estimated mobility effect is also undetermined. Identification via normalization differs from identification via assumption as the former contains no assumption with inherent meaning. There is no scientifically correct way to choose a normalization choice over another, and the normalization choice does not change the model in any way, only how the results are presented.

Finally, a recent article has shown that the mobility effect can be partially identified with mild assumptions, despite that the point identification can only be achieved with very strong assumptions (Fosse and Pfeffer, 2019). With reasonable assumptions on the direction and magnitude of the effects of origin, destination, and mobility, one can roughly identify the interval in which the effect may lie.

We can understand the issue better if we define mobility as a cause in the counterfactual causal inference literature. The unidentifiability arises from the unmanipulability of the independent variables at the same time. Holding origin and destination constant, it is impossible

to manipulate one's mobility status. An analogous example is the interpretation of high-order terms in polynomial models. Imagine that an outcome Y is generated by the model $Y = \alpha + \beta_1 X + \beta_2 X^2$. People would not interpret the coefficient for the quadratic term β^2 as the effect of X^2 net of X, because whenever X changes, X^2 changes accordingly, so the effect of X^2 net of X is ill-defined.

This particular unmanipulability issue is known as the "cross-world" problem in the causal mediation literature (Naimi et al., 2014). Suppose that we denote a person's origin, destination, and mobility status with (o, d, m), and define the person's potential outcome with m as $Y_i(m) = Y_i$. A typical causal inference question is as follows. To learn the unconditional mobility effect of m versus m' for the individual we imagine a potential outcome where the person has a counterfactual mobility status $Y_i(m')$. The problem is that in our world we cannot observe the counterfactual outcome $Y_i(m')$, so we try to learn it from the observed outcome of other individuals. Now suppose that we want to learn the mobility effect conditional on origin and destination instead, so we imagine a counterfactual outcome of this individual $Y_i(o, d, m')$. Now the problem is not just that in our world we cannot observe $Y_i(o, d, m')$, it is that in no possible world can we observe $Y_i(o, d, m')$, since there is not a single world among the infinite worlds where we can observe the individual with m′ but also with (o, d)[4]. To be more concrete, suppose we observe an individual, A, who graduated from college and whose parents also graduated from college. Then A is immobile. Individual mobility effect tries to answer the question: what would happen to A if he/she had experienced a downward (or upward) mobility

[4] Though we borrow the concept "cross-world" from causal mediation to illustrate the logical dilemma of individual mobility effect, we do not think causal mediation is illogical. The reason is that though a potential outcome in causal mediation may not be observable in a certain world, it is still imaginable that a person may receive a treatment but a different version of mediator.

versus immobility given that he/she and his/her parents all graduated from college? The question is ill-posed because in no logical world can we observe counterfactually different mobility experiences of A.

THEORETICAL MOTIVATION OF SME

Given the unidentifiability of the individual mobility effect, in this article, we reconceptualize mobility effects at the society level, aggregated from the individual level. Still suppose an immobile person A who graduated from college and whose parents all graduated from college. We ask an alternative question to individual mobility effect: what will happen to A if we intervene to assign him/her to a destination according to a new mobility scheme – that is the rule for allocating people to destinations based on origin? Instead of asking what happens to an individual if we change the mobility status of that individual, we ask what happens to the population if we change the mobility regime of a society. At the individual level, mobility M is fixed after controlling for origin O and destination D. At the society level, there is a large range of permissible mobility regimes $P(O, D)$ when the origin and destination marginal distributions $P(O)$ and $P(D)$ are fixed, hence creating space for manipulation. There are large potential theoretical and substantive payoffs from a reorientation toward macro-level mobility effects.

Theorists of social mobility have long been concerned with the consequences of macro mobility patterns. For example, Blau and Duncan (1967: 392) stated in their landmark study that "The viewpoint governing the discussion in this chapter...has been concerned with the question of what possible demographic consequences follow from the volume of mobility observed in an economically advanced society." More recently, there has been a debate about the effect of social mobility on health inequality, where some hold that mobility reduces health inequality while

others believe that it increases health inequality (Bartley and Plewis, 1997; Blane et al., 1999; Claussen et al., 2005; Bartley and Plewis, 2007; Boyle et al., 2009). Due to a lack of proper tools, earlier studies conducted individual-level analyses about mobility effects and then made inferences about macro-level consequences by aggregating up individual-level estimates. However, as argued above, the individual-level mobility effect is fundamentally unidentifiable, and the macro-level aggregates de- rived from micro-level effects are therefore flawed. By contrast, our method estimates macro-level mobility effects without identifying individual-level mobility effects, thereby circumventing the problem of individual-level unidentifiability.

A methodological tool for understanding the consequences of mobility patterns has been called for by a revived academic interest in intergenerational mobility in recent years. Social mobility has been found to be stalled or declining in developed countries (Erikson et al., 1992; Chetty et al., 2017; Hout, 2018; Bukodi and Goldthorpe, 2021), against the optimistic prediction of the liberal theory that industrial societies will become more open over time (Treiman, 1970). Scholars have associated this stalled mobility with a wide range of social problems, including political polarization, anti-immigrant and xenophobic sentiments, and the deterioration of physical and mental health (Jackson and Grusky, 2018; Case and Deaton, 2020). Prior studies have employed ad hoc methods to test the effect of "desegregating" educational background by family origin (Chetty et al., 2020). Whether mobility pattern is indeed a cause of these social problems, however, remains to be tested formally.

From a broader theoretical perspective, SME bridges the macro and the micro traditions of research in social stratification and mobility. The macro-structural tradition investigates social relations across different social positions (Blau, 1977). Though the unit of data analysis may be the individual, the analytical focus is on the individual's position in a social structure. One of the

most fruitful branches of the macrostructural tradition is the analysis of social mobility tables (Duncan, 1979; Hauser, 1980; Erikson et al., 1992). Mobility table studies make use of categorical data analysis to model the linkages between parents' social positions and those of their children. The macro tradition is descriptive in nature, as its research goal is to generate proper descriptions of social stratification systems within historical and comparative contexts.

The micro tradition of stratification research, on the other hand, is more oriented to- ward individual-level analyses. Instead of asking the descriptive question of how social positions are interrelated, the micro approach asks the causal question of what would happen to an individual if he/she were placed in a different social position. The status attainment model that attempts to estimate the effects of parental as well as an individual's own early- life factors on one's adult outcomes through structural equation modelling was perhaps the first causal approach to study stratification at the individual level (Blau and Duncan, 1967). More recent examples include research on the effect of college education on income using observational data (Brand and Xie, 2010), the effect of parental education on children's education using twin design (Holmlund et al., 2011), and the effect of parental income on children's education using instrumental variables (Chevalier et al., 2013).

SME bridges the two research agendas by applying counterfactual causal reasoning found in the micro-level tradition to understanding societal consequences of mobility in the macro tradition. Decades of accumulative scientific efforts in the macro tradition have established many important descriptive regularities about social mobility in many national contexts. However, we know little about the influence of mobility patterns on other aspects of a society. The micro tradition of social stratification research has accumulated evidence about the causal effect of social achievements, but the macro implications of the micro estimates have not been

investigated systematically. For example, Brand and Xie (2010) showed that students from disadvantaged family origins have higher gains from college education compared with students from advantaged origins, indicating that a reform of the college system to recruit more disadvantaged students would increase overall social welfare. Micro-level causal inference, though, is not equipped to evaluate the exact magnitude of welfare change induced by different types of reforms, a task SME is designed to accomplish.

SME: DEFINITION, IDENTIFICATION, AND ESTIMATION

Definition of a Mobility Intervention

SME is the effect of a mobility intervention on the population distribution of an outcome. In this section, we discuss the definition, identification, and estimation of SME, followed by a brief discussion of the relevance of SME to the causal inference literature. Here, we will introduce all the main notations before the discussions. We denote Y and $Y(D)$ as the outcome and the potential outcome given treatment D. X denotes covariates (or confounders in the identification of potential outcome) other than social origin O, and L denotes all covariates, which is the union of O and X. ψ is a superscript denoting post-intervention for variables or functions of interest. For example, Y^{ψ} denotes the post-intervention outcome.

In the context of this paper, a "mobility intervention" refers to a hypothetical manipulation of the mobility regime so that individual destinations are redistributed based on origin-specific quotas. A mobility intervention is defined by two components, a macro one and a micro one. The macro component is a hypothetical mobility table whose effect is the objective of research: we want to know what would happen to the outcome at the population level if this social mobility intervention were realized. The micro component concerns the calculation of the

individual propensity scores of receiving treatments after the intervention. Intuitively, the macro component specifies a counterfactual society that we wish to make an inference about, and the micro component refers to what happens to specific individuals in that counterfactual society. Note that for our analysis "intervention" is a different causal concept from "treatment." Whereas a "treatment" happens at the individual level, an "intervention" occurs only at the society level.

A hypothetical mobility table is an externally imposed joint distribution between social origin and social destination $P^{\psi}(O, D)$, whose potential consequence is our primary interest. Here, social origin and social destination refer to important parental and own social characteristics, like educational level or social class. In this proposal, we only consider categorical social characteristics but not continuous ones, like wealth or income.

We focus on mobility interventions that only change relative mobility, i.e., that preserve the marginal distributions of origin and destination. Social mobility scholars separate mobility into two parts, the absolute part and the relative part. Absolute mobility measures the departure of the overall distribution of children's social positions from that of parents' social positions, and it is predominantly driven by changes in social structure. Relative mobility, on the other hand, measures the inequality of opportunity by the association be- tween origin and destination net of changes in social structure. Because relative mobility has been the main focus for social mobility analysts in recent decades (Erikson et al., 1992), we focus our discussion on mobility interventions that only change relative mobility. Our framework, however, can also be used to analyze mobility interventions when structural changes are taken into account.

It has been well established that relative mobility is fully captured by the odds ratios in a mobility table, so a counterfactual mobility table that only changes relative mobility should only change the odds ratios in the observed table but not the marginal origin and destination

distribution. In other words, $P^{\psi}(O) = P(O)$ and $P^{\psi}(D) = P(D)$. Once the marginal distributions are fixed, we manipulate the $(O - 1)(D - 1)$ odds ratios. For illustration, we generate the hypothetical mobility table in one of two ways, either through a "theoretical model" or through a "comparison case." Both ways of generating the table require the use of a log-linear model.

With cell-frequencies in a two-way frequency table as the data input, the log-linear models are used to describe the association between the row variable and the column variable. The general form of a log-linear model is as follows:

$$ln(F_{o,d}) = \lambda + \lambda_o + \lambda_d + \lambda_{o,d}(1)$$

where $F_{o,d}$ denotes the frequency of cell (o, d) in a mobility table, λ is the intercept capturing the frequency of the excluded category, λ_o are the marginal origin effects, λ_d are the marginal destination effects, and $\lambda_{o,d}$ represent the origin-destination associations. As a standard practice, the marginal terms in the log-linear model are saturated because we are uninterested in the marginal distributions, and the researcher chooses ways to specify the association term $\lambda_{o,d}$ depending on substantive understandings about social mobility, where popular methods include the levels model, association model, and crossings model (Goodman and Kruskal, 1979; Hauser, 1980; Xie, 1992; Erikson et al., 1992). If the model is saturated, the $(O - 1)(D - 1)$ association terms $\lambda_{o,d}$ correspond exactly to the $(O - 1)(D - 1)$ odds ratios. In general, log-linear model can be used to express more parsimonious representations of the odds ratios.

The first way of generating a hypothetical mobility table is to specify a theoretical model of the association terms $\lambda_{o,d}$, and use the predicted frequencies of the model to form a

counterfactual mobility table. The simplest theoretical model is total fluidity, when origin and destination are not associated. To estimate the effect of total social fluidity, we specify a log-linear model with the association terms $\lambda_{o,d}$removed and obtain the predicted frequency table from the model; we use this as the intervention ψ in our analysis. More complicated models for the association term $\lambda_{o,d}$ have been developed that link log-linear model results to theoretical conjunctures about social mobility such as homogeneity, hierarchy, affinity, inheritance, and symmetry (Hout, 1984; Hout and Hauser, 1992; Erikson et al., 1992; Bukodi et al., 2017; Zhou and Xie, 2019b). Previous mobility scholars have developed models to facilitate understandings of the landscape of mobility in a society. Using these models as counterfactual interventions, we can analyze the implications of aspects of social mobility on other social outcomes.

Our approach can also be used to make inferences if we hypothetically transport relative mobility from a comparison regime to another, observed regime. Since relative mobility is fully captured by odds ratios in a mobility table, we can extract odds ratios from the reference regime and insert them into the observed mobility table being compared while holding the observed marginal distributions. There are two algorithms to achieve this "extract-and-fit" procedure. The first is to fit a main-effect-only log-linear model with the outcome being the frequencies of the observed data under independence divided by the frequencies of the comparison case. The table formed by the predicted frequencies from this model will have marginal distribution of observed data and associations of comparison case data (Little and Wu, 1991). The second is to use the iterative proportional fitting (IPF) algorithm (Fienberg, 1970). The IPF algorithm takes the comparison case mobility table as the initial point and iteratively adjusts the row and column marginal distributions to approach the marginals of the observed data.

To evaluate the interventional effect, it is necessary to go beyond the knowledge of the counterfactual mobility table $P^{\psi}(O, D)$ and know what happens to treatments of specific individuals. To be more concrete, suppose that a mobility intervention redistributes education opportunities by parental education background. For simplicity also suppose that education is a binary variable, college versus non-college. Then, we may define an intervention that distributes more college opportunities to potential first-generation collegegoers. However, we need to ask: who would be affected by this redistribution policy, i.e., who would now go to college but did not go to college under the old regime, and who would now not go to college but did go to college under the old regime? Among those whose parents did not go to college, it could be that these new opportunities are randomly distributed, or it could be that the opportunities go to students with the highest SAT scores or students satisfying some other criteria. Different selection criteria reflect different hypothetical reforms and likely lead to different inferences about the effect of the intervention.

We propose a simple and reasonable redistribution rule: conditional on origin, the as- sociation between the treatment and the covariates is unaffected by the intervention. While we change the opportunity structure by origin through the intervention (i.e., OD association), the mechanisms through which other individual features affect destination are pre- served. Denote X as all other determinants of college attainment beyond parental college attainment O, such as SAT score, or latent ability. Our specification states that the existing mechanisms leading to college attainment, the association between X and D, are preserved after the intervention, and only the percentage of collegegoers by parental college attainment is manipulated. The specification echoes the well-known maximally maintained inequality (MMI) hypothesis: new resources are normally first distributed among those who are already close to getting the

resources (Raftery and Hout, 1993). Of course, this specification is not the only way of assigning individual-level treatments. Depending on research setting, other scholars may well choose alternative forms of intervention, such as total randomization of resources.

We formally define the specification that maintains the association between X and D in the following way:

Specification (Conditional Odds-Ratio Preservation): for all covariate x, the odds-ratio between x and d conditional on o is unchanged $\frac{P^{\Psi}(d|x,o)}{P^{\Psi}(d'|x,o)} / \frac{P^{\Psi}(d|x',o)}{P^{\Psi}(d'|x',o)} = \frac{P(d|x,o)}{P(d'|x,o)} / \frac{P(d|x',o)}{P(d'|x',o)}$.

Since the association between two variables is captured by the odds-ratio, an intervention with this specification preserves the association between covariates and destination conditional on origin. With the conditional odds-ratio preservation specification, the interventional effect can be expressed as a multiplicative effect on treatment odds, $\frac{P^{\Psi}(d|x,o)}{P^{\Psi}(d'|x,o)} = \delta^{d,d'|o} \frac{P(d|x,o)}{P(d'|x,o)}$, where the multiplier $\delta^{d,d'|o} = \frac{P^{\Psi}(d|o)}{P^{\Psi}(d'|o)} / \frac{P(d|o)}{P(d'|o)}$ is defined by the pre- and the post-intervention mobility table.

The multiplier $\delta^{d,d'|o}$ captures intervention-induced changes in odds of d versus d', that is, the probability of receiving d divided by the probability of receiving d'. When $\delta^{d,d'|o} > 1$, the odds of d versus d' for a person from origin o is increased by the intervention, meaning that people are more likely to end up in d rather than d' because of the intervention; when $\delta^{d,d'|o} < 1$, the odds of d versus d' is decreased, and the person is more likely to end up in d' than d.

There are two advantages deriving from the interventional effect being in the scale of odds instead of probability. The first is that it conveniently preserves the marginal probabilities. Take the simple case when D is a binary variable as an example. The intervention only affects the odds $(P(d \mid o))/(P(d' \mid o))$, but not its marginal distribution. The second is that a shift in odds naturally captures the non-linear effect of intervention on individual probability. Intuitively, those who are very likely and very unlikely to receive a treatment will be less affected by the intervention in the probability scale than who are close to the 50 percent likelihood of treatment. Figure 1 illustrates such non-linear relationship between pre-intervention and post-intervention propensity scores implied by the specification in the setup of a binary treatment. Those whose pre-intervention propensity scores are around 0.5 are most influenced by the intervention, whereas those who are closer to the extremes are less affected.

[Figure 1 about here.]

From the specification, a closed-form solution for the post-intervention propensity score (PIPS) can be derived:

$$P^{\psi}(d|x,o) = \frac{\delta^{d,d'|o} P(d|x,o)}{\sum_1^D \delta^{d,d'|o} P(d|x,o)} \qquad (2)$$

PIPS reflects how a mobility intervention affects the treatment assignment mechanism and is a building block for the identification of SME. Once we know PIPS, the specified mobility intervention will provide no further information that would help us learn about SME. The next subsection discusses how SME can be identified once we know PIPS.

Identification of SME

In the broadest sense, SME is defined as the effect of a mobility regime intervention on the population distribution of an outcome. In this section, we discuss the identification of SME on the mean and dispersion measures. We show that SME is identifiable with two classic assumptions in causal inference with observational data: ignorability and SUTVA. The empirical estimation of SME beyond the mean, however, may require more stringent assumptions, which we will discuss in the next subsection.

The first assumption, ignorability, states that the potential outcomes are independent of treatment conditional on pre-treatment covariates. The ignorability assumption helps link the observed conditional expectations to unobserved expected conditional potential outcomes.

Assumption 1 (*Ignorability*): $Y\ (d)\ \perp\ D|L$ for all $d\ \in\ D$.

This assumption is also called the "no unmeasured-confounder" by some researchers, because it is equivalent to stating that the covariate set L contains all confounders, where a confounder is defined as a covariate variable that is not only associated with the treatment but also causes the outcome. In the empirical setting of SME, D is education or occupation, and Y is an important life outcome like income, happiness, or health. L should include all relevant pre-treatment covariates, which can be roughly classified into three types. The first type includes important demographic information, like gender or race. The second type is early-life situation and achievement. If education is the treatment, then confounders could include standardized test scores, ability test scores, or early-life coresidence with parents. The third type is family

background, which can include any parental feature that is attained before treatment. When we try to identify SME on other distributional statistics beyond the mean, the ignorability assumption required is slightly stronger than is usually needed. The ignorability assumption needed is the conditional independence of the entire distribution, that is, conditional on covariates, the different treatment groups have the same distribution, rather than just the same mean, of potential outcomes.

Assumption 2 (*Stable Unit Treatment Value*): $Y_i(\boldsymbol{d}) = Y_i(\boldsymbol{d'})$ for all $\boldsymbol{d}$ and $\boldsymbol{d'}$ such that $\boldsymbol{d_i} = \boldsymbol{d'_i}$, where $\boldsymbol{d}$ and $\boldsymbol{d'}$ denote vector of treatments of all units and $\boldsymbol{d_i}$ and $\boldsymbol{d'_i}$ denote individual i's treatment.

Another assumption needed to identify SME is the Stable Unit Treatment Value Assumption, or SUTVA. The conventional form of SUTVA is written in Assumption 2, which essentially states that one's potential outcome is defined only by one's own treatment, so that the treatment statuses of other people would not affect one's own potential outcome. For instance, if the treatment is college and outcome is income, SUTVA states that one's potential outcome of going to college is unaffected by the education attainment of any other person, and naturally is unaffected by the proportion of people in the society that go to college. A corollary from SUTVA suitable to our setup is that one's potential outcome is defined only by one's own treatment status but not the treatment assignment regime since the assigned treatment of others does not affect one's own potential outcome. This means that the post-intervention potential outcome is the same as the pre-intervention potential outcome because the intervention only affects treatment

assignments. Formally, this alternative expression of SUTVA states that $Y_i^{\psi}(d) = Y_i(d) = Y_i$ if $D_i = d$ for all $d \in D$.

We first introduce the identification of SME on the mean (SMEM). As in the case of average treatment effect (ATE), we start from the identification of SMEM conditional on covariates. The covariate-specific SMEM is defined as SMEM for a specific covariate stratum, $SMEM(l) = E[Y^{\psi} - Y|l]$. For expositional ease, suppose a setting where the treatment is binary, and the individual potential outcomes when the treatment is 1 or 0 are $Y(1)$ and $Y(0)$. The individual treatment effect is then defined as $\tau = Y(1) - Y(0)$. Denote conditional average treatment effect given x as $\tau(l) = E[Y(1) - Y(0)|l]$. Assuming SUTVA and ignorability, $SMEM(l) = w(l)\tau(l)$, where the conditional ATEs are weighted by $w(l) = P(D^{\psi} = 1|l) - P(D = 1|l)$.

To see why, we start from the following equations that relate potential outcomes to observed outcomes. The equations are derived from SUTVA.

$$Y = Y(0) + D \times \tau$$

$$Y^{\psi} = Y(0) + D^{\psi} \times \tau$$

Based on these two equations, SMEM at the individual level is $Y^{\psi} - Y = (D^{\psi} - D)\tau$. The conditional SME can therefore be written as $SMEM(l) = E[Y^{\psi} - Y|l] = E[(D^{\psi} - D)\tau|l]$. The ignorability assumption means that D and τ are independent conditional on L. Therefore $SMEM(l) = E[(D^{\psi} - D)|l] \times E[\tau|l] = (P(D^{\psi} = 1|l) - P(D = 1|l))\tau(l)$[5].

[5] $P(D^{\psi} = d)$ and $P^{\psi}(d)$ denote the same quantity and will be used interchangeably.

This relationship between SMEM and CATE can be extended to the general situation where there are multiple treatment conditions, handled as the "multinomial treatment" for simplicity. Let $\tau_{d,1} = Y(d) - Y(1)$ denote the treatment effect of category d versus a baseline group ($d = 1$) ($d = 1, \ldots d+$). SMEM(l) with a multinomial treatment can be expressed as follows. Within each covariate stratum, the conditional SMEM is a weighted sum of effects of treatments versus a common reference group ($d = 1$), weighted by the propensity score shifts induced by the intervention.

$$SMEM(l) = E\left[Y^{\psi} - Y \middle| l\right] = \sum_{1}^{D} \left(P\left(D^{\psi} = d \middle| l\right) - P(D = d|l)\right) \tau_{d,1}(l) \quad (3)$$

The treatment effect $\tau_{d,1}(l)$ is identified by SUTVA and the ignorability assumption as $\tau_{d,1}(l) = E[Y|d, l] - E[Y|1, l]$. Therefore, SMEM(l) is identified, meaning all its components are estimable from empirical data.

It is important to discuss the interpretation of SMEM before proceeding. PIPS determines how the mobility intervention, which is a stochastic intervention, would affect individuals. Stochastic interventions are interventions which assign treatments in a presubscribed probabilistic way based on some pre-treatment variables (Mun˜oz and Van Der Laan, 2012; Diaz and van der Laan, 2013; Young et al., 2011; Haneuse and Rotnitzky, 2013; Kennedy, 2019). In contrast to a deterministic intervention, a stochastic intervention does not assign a priori individuals to treatment conditions with certainty, but merely the proportion of people that receive the treatment, conditional on covariates.

In our particular setup, the source of stochasticity is twofold. The first is the indeterminacy from the hypothetical mobility table to the fully specified PIPS. We addressed this indeterminacy with our conditional odds-ratio preservation specification. The second is from PIPS to specific individual treatment. As is true in general, even for deterministic interventions, we are interested only in aggregate quantities of interest for a population or subpopulation resulting from an intervention. Therefore, the correct interpretation of SME is the following: given the specification and the assumptions, SME estimates the expected result of the hypothetical stochastic intervention, where the expectation is taken over the specific realized treatment assignments.

So far we have only considered SMEM, that is, how the intervention might affect the mean of the observed outcome. However, we may not only be interested in the interventional effect on the mean, but also the effect on other distributional features of the outcome. For instance, we may be interested in whether increasing social mobility could decrease income inequality, which can be measured by variance, the Gini coefficient, or the Theil index. More generally, we consider a broader set of distributional effects:

$$SME^{*} = \nu\left(f_{Y^{\psi}}\left(Y^{\psi}\right)\right) - \nu\left(f_{Y}\left(Y\right)\right) \qquad (4)$$

The probability density functions of the observed and post-intervention outcomes are written as $f_{Y}\left(Y\right)$ and $f_{Y^{\psi}}\left(Y^{\psi}\right)$, and $\nu()$ is just a summary measure of the distribution ν :

$F \rightarrow R$. For instance, when ν is the expectation function, SME$*$ is simply SMEM as defined earlier.

The assumptions we have made are sufficient for the identification of SME∗ with an infinite sample size. To see why, note that the post-intervention outcome distribution is a mixture of conditional distributions of potential outcomes, $f_{Y^{\Psi}}(Y^{\Psi}) = \sum\sum P(l)P^{\Psi}(d|l)f_{Y^{\Psi}(d)|l}(Y^{\psi}(d)|l)$. By Assumptions 1 and 2, $f_{Y^{\Psi}(d)|l}(Y^{\psi}(d)|l) = f_{Y|d,l}(Y|d,l)$. Therefore, all quantities in the definition of SME∗ are non-parametrically identifiable.

The empirical estimation of SME∗ beyond the mean, however, is challenging without imposing additional strict assumptions. The reason for the difficulty is twofold. The first is that because that the law of conditional expectations does not apply to more general distributional statistics, we cannot simply aggregate covariate-specific estimands to population-level estimands. The second is that we often need to estimate not just the conditional mean, but other higher moments. We discuss the identification of the interventional effect on variance to illustrate these two points, and in the next section discuss possible assumptions to impose to make the estimation feasible.

The interventional effect on the variance can be written as $SMEV = Var(Y^{\psi}) - Var(Y)$. Given the law of total variance, the effect can be decomposed into two parts, i.e., the effect on within-group variance and the effect on between-group variance,

$$SMEV = (E[Var(Y^{\psi}|L)] - E[Var(Y|L)]) + (Var(E[Y^{\psi}|L]) - Var(E[Y|L])) \quad (5)$$

As mentioned above, the first difficulty of identifying SME∗ is that it cannot be aggregated from conditional SME∗s. Comparing SMEV with SMEM, it is clear that the effect on variance is not simply the aggregate of conditional effects, which is the case in the effect on the

mean, $SMEM = E[E[Y^{\psi}|L]] - E[E[Y|L]]$. Rather, an extra term, $Var(E[Y^{\psi}|L]) - Var(E[Y|L])$, needs to be considered. This is generally the case for distributional statistics ν beyond the mean: that the marginal effect cannot be directly aggregated from the conditional effects. This adds to the complication of the identification of SME∗.

The second difficulty lies in the estimation of higher moments of the distribution. In the case of SMEV, the full identification of Equation 5 is given by the following formula[6].

$$\begin{aligned} SMEV = & \sum\sum\{P^{\psi}(d|l) - P(d|l)\}\{Var(Y|d,l) + E^2[Y|d,l]\}P(l) \\ & - \sum\{E^2[Y^{\psi}|l] - E^2[Y|l]\}P(l) \\ & + (Var(E[Y^{\psi}|L]) - Var(E[Y|L])) \end{aligned} \quad (6)$$

To estimate SMEV , we need to estimate not only conditional probability and mean, $P(D|L)$ and $E[Y|L]$, but also conditional variance, $Var(Y|L)$. It is generally true that, in order to estimate SME beyond the mean, we need to have knowledge about the conditional distribution $f_{Y|L}(Y|L)$ beyond its mean. If we have an estimator for the full conditional distribution, we would be able to estimate all forms of SME∗. The next section will discuss one possible assumption that enables such estimation.

Estimation of SME

The estimation of SME conforms to the common standard in the mainstream causal inference literature. Inspecting Equation 3, we note that the two quantities to be estimated for

[6] Full proof in appendix.

SMEM are the observed propensity score, $P(D|L)$, based on which we can calculate PIPS $P^{\psi}(D|L)$, and the observed conditional expectation $E[Y|D,L]$.

The most straightforward estimation is to run two regressions with D and Y as outcomes separately and plug the estimates in to $P(D|l)$ and $E[Y|D,l]$ to estimate SMEM(l). Denote the estimates from the two regressions as $\hat{P}(D|l)$ and $\hat{E}[Y|D,l]$. Then SMEM(l) is simply estimated as $\widehat{SMEM}(l) = \sum(\frac{\delta^{d,d'|o}\hat{P}(d|l)}{\sum\delta^{d,d'|o}\hat{P}(d|l)} - \hat{P}(d|l))(\hat{E}[Y|d,l] - \hat{E}[Y|d',l])$.

The regressions can be in the form of common parametric models. For instance, $\hat{P}(D|l)$ can be generated by a multinomial logistic regression, and $\hat{E}[Y|D,l]$. can be generated by an OLS. However, estimation via parametric models has been increasingly criticized because it imposes functional constraints that cannot be justified (Van der Laan and Rose, 2011; Chernozhukov et al., 2018; Hines et al., 2021). When a model is misspecified, the expectation and probability estimates are potentially biased, and so is SME. What appears to be a better solution is to estimate the conditional expectation and probability using data-adaptive/machine learning algorithms, such as Lasso or random forest. The apparent advantage of using machine learning algorithms for estimation is that they often impose little to no a priori parametric assumptions about data distribution and are designed to produce the best data fit. Since they produce more accurate estimates of the conditional probability and expectation than parametric models, the causal estimates are also more accurate, as shown in a recent data competition (Dorie et al., 2019). One cost of the non- parametric approach, however, is the requirement that the data set being analyzed be very large with good support over the data space.

The estimation of SMEV and the generally defined SME is more complicated. As the previous section illustrates, a plug-in estimator for SMEV requires an estimator for $E[Y|L]$, an estimator for $P(D|L)$, and an estimator for $Var(Y|L)$. We discussed above the estimators for the

conditional mean and probability, so the only task remaining is the estimation of conditional variance. The most straightforward approach for estimating conditional variance is to employ the fact that $Var(Y|L) = E[Y^2|L] - (E[Y|L])^2$. Apart from an estimator for $E[Y|L]$, we train another estimator of the conditional expectation of Y squared. Then, a conditional variance estimator can be constructed as $\widehat{Var}(Y|L) = \hat{E}[Y^2|L] - \left(\hat{E}[Y|L]\right)^2$ (Yao and Tong, 1994). Otherwise, we can implement the "residual-based estimator" (Fan and Yao, 1998). Using the conditional expectation estimator, we first calculate all individual residuals $Y_i - \hat{E}[Y|l_i]$. Then, we train another model with squared residuals as the out- come, and the mean of predicted squared residuals would be the estimator for conditional variance $\widehat{Var}(Y|L) = \hat{E}[\left(Y - \hat{E}[Y|L]\right)^2|L]$.

If we wish to estimate the generally defined SME$*$, that is the SME on any arbitrary distributional characteristics of the outcome, we need an empirical estimator for the entire conditional distribution of Y on L. Several approaches have been developed by earlier statisticians to accomplish such a task. Among them, we introduce the "regression imputation" approach, as it is intuitive and highly flexible (Juhn et al., 1993). In the regression- imputation approach, after invoking one additional assumption, we are able to impute the exact counterfactual outcomes for all individuals. The one additional assumption invoked is "Conditional Percentile Preservation":

Assumption 3 (*Conditional Percentile Preservation*): For all units $i \in I$ and all possible $d \in D$, $\rho_i(Y_i(d)) = \rho_i(Y_i(d_i))$, where $\rho_i\left(Y_i(d)\right) = F_{Y(d)|l_i}(Y_i(d)|L = l_i)$ is the percentile of the individual i's potential outcome given i among those with the same covariates l_i.

The assumption states that, for all individuals, conditional on their respective covariates, their percentile of potential outcome given any treatment is the same as their per- centile of the observed outcome. Given Assumptions 1 and 2, it follows immediately from Assumption 3 that the conditional percentile of individual potential outcomes is identified as $\rho_i\big(Y_i(d)\big) = F_{Y|l_i,D=d_i}(Y|L = l_i, D = d_i)$. The individual potential outcome can then be imputed by reversing the conditional cumulative distribution function $Y_i(d) = F^{-1}_{Y|l_i,D=d_i}(\rho_i\big(Y_i(d)\big)|L = l_i, D = d)$. The imputed potential outcomes can be used to derive the full distribution of the post-intervention outcome, which is then used to calculate SME$*$.

To put the assumption in substantive terms, suppose that D is education and Y is income. The assumption means that if a collegegoer ranks at the 75th percentile in income among their comparable peers with the same covariates (family background, ability, etc.), then the person would also have the income at the 75th percentile among the comparable peers if they did not go to college.

With this assumption, to impute this person's counterfactual income if he/she did not go to college, we simply place the person at the 75th percentile income among all persons who did not go to college with the same covariates. This assumption is plausible in settings where the influences of the covariates are similar (say, monotonic) across different treatment conditions. For instance, in the case of the college-income effect, if the covariate is ability, Assumption 3 would imply that ability has similar effects on income for both high school and college graduates, therefore the collegegoers who have high income would still have relatively high income if they went to high school. If, however, the abilities required in high school are different from those required in college (Willis and Rosen, 1979), then the assumption is likely to be violated.

To estimate the broadly defined SME$*$ under the Conditional Percentile Preservation Assumption (Assumption 3), we need to further estimate the conditional percentiles and the inverse of the conditional cumulative distribution function, $\rho_i\big(Y_i(d_i)\big)$ and $F^{-1}_{Y|l,d}(.|l,d)$. To estimate them non-parametrically would require us to rank every individual's outcome within people with the same covariates. Empirically, such a task is normally impossible because there would not be a sufficient number of data points within each covariate stratum l. To circumvent the sparsity problem, the simplest solution is to further assume that residuals are independent from the covariates, $Y - \mu(d, l) \perp L$. This way, the individual residuals can be ranked not within the same covariate stratum but within the same treatment group. If this is the case, to impute the individual i's potential outcome $Y_i(d)$, we simply get the percentile of individual i's residual among the entire population τ_i, identify the individual j with treatment d who has the same residual percentile, and impute $Y_i(d) = Y_j$.

Without imposing extra assumptions beyond Assumption 3, we can also choose to coarsen the covariate space. For instance, we can find a way to discretize all covariates into a few categories and estimate the percentiles within each category. One way to achieve this is to use the propensity score. We can split all units into five groups based on propensity score values, calculate within-categories percentiles, and do imputation. Another option is to extract the first component of the covariates using principal component analysis, and then divide the data based on the principal component.

Relevant Literature and Estimands in Causal Inference

The motivating ideas behind SME are relevant to some earlier strands of literature in causal inference. First, SME is a type of interventional effect. In causal inference, an intervention

is defined as a counterfactual way to assign treatments. Conventional causal estimands, like average treatment effect (ATE), assess the consequence of deterministic interventions. For instance, ATE is the contrast of the consequence of two hypothetical deterministic interventions, one assigning everyone to treatment and the other assigning everyone to control. The deterministic intervention setup has some well-known shortcomings. The first is that such target intervention is often unrealistic (Munoz and Van Der Laan, 2012). In the case of the estimation of ATE, sending everyone to the same treatment group is often impractical if not also undesirable. Related to the first point, the second shortcoming is that it requires the positivity assumption. If certain individuals have zero probability of receiving a treatment, the effect of an intervention that assigns these individuals to the impossible group is unidentified. The third shortcoming is that it does not address the uncertainty in the assignment process of treatment. Social reforms are often carried out by changing the likelihood of targeted social groups to receive the treatment rather than directly assigning treatments to individuals (Angrist et al., 1996). Contrasting the deterministic intervention case, a stochastic interventional effect evaluates the consequence of a hypothetical intervention that assigns individual treatments randomly by a mechanism that is specified by the researcher (Munoz and Van Der Laan, 2012; Diaz and van der Laan, 2013; Young et al., 2011; Haneuse and Rotnitzky, 2013; Kennedy, 2019). In SME, the treatment assignment mechanism is given by the mobility intervention, whose main component is the mobility table.

Second, the setup of SMEM resembles the widely known local average treatment effect (LATE), as they both deal with population heterogeneity of responses to an external shock. Following Angrist et al. (1996), we can split a population into four strata based on their differences in responses to an intervention: the ”always-taker,” whose treatment is always 1;

the ”never-taker,” whose treatment is always 0; the ”complier,” whose treatment is 0 if assigned no intervention but 1 if assigned intervention; and the ”defier,” whose treatment is 1 if assigned no intervention but will be 0 if assigned intervention. Assuming away defiers, the instrumental variable estimator recovers the LATE, the average treatment effect only among the compliers.

SMEM, on the other hand, is the normalized difference between LATE among the compliers and LATE among the defiers. By intervention specification, we know a priori whether respondents from covariate stratum l are compliers or defiers. By rearranging the

covariate stratum order so that the compliers are those from the covariate strata $L \in (1, l')$, and the defiers be those from the other covariates strata, we have $SMEM = \sum_1^{l'}(P(D^{\psi} = 1|l) - P(D = 1|l))\tau(l)P(l) - \sum_{l'+1}^{L}(P(D = 1|l) - P(D^{\psi} = 1|l))\tau(l)P(l))$. In general, SMEM is the difference between the weighted conditional LATEs between the compliers and the defiers. In the case where the propensity score shift is constant within-stratum, the difference is weighted by the propensity score shift.

The comparison between SMEM and LATE illuminates that SMEM reflects the population consequence of treatment effect heterogeneity. Since the total number of treatments is fixed in our setup, whether there is an interventional effect depends on whether the average treatment effect among compliers is different from the average effect among the defiers. In the case where the units who already receive treatment have larger treatment effects than those who would receive treatment under the intervention on average, the intervention effect would be negative. Whereas when the units who might be induced by the intervention into treatment have a larger average treatment effect, the intervention effect would be positive (Brand and Xie, 2010).

The final two strands of the literature that are relevant to our approach are policy-relevant treatment effect (PRTE) and causal decomposition. PRTE is the potential effect of hypothetical policies that take into account of self-selection into policies (Heckman and Vytlacil, 2001, 2005; Carneiro et al., 2010; Zhou and Xie, 2019a). SME can be understood as one type of policy-relevant treatment effect where the policy is to redistribute social destination based on social origin. We contribute to the PRTE literature by discussing the consequence of a new type of policy unconsidered in earlier literature. Causal decomposition evaluates the disparity-reduction power of hypothetical interventions (Jackson and VanderWeele, 2018; Jackson, 2020; Lundberg, 2020, 2021; Zhou and Pan, 2021; Yu and Elwert, 2021). The decomposition framework usually only considers simple interventions like the total randomization of treatments. Our framework allows for much more complicated and theoretically grounded specifications of interventions.

APPLICATIONS: SME ON AVERAGE FERTILITY AND INCOME INEQUALITY

Application 1: SME on Average Fertility

We present two examples to illustrate respectively SME on the mean and on dispersion measures. First, we illustrate SME by analyzing the effect of social mobility on aver- age fertility using data from the General Social Survey. The mobility-fertility hypothesis, positing a relationship between individual fertility and mobility status, has been extensively discussed in the literature on mobility effects (Blau and Duncan, 1967; Sobel, 1985; Luo, 2021). We reexamine this hypothesis not by asking whether mobility status affects fertility at the individual level, but by asking whether social mobility affects average fertility. To preview our results, using data from GSS, we find a small and insignificant negative macro mobility effect on

fertility. More specifically, our results show that if the society is completely fluid, the average fertility would decrease by about 0.01. The net effect is negative because the fertility increases by those who are pushed downward is counterbalanced by the fertility decrease by those who are pushed upward in a hypothetical mobility intervention.

Earlier studies hypothesized that mobility status, net of the influence of origin and destination, has an impact on individual fertility. The stress hypothesis predicts that mo- bile individuals have lower fertility because they are detached from their social origin. The detachment from social origin means lower social support, higher life pressure, and worse subjective well-being, which all cause low fertility (Sorokin, 1927; Duncan, 1966). The economic hypothesis also predicts a negative mobility effect on fertility for a similar reason. It posits that mobile individuals are under not only greater psychological but also greater economic stress compared with non-mobile people. An upwardly mobile individual needs to maintain status, whereas a downwardly mobile individual needs to regain status, so they have less resources and energy to raise children (Westoff, 1953). Finally, the isolation hypothesis predicts that mobility has a positive effect on fertility (Ellis and Lane, 1963). Scholars holding this opinion agree with the observation that mobile individuals are under greater stress, but they think that as a response to stress, mobile people tend to raise more rather than fewer children for emotional support.

Hypotheses about individual mobility effects cannot be directly tested, but the theoretical considerations behind these hypotheses can be viewed from the perspective of treatment effect heterogeneity, which is testable. For instance, assuming that the treatment is education, which has been widely shown to have a negative effect on fertility, stress and economic hypotheses imply that this negative treatment effect is larger for people from lower-SES backgrounds

because they are under more psychological and economic pressure than peers from more advantaged origins.

As described in Section 4.4, SME reflects treatment effect heterogeneity, so hypotheses about treatment effect heterogeneity by family origin are naturally hypotheses about macro-level mobility effects. For instance, if the stress/economic hypotheses are true, giving more education opportunities to lower-SES background people, which is equivalent to increasing mobility, would decrease fertility. On the other hand, if the isolation hypothesis is true, increasing mobility by giving better education to lower-SES background people would increase total fertility.

To test these hypotheses, we use data from the General Social Survey (GSS) between 1972 and 2018. To facilitate comparison with results from earlier studies, we only included white males over 40 years old in our sample. Our social destination or treatment measure is education, which is coded into three categories, less than a high school diploma, a high school diploma, and a college diploma. Our origin measure is also education, following the same categories for destination. The control variables, which we assume to include the full set of confounders of the relationship between own education and fertility, include age, number of siblings, ethnicity, type of residence at the age of 16, region of residence at 16, self-rated family income at 16, religion in which raised, religious environment in which raised, coresidence with parents at 16, parental employment status, whether parents are immigrants, father's education, and mother's education. After excluding missing values in all variables reported above, there is a total of 5,152 respondents. Since most control variables are categorical, we use a random forest to estimate the conditional expectation. Random forest is a tree-based prediction algorithm. Intuitively, what it achieves is similar to a regression where the independent variables are fully interacted, but the

interactions are regularized, meaning that uninformative interactions are automatically dropped. There are 5,152 observations in our analysis. All standard errors are bootstrapped.

Figure 2 shows estimates of individual-level treatment effects of education on fertility. Both high school completion and college completion have negative effects on fertility, but the effect of college is smaller. However, the effects reported are insignificant, suggesting that the association between education and fertility is driven mostly by selection. Moreover, there is no detectable treatment effect heterogeneity by family background. The results suggest that marginal educational expansion, particularly a further expansion of opportunities to attain high school education, is likely to reduce fertility. The effect of redistributing education opportunities based on family origin, on the other hand, is likely to be small, because the effect of education is similar for people with different father's education.

[Figure 2 about here.]

The lack of treatment effect heterogeneity implies that there is no interventional effect. We report the SME of the intervention imposing complete independence of father's education and children's education. This is the most extreme type of intervention, in that the association between father's and children's education is set to zero. Figure 3 shows the effects of this intervention. The population is divided into two halves, those who have better education opportunities under the intervention (gainers), and those who have worse opportunities (givers). Among the gainers, their average fertility is decreased by about 0.13. Among the givers, the positive effect is about 0.11. As implied by results from Figure 2, the treatment effects among the gainers and among the givers are similar in magnitude, albeit in opposing directions. The total

effect is the weighted difference of the two separate effects, which is slightly negative and insignificant. The results suggest that the intervention is likely to cause a negligible average effect on fertility in the population.

[Figure 3 about here.]

We may be interested not only in the interventional effect on average fertility, but also fertility difference between social groups, the "gap-closing estimand" (Lundberg, 2021). Figure 4 reports the expected fertility by father's education and own education before and after the intervention. The left panel shows that the fertility gap by father's education is de- creased by the intervention. Before the intervention, people whose father did not graduate from high school have significantly higher fertility than others. After the intervention, the gap becomes much smaller. The right panel demonstrates the interventional effect on the observed gradient of own education on fertility. Similar to the gradient of father's education, those who attain less than a high school diploma have a significantly higher fertility than others. After the intervention, the gap is largely closed.

[Figure 4 about here.]

Application 2: SME on Income Inequality

Researchers may be interested in the mobility effect on inequality. For example, the empirical negative association between income inequality and intergenerational mobility, a conjecture commonly known as the "Great Gatsby Curve" (Krueger, 2012), has piqued much

scholarly and public interest in recent years. Prior theoretical and empirical evaluations of the phenomenon have tried to unpack the causation in one direction, that is, how inequality causes mobility (see DiPrete (2020) for a recent review). An alternative theoretical possibility, that social immobility is a cause of income inequality, has received scant attention.

We hypothesize social immobility to be a cause of inequality because of the additive and interactive origin and destination effects on income. First, when origin effect and destination effect are additive, social immobility creates a polarized income distribution. Immobility means that family origin and social destination are highly associated so that both high-origin/high-destination and low-origin/low-destination individuals are common. Numerous studies in sociology have reported that family origin has a "long-arm" on children's outcomes, net of its influence on children's education or occupation. Since family origin has a direct effect on personal outcome, those who come from a high-origin/high- destination background enjoy a "double status advantage," whereas the low-origin/low-destination population face a "double status penalty". Therefore, social immobility has a potential polarizing effect that increases the gap between the disadvantaged subpopulation and the advantaged subpopulation.

Second, when origin effect and destination effect are interactive, increasing social mobility may decrease inequality by allocating opportunities to those who benefit the most. Increasing social mobility means more people from disadvantaged social origins end up in high-status destinations, like getting a college degree. The benefit of a college degree may differ for different population groups. In some cases, people who come from a disadvantaged background benefit more from college (Brand and Xie, 2010). When there is such "negative selection" into college, increasing mobility can reduce inequality, because the disadvantage of lower-origin students can to a large extent be compensated by their college degree.

We use data from the Panel Study of Income Dynamics (PSID) to estimate the impact of social mobility on income inequality. PSID is a nationally representative survey of more than 5,000 families in the United States since the 1960s. It has been widely used in studies of income because of its comprehensive collection of relatively reliable income information. We use samples of individuals that were active in the workforce in any wave of PSID between 2009 and 2019. Family origin and destination are measured by education in three categories: less than a high school diploma, a high school diploma, and a college diploma. The outcome variable is the average full-time labor income between 2009 and 2019. There are 12,238 observations in our analysis.

There are two groups of control variables. The first group includes demographic characteristics such as race and gender. The second group indicates childhood conditions including childhood mental health, juvenile substance abuse, childhood mental health, childhood emotion, childhood residential region, childhood economic situation, and co- residence status with parents. Missing values in the control variables are imputed. Both men and women are included in the analysis.

Again, we report the effect of a mobility intervention that sets family origin and destination to be independent. This estimate provides an upper bound of what a mobility intervention can possibly achieve. We report the mobility effect on three types of dispersion measures, the variance of logged income, the percentile ratios, and the Gini coefficient. For the estimation of the effect on all four measures, Assumption 3 is invoked to make the empirical estimation tractable. The estimator for all conditional expectations is linear regression.

Figure 5 reports the treatment effect of education by family origin. The issue of the treatment effect heterogeneity of college on income by family origin has long been de- bated.

While many find that children from disadvantaged backgrounds benefit more from college than peers from advantaged backgrounds (Brand and Xie, 2010; Bowen and Bok, 2016; Attewell et al., 2007; Dale and Krueger, 2002), another group of studies report that it is advantaged children who benefit more from college (Willis and Rosen, 1979; Carneiro et al., 2001, 2003; Heckman et al., 2006). A more recent article reported a curvilinear relationship between family origin and college effect (Cheng et al., 2021). According to our results, for respondents from all family origins, the effect of a high school diploma versus having less than a high school diploma is slightly larger than the effect of a college diploma versus a high school diploma. There is no significant treatment effect heterogeneity of a high school diploma among respondents of different origins. The effect of a college diploma, however, is largest for people whose fathers have a college diploma. This interaction between origin and destination effect potentially undermines the inequality-reduction power of a mobility intervention, because people from less advantaged backgrounds do not benefit as much from education opportunities.

[Figure 5 about here.]

Figure 6 presents the social mobility effect on the distribution of logged income. The imputation of specific income values is enabled by invoking Assumption 3 as described above. The first graph of Figure 6 shows that the average logged income for the whole population is barely affected by the intervention. Since SME is only present when there is treatment effect heterogeneity by family origin, the limited treatment effect heterogeneity reported earlier means a negligibly small SME on average income. The following three panels show SME for people from three family origins respectively. For people whose fathers did not graduate from high

school, their income is significantly increased by the mobility intervention. SME on those whose fathers have a high school diploma is insignificant, and on those whose fathers have a college diploma it is negative and significant[7]. As far as dispersion, it is clear that both the overall and the conditional income distributions are more concentrated near the center after the intervention.

[Figure 6 about here.]

Figure 7 reports the statistical test results of the interventional effect on the dispersion of income. Both the variance of logged income and the Gini coefficient are decreased by the intervention, but the effects are not statistically significant. The 90/10 ratio of income is also not significantly affected. By contrast, the 80/20 ratio is significantly decreased by the intervention, from 5.6 to 4.8. In general, the effect of the mobility intervention on income dispersion is inconclusive.

[Figure 7 about here.]

Figure 7 reports the effect of an extreme intervention that sets father-children association to zero. A researcher might be curious about the effect of less radical interventions and how the effect varies with the strength of the intervention. Figure 8 shows that intervention effect increases roughly linearly with intervention strength. To define intervention strength, we borrow the traditional "unidiff" model (Xie, 1992): $ln(F_{o,d}) = \lambda + \lambda_o + \lambda_d + \varphi\lambda_{o,d}$, where φ describes the extent of regime-specific departure of observed association between O and D from

[7] The significance tests of the difference-in-means are not shown in the Figure.

the independence model, normalized between 0 and 1. We define intervention strength as $(1 - \varphi)$. Thus, intervention strength varies between 0 (no intervention) and 1 (complete independence). When intervention strength is 1, all post-intervention log odds ratios are zero.

[Figure 8 about here.]

DISCUSSION AND CONCLUSION

Students of social stratification have long been interested in the consequences of social mobility. However, substantive research on the topic has been hindered by methodological obstacles. Acknowledging the unidentifiability of individual-level mobility effects, this paper proposes to shift the focus from assessing consequences of individual-level mobility to consequences of social mobility at the societal level. We discussed the definition, identification, and estimation of the consequences of social mobility on both the average and the dispersion of the population distribution of an outcome. We illustrated the method by analyzing the mobility effect on average fertility and income inequality in the United States. We find little effect on fertility and an inconclusive effect on different measures of income inequality, but between-group disparities of both fertility and income are significantly reduced by the intervention.

Our proposal, SME, bridges the macro and the micro traditions of social stratification research. Substantively, the macro tradition investigates the associations of social positions in a stratification system, and the micro tradition studies the linkages between social position and key social outcomes. Methodologically, the macro tradition employs categorical data analysis techniques, especially log-linear models, and the micro tradition usually conducts research under a causal inference framework. SME links the two traditions substantively by asking the question

of what would happen to key social outcomes if we change the association of social positions. It also links the two traditions methodologically because it investigates the causal effect of interventions that are specified by log-linear models.

Our macro-level solution to the unidentifiability issue of the mobility effect may also be borrowed in methodological settings that share a similar problem. For instance, status inconsistency theory predicts that the inconsistency of social statuses along different dimensions has an independent influence on an individual social behavior net of the influence of social statuses (Lenski, 1954, 1956; Stryker and Macke, 1978). The assessment of this theoretical hypothesis is hindered by the same unidentifiability issue that we reviewed in the context of mobility effects. Though the individual-level consequence of status inconsistency is unidentifiable, the society-level consequence is identifiable. One can specify an intervention that reflects a hypothetical level of status inconsistency in a society and evaluates its consequence on social outcomes. However, the macro interventional approach may be inappropriate for some other settings, like the APC problem. The reason is that APC is an inherently descriptive task, and it is hard to imagine how to manipulate any of the three APC dimensions. Also, the dimensions of APC are logically simultaneous so it is unclear which association should we manipulate, unlike the mobility effect, where mobility is the logical byproduct of origin and destination, or status inconsistency, where the inconsistency is the byproduct.

Though our article has only discussed interventions relating to social mobility, i.e., redistributing resources by family origin, it is obvious that the framework is also applicable to hypothetical redistributions based on other characteristics. For instance, we may be interested in the social consequence of reducing occupational segregation by gender or by race (Lundberg, 2021). One important form of knowledge accumulation in social stratification and demographic

research is the association between important demographic features. Under the SME framework, we can reach one step further and evaluate the consequences of hypothetical interventions that change such associations.

The intervention effect approach also poses new methodological challenges. Unobserved confounders, the biggest problem that causal researchers of observational data face, are still an issue for intervention-based causal inference. Failure to control for all confounders compromises the accuracy of treatment effect estimation and therefore intervention effect estimation, because the latter is a weighted version of the former. The often-overlooked identification assumption SUTVA also needs to be handled with caution in the intervention effect setup. SUTVA essentially assumes an insensitivity of potential outcomes to the environment, but a large-scale social intervention is likely to change the environment. For instance, SUTVA implies that individual college returns are unaffected by how many and what kind of people enroll in college, which is unrealistic. The methodological work on the correction and sensitivity analysis of SUTVA violation is lacking, and we see this as a potential next step of research that contributes not only to our approach but also the causal inference literature in general.

Asking causal questions about the effect of macro interventions instead of micro treatments opens up new, exciting opportunities for both the substantive and theoretical development of social stratification research. Heterogeneity in causal effects is ubiquitous in social science (Xie et al., 2012). The mainstream causal estimand, average treatment effect by subgroup, can be informative in revealing such heterogeneity but is oftentimes not theoretically interesting because it calculates the contrast of impractical interventions. More theoretically- and policy-relevant quantities of interest can be attained with a reweighting of heterogeneous individual treatment effects with different macro-level interventions on individual treatment

probability. This approach allows the researcher to evaluate the consequences of hypothetical interventions that are flexible and realistic.

The consequences of mobility have been a perennial topic in the field of social stratification. Echoing earlier concerns over the individual-level mobility effect approach, we explicated the unidentifiability of the individual-level mobility effect. Consequently, we call for further attention from methodological and mobility scholars on the societal mobility effect, which is well defined and also closely aligned with classical sociological theories about social mobility. In this paper, we provide a preliminary framework as well as its illustration for studying the effect of social mobility at the societal level. We encourage future works to expand and apply this framework to study the mobility effect on social outcomes in different contexts.

APPENDICES

A. PROOF FOR THE IDENTIFICATION OF SOCIAL MOBILITY EFFECT ON VARIANCE

The identification of SMEV is derived most using the law of mixture distributions and Assumptions 1 and 2. SME on variance is defined as: $SMEV = Var(Y^{\psi}) - Var(Y)$. The first step is to decompose the two components with the law of total variance, which leads us to Equation 5, $SMEV = (E[Var(Y^{\psi}|L)] - E[Var(Y|L)]) + (Var(E[Y^{\psi}|L]) - Var(E[Y|L]))$. The second part of the equation is identifiable as we have already talked about the identification of the post-intervention conditional mean. What remains unidentifiable is the post-intervention conditional variance, that is in the first part.

First, by definition of variance, $Var(Y^{\psi}|l) = E[Y^{\psi 2}|l] - \left(E\left[Y^{\psi}\middle|l\right]\right)^2$. Given the law of the moment of mixture distributions , we have that $Var\left(Y^{\psi}\middle|l\right) = \sum P^{\psi}(d|l)E[Y^{\psi 2}(d)|l] - \left(E\left[Y^{\psi}\middle|l\right]\right)^2$. Now, we focus on the term $E[Y^{\psi 2}(d)|l]$. Using the definition of variance again, we have $E[Y^{\psi 2}(d)|l] = Var(Y^{\psi}(d)|l) + \left(E\left[Y^{\psi}(d)\middle|l\right]\right)^2$. Organize these results a bit we have that:

$$Var\left(Y^{\Psi}\middle|l\right) = \sum P^{\Psi}(d|l)\{Var(Y^{\Psi}(d)\}l) + (E\left[Y^{\Psi}(d)\middle|l\right])^2\}$$

Among the terms, $P^{\psi}(d|l)$ is known by specification. The identification $E\left[Y^{\psi}(d)\middle|l\right] = E[Y|d,l]$ is done with ignorability and SUTVA, and the identification $Var(Y^{\psi(d)}|l) = Var(Y|d,l)$ is done with ignorability and SUTVA. Organize the equation yields the result as in the main text.

B. RELATION OF CONSTANT CONDITIONAL EFFECT ON ODDS SPECIFICATION WITH DISCRETE CHOICE THEORY

The specification of Constant Conditional Effect on Odds can also be interpreted from discrete choice theory. Discrete choice theory is a set of theories that tries to connect mathematical and statistical models for discrete outcomes and theories about rational human behavior (Luce, 1959; Mcfadden, 1973; McFadden, 2001). Under discrete choice theory, the linear predictor of a GLM for discrete outcome $\beta^T X$ is interpreted as the structural part of a latent utility function that determines individual choice. The conventional latent individual utility of choosing an alternative d is denoted as $U_d = \eta_d + \varepsilon_d$ where η_d is the structural utility described by a linear term and ε_d is the stochastic unobserved heterogeneity.

Following Heckman and Vytlacil (2005), we modify the latent utility into latent tendency by adding a resistance or constraint term νd to the function $U_d = \eta_d - \nu_d + \varepsilon_d$. Intuitively, the latent tendency function means that the tendency of choosing an alternative d is affected by the utility of choice d that can be explained by observed features, constraint that prevents one from obtaining choice d, and latent individual heterogeneities. Individual probability of choosing an alternative d over d' equals the probability that $U_d > U_{d'}$, that is, $P(d|D = d, d') = P(U_d > U_{d'})$. Assuming that the error terms εd both before and after the intervention follow a Gumbel distribution, the probabilities have the IIA property $\frac{p(d)}{p(d')} = e^{(\eta_d - \nu_d) - (\eta_{d'} - \nu_{d'})}$. Also, assuming that the utility functions are unaffected by the intervention, $\eta_d = \eta_d^{\psi}$, Assumption 1 leads to $\delta^{d,d'|o} = e^{\left(\nu^*_{d'|o} - \nu^*_{d|o}\right) - \left(\nu^*_{d'} - \nu^*_{d|o}\right)}$. Arranging the equation slightly renders that $\nu^*_{d|o} - \nu^*_{d'|o} = \nu_{d|o} - \nu_{d'|o} - log\delta^{d,d'|o}$.

In other words, under Assumption 1 and associated assumptions about latent tendency, the intervention induces a common constrant structural constraint to people from the same family origin. The intervention does not affect the extent to which people desire certain choices but imposes structural constraint on people's ability to pick those choices. Conditional on origin, such constraint induced by intervention is the same constant for individuals. When $\delta^{d,d'|o} > 1$, then the structural distance between d and d' is smaller, and it becomes easier to choose d compared to choose d'.

C. MULTIVARIATE INTERVENTION SPECIFICATION

The setup can be generalized to scenarios where we wish to redistribute our treatments D based not only on one variable but by a set of variables. For instance, in the case of education mobility, SME sets quota of educational opportunities based on parental education background, but we may also want to have more detailed redistribution rules that also takes into account gender or race. One special case of extra condition is observed treatments, that is we may want to redistribute treatment based on both family origin and observed treatment in some cases. To see why setting observed treatment as a condition can be desirable, note that there are two possible ways of understanding an intervention. The first is that we take back all resources that people have already obtained, and then redistribute them. The second is that we achieve the redistribution goal only by switching the resources of some units, leaving the attainment of other units unaffected. The setup where only origin is set as a condition simulates the first scenario, whereas to simulate the second scenario we need to set observed treatment as a condition as well. That is, we need to specify how we wish to distribute treatment considering both family origin and observed treatment. See the appendix for details about how an intervention that considers observed treatment is specified.

The notational change is modest if we wish to redistribute based on more than one conditions. Denote O as a set of variables based on which we wish to redistribute treatment. The observed treatment D can be a part of O as well. Then the specification is generalized as:

$$\text{Generalized Specification: } \delta^{d,d'|o} = \frac{P^{\psi}(d|o)}{P^{\psi}(d'|o)} \Big/ \frac{P(d|o)}{P(d'|o)} = \frac{P^{\psi}_{i\in\{o_i=o\}}(d)}{P^{\psi}_{i\in\{o_i=o\}}(d')} \Big/ \frac{P_{i\in\{o_i=o\}}(d)}{P_{i\in\{o_i=o\}}(d')}$$

Correspondingly, the solution for PIPS is:

$$P^{\psi}_{i\in\{O_i=o\}}(d) = \frac{\delta^{d,d'|o} P_{i\in\{O_i=o\}}(d)}{\sum \delta^{d,d'|o} P_{i\in\{O_i=o\}}(d)}$$

The difficulty, however, lies in the specification of the intervention even when O only contains a small number of variables, because now the researcher needs to supply the multivariate probability distribution $P^{\psi}(D, O)$. Recall that in the case where O is unidimensional, to construct an intervention, we need to supply the marginal distributions of O and D, which are assumed to be constant, and the joint distribution of O and D. If O is n-dimensional, we need to supply all the two-way, three-way, and up to n-way interactions, which is a total of $\sum_i \binom{n}{i}$ tables, most of them high-dimensional. One realistic solution is to give up all high-dimensional tables and instead supply a set of two-way tables, that is there is no interaction in the distributional rule. If O includes parental education and gender, no interaction implies that we redistribute resources based on parental education in the same way for male and female. The two-way tables can be generated based on theory or empirical data as described in Section 4.1, and once the tables are generated, the joint distribution $P^{\psi}(D, O)$ can be generated using the log-linear model following the "extract-and-fit" procedures laid out in Section 4.1.

FIGURES

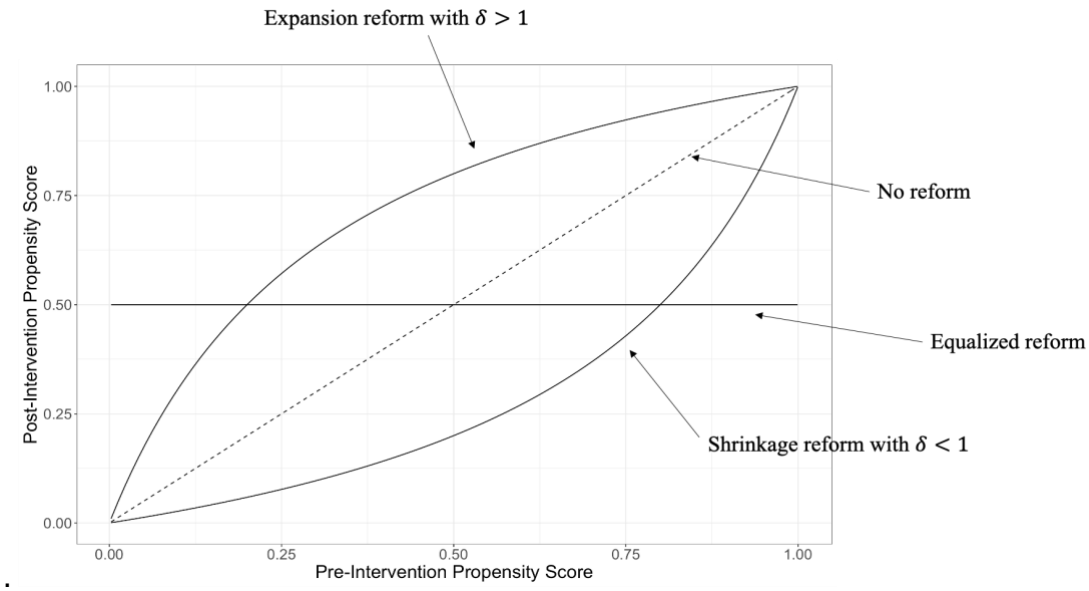


Figure 1: Pre-Intervention Propensity Score and Post-Intervention Propensity Score in Equalized Reform and Realistic Reform with Conditional Odds-Ratio Preservation Specification

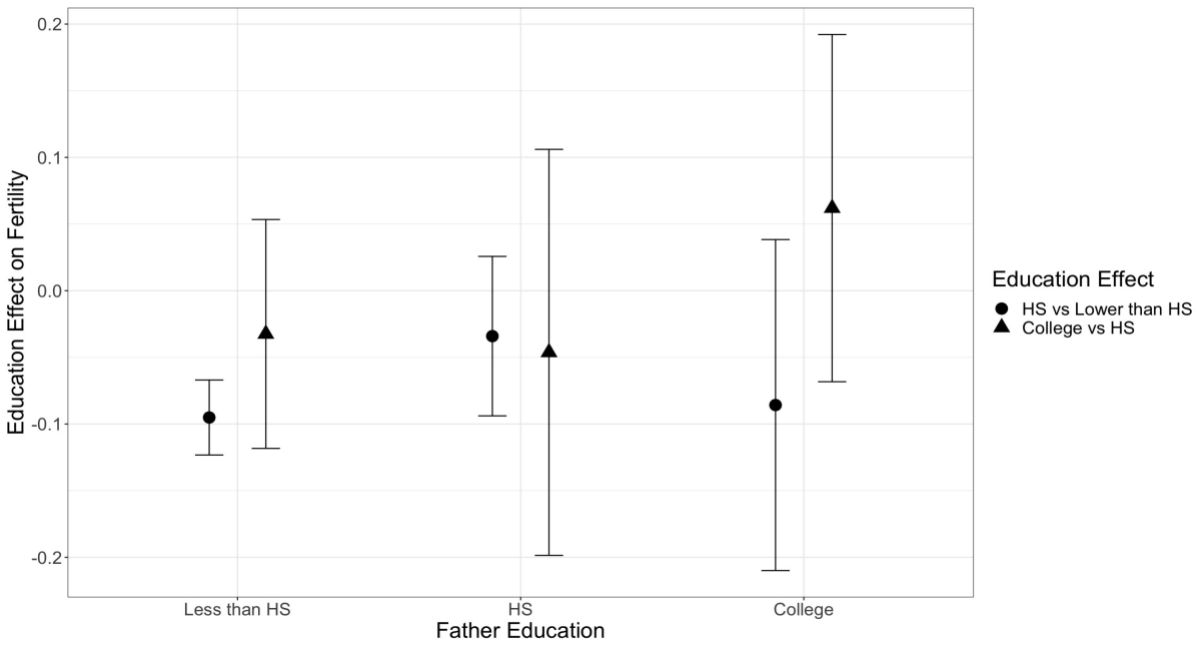


Figure 2: Treatment Effects of Own Education on Fertility by Father Education

Note: GSS 1972-2018. Bars represent bootstrapped 95 percent confidence intervals. Education effect estimated as expected group difference predicted by a random forest.

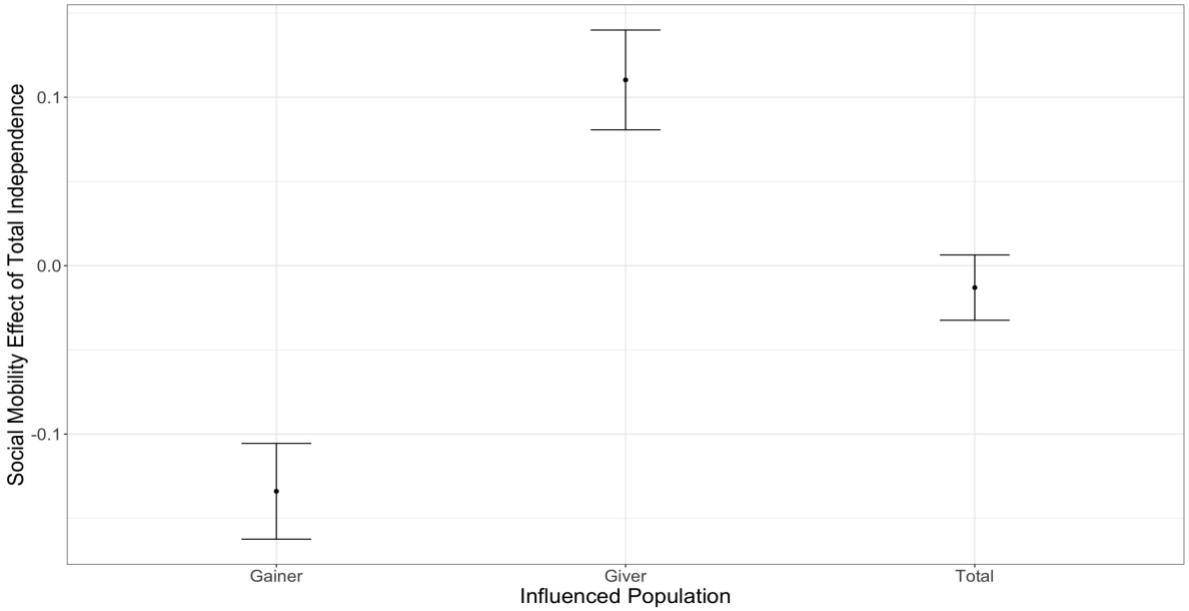


Figure 3: Social Mobility Effect of Total Equalization on Fertility by Population

Note: GSS 1972-2018. Bars represent bootstrapped 95 percent confidence intervals. The estimates reflect the mobility intervention where children and father education are set to be independent. “Gainer” refers to the population group whose education would be in- creased by the intervention; “giver” refers to the population group whose education would be decreased by the intervention.

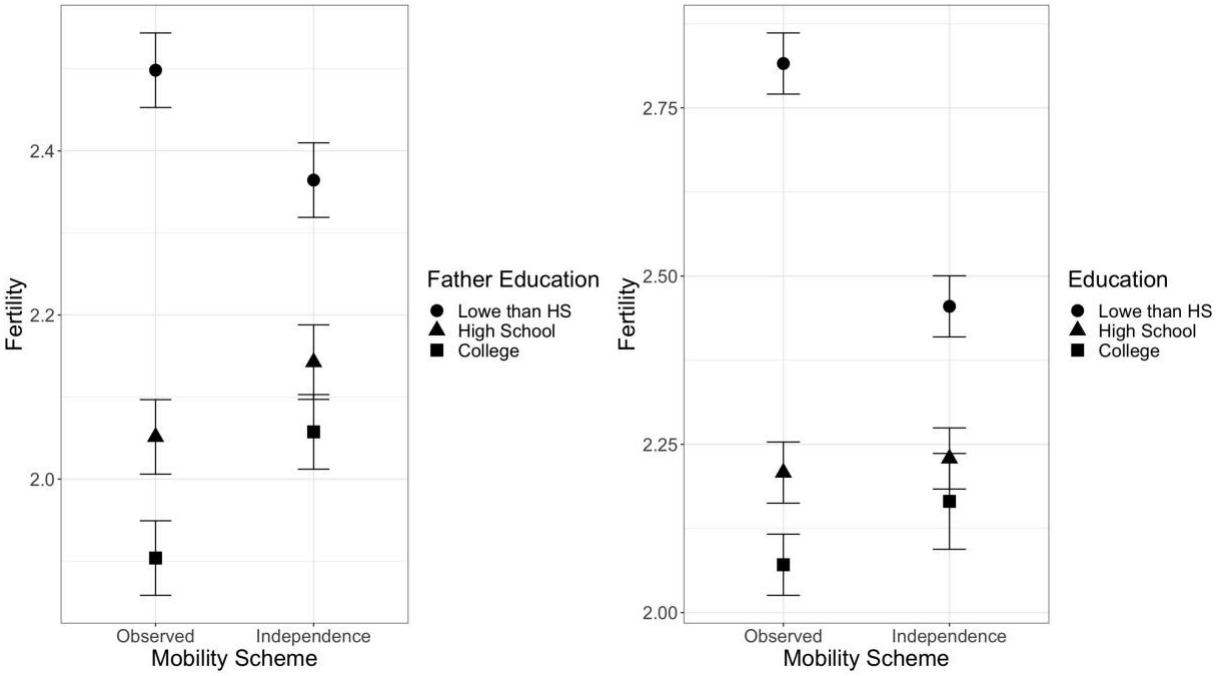


Figure 4: Post-Intervention Education and Father's Education Gradient in Fertility

Note: GSS 1972-2018. Bars represent 95 percent confidence intervals. The "observed" column shows observed group-specific average fertility. The "independence" column shows counterfactual group average fertilities under the intervention that sets father-child education association to be independent.

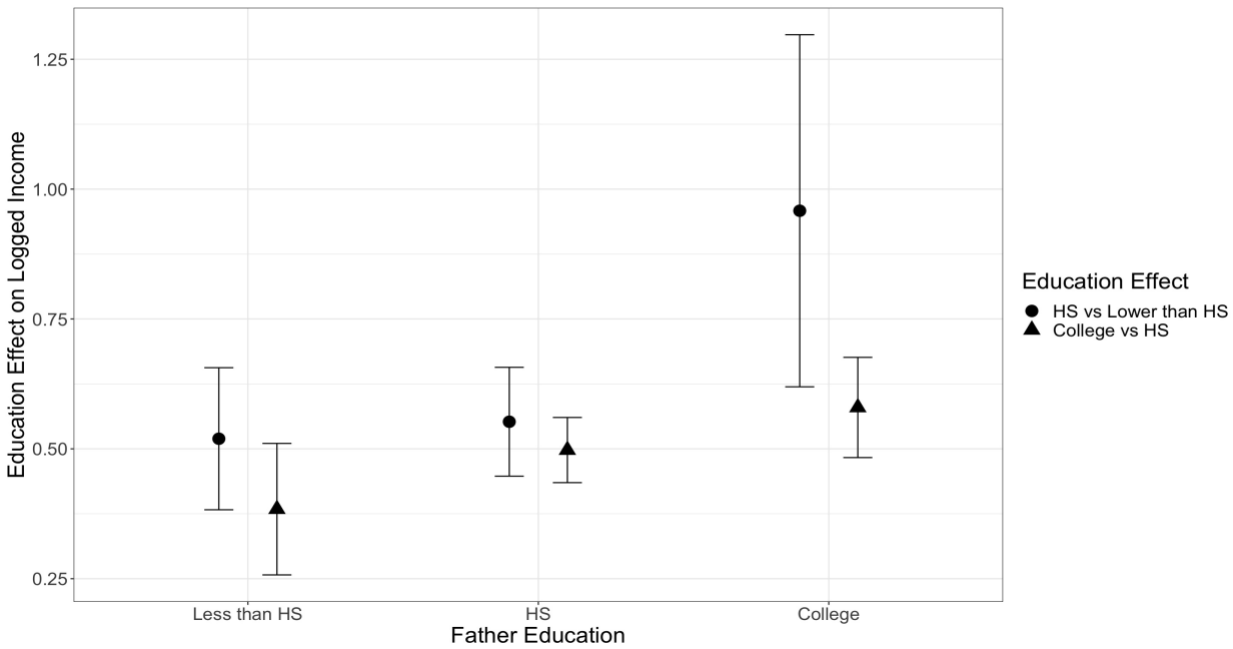


Figure 5: Education Effect on Logged Income by Father's Education

Note: Panel Studies of Income Dynamics 2009-2019. Bars represent 95 percent confidence intervals. Education effect estimated as expected group difference predicted by the OLS model.

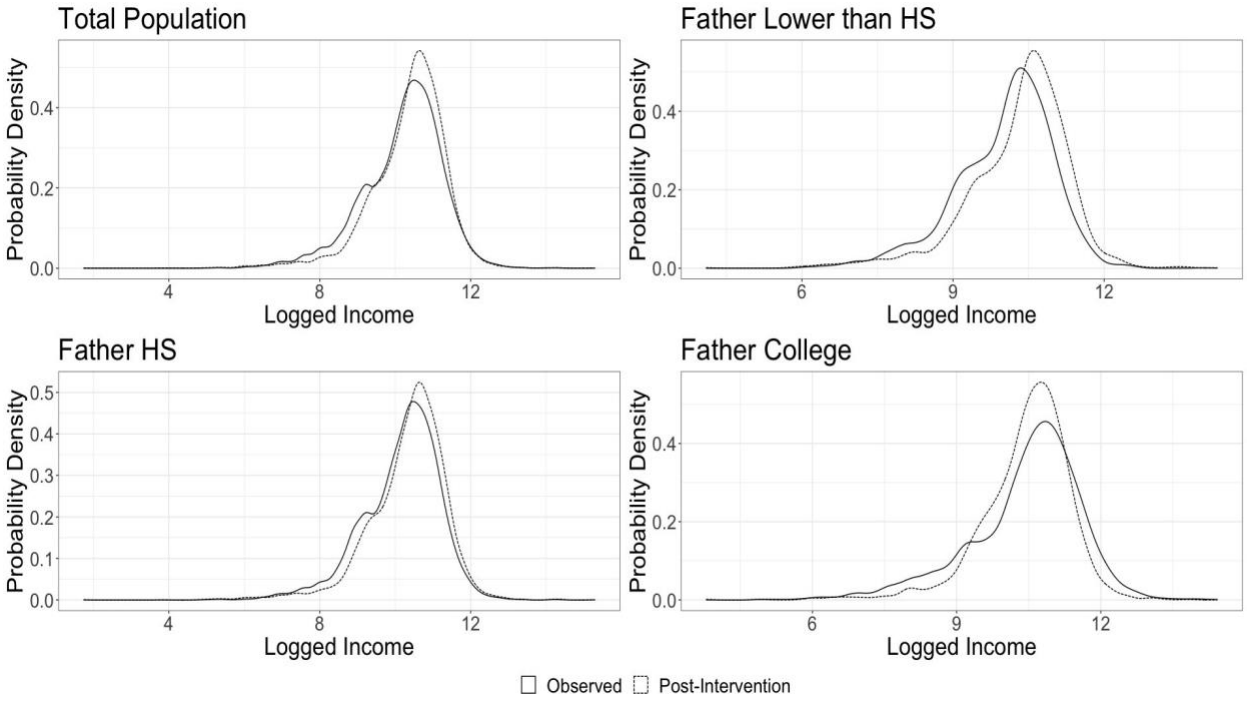


Figure 6: Social Mobility Effect on the Distribution of Logged Income by Father Education

Note: Panel Studies of Income Dynamics 2009-2019. Different panels show conditional effect by father education. The intervention is to set father-child education to be independent. All income values are imputed after invoking Assumption 3 as described above.

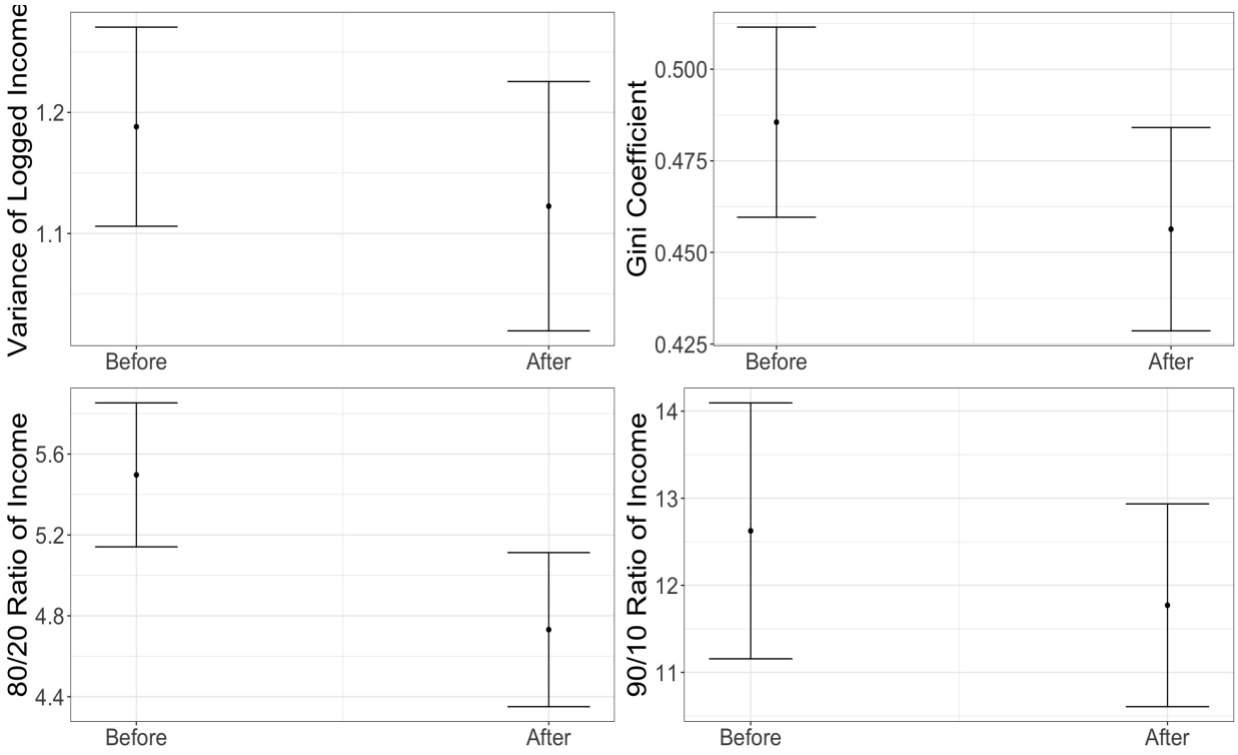


Figure 7: Social Mobility Effect on Dispersion of Income

Note: Panel Studies of Income Dynamics 2009-2019. Bars represent 95 percent confidence intervals. Different panels denote conditional effect by father education. In each column, the left bar is pre-intervention, the right bar is post-intervention. Dispersion measure effects are estimated with extended Assumption 3.

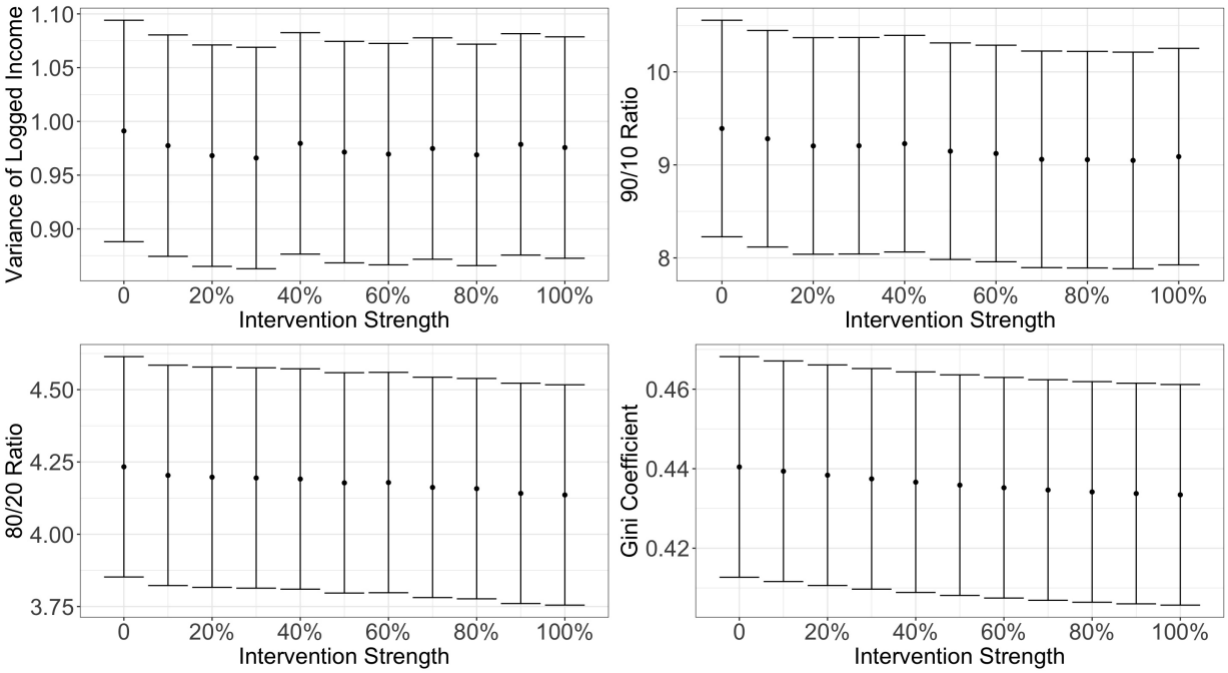


Figure 8: The Relation Between Post-Intervention Dispersion Measures of Income Distribution and Strength of Intervention

Note: Panel Studies of Income Dynamics 2009-2019. Bars represent 95 percent confidence intervals. To define intervention strength, we borrow the traditional "unidiff" model (Xie,1992): $ln(F_{o,d}) = \lambda + \lambda_o + \lambda_d + \varphi\lambda_{o,d}$, where φ describes the extent of regime-specific departure of observed association between O and D from the independence model, normalized between 0 and 1. We define intervention strength as $(1 - \varphi)$. Thus, intervention strength varies between 0 (no intervention) and 1 (complete independence). When intervention strength is 1, all post-intervention log odds ratios are zero.

CHAPTER 2. LABOR MARKET TRANSFORMATION AND EFFECTIVELY MAINTAINED IMMOBILITY IN THE UNITED STATES

Abstract

Intergenerational income mobility in the US has remained stable in recent decades. Seeing education as the primary mechanism of intergenerational persistence, earlier literature in social stratification tried to explain the stability by exploring the various forms of continuing SES-based inequality in the education arena. Since education is their main focus, these earlier theories overlook how changes in labor market conditions help maintain immobility. The US labor market has experienced a transformation in skill requirements in the recent decades, with cognitive skill becoming less and non-cognitive skill becoming more important, which suggests a corresponding switch of the channel of income transmission. To evaluate the changing mechanisms of intergenerational income persistence, I developed a novel sequential decomposition that extends the canonical Kitagawa-Blinder-Oaxaca decomposition. Using data from various sources, including the National Longitudinal Survey of Youth 79 and 97, multiple waves of the US census, the American Community Surveys, the Occupational Information Network (O*NET), and information from newspaper job ads, I find that the constancy in the income mobility rate is the result of two opposing forces: intergenerational persistence through the cognitive channel has decreased, while that through the non-cognitive channel has increased. I report these trends in the intergenerational transfer of both skills and occupational skill content.

INTRODUCTION

Intergenerational transmission of status, or intergenerational mobility, is one of the most active and fruitful research areas in sociology. Over the past couple of decades, scholars have accumulated ample evidence about social mobility along a number of different dimensions and across many different temporal and national contexts. The canonical theoretical lens through which these empirical studies are evaluated is the modernization theory, also known as the industrialization, functionalism, or liberal theory (Becker, 2010; Bell, 1976; Blau and Duncan, 1967; Treiman, 1970; Hout, 1988). This theory, which adopts children's education as the central mediating mechanism of intergenerational status transmission, envisions an increasingly mobile society as the equalization of education opportunities constrains the total impact of family origin on children's outcomes. Such an optimistic belief in the "meritocratization" process echoes the classical economics theory that poses human capital (ability, or skill) as the fundamental determinant of labor market success (Goldin and Katz, 2010).

However, against the predictions of the canonical perspective, income mobility has remained almost constant in the United States in the recent decades (Chetty et al., 2014; Bloome et al., 2018). This is a noteworthy and somewhat surprising phenomenon that has prompted numerous explanations, some of which continue to be hotly debated. The prevailing explanations for this phenomenon in sociology are still largely based on the canonical perspective, which is represented by the "OED" (origin-education- destination) triangle. One line of thought suggests that family-based educational in- equality, represented by the "OE" association, has not declined despite the expansion of educational opportunities. For instance, children from advantaged backgrounds are more likely to benefit from the expanded opportunities and are more likely to secure positions in higher-quality institutions within the same educational level (Raftery and

Hout, 1993; Lucas, 2001). This line of scholarly inquiry offers a more nuanced understanding of the intricate ways in which parental factors shape their children's educational attainment, but they still regard education as the sole mechanism in the intergenerational transmission of status.

While education remains perhaps the most significant predictor of individual achievement, the one-dimensional approach to status transmission may have lost its analytical efficacy as both class advantage and labor market requirements become increasingly multidimensional. Family background can influence children's outcomes through channels that might have overlap with but cannot be fully accounted for by education. For example, the Wisconsin model of status attainment considers the influence of psycho- logical variables like aspiration, expectation, and intelligence on personal achievement (Sewell et al., 1969, 1970; Woelfel and Haller, 1971; Haller and Portes, 1973; Hauser et al., 1983). More recently, a series of empirical studies pointed out the cultural mechanisms of elite reproduction through the transmission of cultural capital, leadership and social skills, and habitus suited for high-status jobs (Jackson et al., 2005; Jackson, 2009; Rivera, 2012; Armstrong and Hamilton, 2013; Rivera, 2016; Laurison and Fried- man, 2016; Lareau, 2015; Friedman, 2016). In economics, there is also a burgeoning literature on the economic returns to non-cognitive skills, challenging and extending the traditional perspective which views cognitive skills and educational credentials as the core elements of human capital (Heckman et al. (2006) as a pioneer example).

Earlier scholars who focused primarily on education as the central mechanism of intergenerational transmission may have overlooked an alternative explanation for the stability of immobility rates. It is possible that immobility is maintained not just in the educational realm, but rather emerges as a combined result of various interconnected changes occurring in the labor market. While income mobility stays stable over time, labor market conditions in the US have

experienced a transformation, known as “the great reversal of skill demand” (Beaudry et al., 2016). Starting in the 2000s, the return to cognitive skills has gradually declined, whereas the return to non-cognitive skills has increased (Castex and Kogan Dechter, 2014; Deming, 2017; Edin et al., 2022).

In light of this labor market transformation, I propose an alternative explanation for the persistence of immobility that focuses on the role of skills and occupational content. To account for the multidimensional nature of persistence, I split the labor market paths into a “cognitive” path and a ”non-cognitive” path. Within each path, I distinguish two types of inequality maintenance: “primary maintenance,” that is, through parental influence on children’s skills; and “secondary maintenance,” that is, through direct parental influence on children’s occupational content, independent of their skills. Given the “great reversal of skill demand”, I hypothesize that immobility is “effectively maintained” in the changing labor market: whereas family transmission happens less through the cognitive path, the decline is counterbalanced by an increase in transmission through the non-cognitive path.

To evaluate this perspective, I developed a novel sequential decomposition of intergenerational income transmission that provides better-defined estimands and more interpretable results compared with earlier ad-hoc approaches. Treated as a purely descriptive summary, it can be understood as an extension of the Kitagawa-Blinder- Oaxaca decomposition, one that is non-parametric, sequential, and custom-tailored for a scenario involving a continuous group indicator. In the Appendix, I discuss the causal interpretation of the decomposition, which makes the decomposition a novel contribution to a nascent literature within causal inference (VanderWeele and Robinson, 2014; Jackson and VanderWeele, 2018; Jackson, 2018, 2021; Lundberg, 2021; Wei and Xie, 2022; Zhou and Pan, 2023). This approach can be viewed as a

generalization of the recently proposed “gap-closing estimand” to a context with continuous and sequential mediators (Lundberg, 2021).

Using data from a variety of sources, including the National Longitudinal Survey of Youth 79 and 97 (NLSY79 and NLSY97), multiple waves of the US census, the Ameri- can Community Surveys (ACS), the Occupational Information Network (O*NET), and information from newspaper job ads, I find that constancy in the income mobility rate in the US is the product of two countervailing forces: whereas intergenerational persistence through cognitive channels is declining, that through non-cognitive channels is rising. The result is a product of the shifting importance of skills in the labor market and the shifting skill advantage of children from higher-income families. The trends are observed in parental transfer of skills (primary maintenance), but not in direct parental influence on occupational content (secondary maintenance). A closer look reveals that the trends in transmission mechanism are driven mostly by the changing skill portfolio of children from the richest families in the top 10 percentile rank. For the older 1979 cohort of the NLSY, there is a significant return to the top level of cognitive but not non-cognitive skill, which favors children from top-income families because they have better cognitive but not non-cognitive skill. However, for the more recent 1997 cohort of the NLSY, the situation is reversed. There is a return to the top level of non-cognitive skill, which the richest children are better equipped with, but no clear return to top cognitive skill, in which children from wealthy families have no apparent advantage.

This article contributes to multiple related strands of research. First, it expands the well-documented impact of labor market changes on economic inequality, which has mainly been studied in a cross-sectional context, by investigating their implications for intergenerational inequality. Second, though it is well known in the area of child development that differential

parental practices between socioeconomic groups reinforce achievement gaps among children, the long-term labor market outcomes associated with these gaps and their connection to macro-level changes is less understood, which this research helps illuminate. Third, this research documents a trend in intergenerational mobility that goes beyond the traditional focus on transmission through education by examining the roles of skill formation and job market dynamics. Lastly, it provides a methodological alternative to studying channels of intergenerational transmission, a once-illustrious research agenda exemplified by the status-attainment model that has lost popularity partly due to methodological limitations.

The article is organized as follows. Section 2.1 reviews the canonical theory of intergenerational mobility that centers on the role of education. Section 2.2 discusses the recent empirical observation of the changing economic return to skills and occupational skill content, which poses challenge to the education-centric view of intergenerational transmission. Section 2.3 and Section 2.4 review the literature around the evolving relationship between parental income, children's skills (2.3), and children's occupational content (2.4). Section 3 discusses the main research hypothesis (3.1), defines the estimands (3.2), and introduces the estimation procedure (3.3). Section 4 introduces the data, Section 5 presents results, and Section 6 concludes.

MOTIVATION

The Puzzling Persistence of Intergenerational Immobility and the Canonical Account

Intergenerational mobility, measured as the association between the social statuses of parents and children, is an important indicator of the openness of a society and has long been an

important subject of sociological research. A central theoretical perspective behind the voluminous literature studying intergenerational mobility is what I refer to as the canonical theory, which mainly originated within sociology but was also influenced by labor economics (Becker, 2010; Bell, 1976; Blau and Duncan, 1967; Treiman, 1970; Hout, 1988). The theory holds, though sometimes only implicitly, that one's achieved social status results from one's human capital, and human capital is best represented by educational qualifications. High-status parents are more likely to have high-status children because their children have better educations. Since education is the core mediating mechanism of intergenerational transmission, the canonical model predicts an increase in mobility in the long-term because education opportunities have generally become more available to children from all social origins. The influence of family origin on children's achievement is believed to be limited, beyond the mechanism of children's education. The framework is often summarized as the "OED (origin-education-destination) triangle". Though the main parts of the theory are developed in the context of class mobility, the perspective has also been employed in the analyses of other mobility indicators, including income mobility, the focus of this research.

The picture drawn by the canonical theory of social mobility echoes the widely held narrative of "the American Dream" or "the land of opportunity" in the United States, which pictures the US as a place with equality of opportunity, where one's chances of success are determined only by one's own effort but not one's family origin. Empirical evidence, however, in general rejects such a picture. Occupation-based mobility has been shown to be stable or even declining over time if the horizon is expanded to the 19th century (Erikson and Goldthorpe,

1992; Song et al., 2020). In terms of income mobility, based on large-scale administrative data and survey data, scholars have also shown a trend of stability over time (Chetty et al., 2014)[8].

The persistence of intergenerational immobility invites explanations that diverge from the canonical theory. The first and principal observation that goes against the prediction of the canonical theory is that education is not becoming more equalized: the expansion of education opportunities benefits children from advantaged backgrounds disproportionately (Raftery and Hout, 1993).A more recent line of empirical research shows that the education gap by family wealth and income has widened in recent years (Reardon, 2011; Pfeffer, 2018; Chmielewski, 2019), even as the education gap by family class appears to be stable or weakening over time (Shavit and Blossfeld, 1993; Breen and Mu¨ller, 2020). In the mean time, consistent with the prediction of the canonical theory, economic returns to education have increased in the recent decades (Goldin and Katz, 2010), so the persistent immobility may be a joint product of both persisting OE and ED associations.

Another important and related line of literature acknowledges the expansion and equalization of nominal education opportunities, but points out that such equalization is insufficient to reduce immobility, because education is a positional good – the value of one's educational credentials depends on the amount of education held by others (Thurow, 1972; Jencks et al., 1972; Boudon, 1974; Raftery and Hout, 1993; Lucas, 2001). EEducation is not just a proxy for human capital, but also a credential whose value hinges on institutional arrangements and social approval. Proponents of the relative education thesis like to use the metaphor wherein status attainment is described as a "queuing" process with a fixed number of high-income or high-status positions. Competitors are ranked and admitted to these positions based on their

[8] Throughout the article we focus on relative mobility. The trend of absolute mobility is different as shown in Chetty et al. (2017).

education credentials. In this process, it is one's rank, not level, of educational attainment that determines class destination. Education expansion only equalizes the nominal level but not the rank of education among children from different class backgrounds, because the quality, though not the level, of the schools that high-SES children attend is still higher than that of children from disadvantaged backgrounds (Lucas, 2001). The relative nature of education attainment has led some scholars to famously contend that expanding education will not reduce class-based inequality (Jencks et al., 1972; Boudon, 1974).

A recent article, Bloome et al. (2018), serves as an exceptional example of efforts geared towards elucidating how the rate of income immobility has been sustained within the educational sector over the past several decades. Utilizing a decomposition technique, Bloome et al. (2018) systematically articulates both the equalizing and unequalizing roles of education. The expansion of education has mitigated income persistence, generating a greater number of high-income adults from low-income back- grounds and also diminishing the direct impact of parental income on adult earnings. However, this equalizing effect is offset by an escalating educational inequality driven by parental income, as well as an increasing returns to education.

The canonical theory and the following explanations discussed above treat education as the main mechanism of status transmission, relegating other channels of parental influence to residual status. This theoretical approach yields a powerful account of how inequality is maintained in the educational realm, but is almost totally agnostic to the labor market aspect of inequality persistence. Parents pass down their advantages by providing their children with the skills and knowledge necessary to excel in the job market and by securing them high-status occupational positions. While the skillsets that lead to better occupations and the occupations themselves are undeniably linked to education, the education-centric model only reflects the

complex dynamics of the labor market implicitly, potentially overlooking crucial empirical trends in the labor market that are of interest to researchers.

The need for an explicit multidimensional labor market account of income trans- mission is emphasized by recent empirical observations which suggest that the persistence of intergenerational immobility may be related to changing labor market demands and parenting styles. The next sections discuss these empirical trends.

The Great Reversal of Skill Demand

"The 20th century is the century of human capital", triumphantly declared in the influential book Race between Education and Technology (Goldin and Katz, 2010). The theory represented by the book, skill-biased technological change (SBTC), states that human capital, as conceptualized as the ability to perform job tasks and as measured by educational qualifications, is the most important indicator for the success of both individuals and nations. The theory shows its strong predictive power by explaining most of wage inequality trends in the 20th century.

SBTC's emphasis on human capital as the primary determinant of success in the labor market is consistent with the canonical theory of mobility, which posits that education is the primary means of achieving intergenerational mobility. However, this perspective overlooks the fact that abilities and human capital are multidimensional, and the skills that are most valued in the labor market can change depending on various social and economic factors. Factors such as technological advancements, globalization, and shifts in industries can alter the occupational structure and favor certain skills over others (Acemoglu and Autor, 2011). Even within occupations, the nature of work may change as a response to technological development, such as automation, which can result in changes in skill demands. Thus, the relationship between human

capital and labor market success is more complex than the "meritocratization" narrative suggests (Deming and Noray, 2020; Atalay et al., 2020).

Recent empirical studies documented a reversal of skill demand in the US labor market (Beaudry et al., 2016). Castex and Kogan Dechter (2014) shows that, though returns to college education have increased in the recent decades in the US, as predicted by the canonical theory and SBTC, the return to specific skills shows different trends. The return to cognitive ability declined dramatically in the early 2000s compared with the 1980s. On the other hand, returns to non-cognitive abilities have been shown to have increased during the same time period (Deming, 2017).

[Figure 1 about here.]

I replicate, update, and extend the finding of "the great reversal of skill demand" as reported by previous studies. First, high-skill jobs are constituting a bigger proportion of the labor market. Figure 1 shows the changing employment shares of occupations with different skill levels since the 1980s. The share of high-skill occupations, those whose skill requirements are in the first quartile, has increased substantially. Occupations with a high cognitive component increased from about 10 percent to the 20 percent, whereas non-cognitively demanding jobs experienced an even bigger increase, that is from about 18 percent to over 30 percent. The third column shows that the two trends are concurrent. The occupations with only a high cognitive or non-cognitive com- ponent are not taking a higher share of the labor market. By contrast, the occupations with both high cognitive and non-cognitive requirements have expanded substantially.

[Figure 2 about here.]

Second, as Figure 2 shows, after 2003, the importance of cognitive skill has been declining whereas the importance of non-cognitive skills keeps increasing. Though across-occupation both high-cognitive and high-non-cognitive skill jobs are taking a bigger proportion, the within-occupation trend of cognitive skills and non-cognitive skills diverge as one is declining in importance and the other is increasing.

[Figure 3 about here.]

The labor market reversal is characterized not only by the changing importance of skills but also their changing economic returns. Figure 3 reports the trend in wage returns to skills as measured by the interaction between skill importance and time with wage as the outcome. After 2003, the return to cognitive skills is decreasing whereas the return to non-cognitive skills is increasing. Figure 3 conveys a similar message to Figure 2: the returns to and prevalence of non-cognitive skills are increasing, while those of cognitive skills are decreasing.

Various explanations for the decline of the return to cognitive ability have been proposed, with many focusing on the content and pace of technological advancements. One straightforward explanation is that the pace of the development of the IT industry has slowed down significantly, which was a main source of the increase in cognitive skill demand in the 1980s (Bound and Johnson, 1989; Gordon, 2012). Since there is not as many new job positions in high-tech industries, the demand for cognitive skills has shrunk. Furthermore, the reduced rate of

technological advancement lowers worker's need to constantly update their skills, therefore lowering the demand for cognitive ability which is consequential in commanding new tools (Castex and Kogan Dechter, 2014). Finally, though technological advancement like automation is traditionally believed to have the biggest impact on low-skilled workers who perform routine tasks, scholars have begun to realize that high-skilled workers with cognitively intensive tasks are also at the risk of being replaced by computers and artificial intelligence (Brynjolfsson and McAfee, 2014; Susskind and Susskind, 2015; Grace et al., 2018). The recent rise in machine learning enables computers for the first time to accomplish complicated tasks that used to be completed only by human beings with professional training, like medical diagnosis, tax processing, or driving. The replacement of high-skill jobs by AI is one possible explanation for the decline in the economic return to high cognitive ability.

Technological development has curbed the need for cognitive skills but spurred the demand for other skill sets (Autor, 2015; Acemoglu and Restrepo, 2019). As machines replace labor in a lot of tasks, new tasks where people still have a comparative advantage to are created. Human interaction is one task that machines still have difficulty performing. Deming (2017) found that, comparing the 2000s and 1980s, there are more jobs that require high social skills, and economic returns to social skills have increased. By contrast, jobs that require math skills but low social skills have shrunk in the same period. This phenomenon of technologically increase in social skill demand is best illustrated in the case study of bank tellers in Bessen (2015). After the invention of ATM, the need to interact with a human bank teller in order to withdraw money be-came almost non-existent. A natural prediction was that the bank teller as a job would become extinct. However, Bessen (2015) found just the opposite, that the employment of bank tellers increased modestly after the invention of the ATM. The reason is that the tasks of bank tellers

completely transformed after the invention of the ATM, from a technical job to a socially intensive job whose main task is to respond to customer questions and promote financial products.

Class-Based Inequality in Non-Cognitive Skills

As non-cognitive skills have emerged as a crucial predictor of success in the labor market, researchers have increasingly recognized the role of parents in facilitating their children's development of these skills. In sociology, there is a long tradition of studying non-cognitive determinants of achievement and the non-cognitive dimension of intergenerational status persistence. The Wisconsin model of social stratification poses psychological variables such as aspiration and expectation as important components of the process of status attainment (Sewell et al., 1969, 1970; Woelfel and Haller, 1971; Haller and Portes, 1973; Hauser et al., 1983). More recently, in economics, there is a burgeoning literature that studies the development of and returns to personality traits (Heckman et al., 2006; Cunha and Heckman, 2007; Heckman and Mosso, 2014). Inspired by the Wisconsin model and the "non-cognitive" turn of labor economics, a recent line of empirical studies in sociology assesses the importance of psychological measures in many life outcomes ranging from school performance to labor market success (Mood et al., 2012; Blanden, Gregg, and Macmillan, Blanden et al.; Blanden et al., 2014; Doren and Grodsky, 2016; Hsin and Xie, 2017; Bourne et al., 2018; Esping-Andersen and Cimentada, 2018; Liu, 2020; Holtmann et al., 2021; Huang et al., 2022).

Another literature provides a more detailed examination of the role of parenting in children's non-cognitive development and how that contributes to their success in the labor market. Lareau (2011) describes a style of parenting, "concerted cultivation", that is typically

associated with middle-class families in the United States, and it involves parents actively and intentionally cultivating their children's talents, interests, and skills by organizing and supervising structured activities such as sports teams, music lessons, and tutoring sessions. The goal of concerted cultivation is to develop children's social and cultural capital, which can be used to gain advantages in school and in later life. This parenting style emphasizes the development of communication skills, negotiation abilities, and comfort with authority figures, among other things.

As class-based differences in parenting culture gains scholarly attention, a growing line of literature has pointed out the important roles of cultural and personality characteristics in labor market stratification processes and intergenerational inequality maintenance (Jackson et al., 2005; Jackson, 2009; Rivera, 2012; Armstrong and Hamilton, 2013; Rivera, 2016; Laurison and Friedman, 2016; Lareau, 2015; Friedman, 2016). Employers in elite industries often prioritize cultural compatibility when selecting candidates for their organization. This is because employees who share similar interests, hobbies, and values are better able to form strong social bonds with their colleagues, leading to a greater sense of solidarity and a more positive work environment. Further- more, culturally savvy employees can help to signal the company's status to clients and customers, which is important in maintaining the company's reputation and attracting new business.

Although to my knowledge, there is no definitive evidence on whether the class gap in non-cognitive skills is widening in the United States or if such a shift, should it exist, is in response to changes in labor market demand, recent studies indicate a growing class divide in the non-cognitive development of children. Indicative evidence comes from changes in parenting styles over time. The engagement of parents in their children's development, especially through

extracurricular activities, is a good indicator of non-cognitive development (Covay and Carbonaro, 2010), and it appears that parents with higher socioeconomic status (SES) are spending more time with their children in leisure and cultural activities and becoming more involved in their education (Bassok et al., 2016; Kalil et al., 2016). Ryan et al. (2020) reported changes in parenting style over time, which they linked to a changing labor market context. The class gap in mother's engagement in their children's education has widened, with the difference in time spent with children for leisure and studying activities growing significantly between college-educated and non-college-educated mothers since the early 1990s (Ramey and Ramey, 2009).

Intergenerational Income Persistence through Occupational Content

Intergenerational inequality can persist not only through the transfer of skills or abilities but also through the more direct transfer of occupational content. Sociologists have long been interested in the occupational aspect of intergenerational status persistence and have examined extensively the association between parental occupation and children's occupation. The association suggests that children often end up in occupations similar to their parents, regardless of how "similarity" is defined - be it at the macro-level (Erikson and Goldthorpe, 1992; Xie, 1992), micro-level (Grusky and Sorensen, 1998; Grusky and Weeden, 2001; Weeden and Grusky, 2005), or by continuous characteristics (Blau and Duncan, 1967; Sewell et al., 1969, 1970; Hout, 1984).

Another more recent strand of literature tackles a related but distinct question: how children's occupation serves as a mediator for the transfer of intergenerational economic advantage (Breen et al., 2016; Mood, 2017). They find that, high-income parents have high-

income children because their children have “better” occupations, as defined by higher average income.

The extant findings on the mediating role of occupation for income persistence and the correlation between occupational and income mobility provide valuable insights. However, the current discussion is focused on the mediating role of “class”, or clusters of occupation categories. As a result, it is also subject to the well-known limitations associated with the class approach to inequality, particularly the macro-class approach. These limitations have been highlighted by advocates of “the death of class” (Clark et al., 1993; Pakulski and Waters, 1996) and ”the decomposition of class” perspectives over the years (Grusky and Sorensen, 1998; Grusky and Weeden, 2001; Weeden and Grusky, 2005).

In an increasingly “upgraded” class structure, the occupation of children from advantaged economic backgrounds have been more and more homogeneous in the sense of macro-class (Pakulski and Waters, 1996). Meanwhile, there has been a significant growth in within-class differentiation along a series of occupation characteristics (Jonsson et al., 2009). If these trends are overlooked, the macro-class approach may underestimate or misreport the extent of income persistence through occupational achievement. To address these issues, researchers can also choose to analyze the mediating role of occupation with a more detailed occupational scheme. The micro-class school of occupational mobility emphasizes “local” mechanisms of social reproduction and highlights that many channels of status transmission occur specifically at the occupation level, which cannot be simply aggregated to classes. These mechanisms include occupational association, cultural habits, unions, and more (Grusky and Sorensen, 1998; Grusky and Weeden, 2001; Weeden and Grusky, 2005).

The categorization of occupations on macro, meso, or micro scales confronts a common limitation: the compartmentalization of occupations easily disregard the underlying similarities and existing interconnections between various occupations across a range of dimensions. This gap is especially salient in the analysis of income mobility mediated by occupation, more so than in the study of occupational mobility.

In the context of occupational mobility, mechanisms such as familial social net- works and micro-level interactions could predispose children to landing in the same professions as their parents. In these instances, parents' occupational influence is "incommensurable," as its impact specifically unfolds when children venture into the same professional fields, reducing its applicability when children enter a different occupational context from their parents.

Regarding income persistence mediated by occupation, parental income assumes a "commensurable" role. This implies that parental resources, in this case financial, can be flexibly employed in diverse scenarios associated with guiding their children into varying occupations. Parents might not necessarily wish to steer their children exclusively towards certain professions. A more plausible assumption draws upon the tendency for parents to encourage their offspring towards career paths characterized by attractive features - these may include prospects for high income, advanced skill requirements, or carrying a high-status (Irwin and Elley, 2013). Essentially, it is the characteristics of the profession, rather than the profession itself, that often become the focal point of parental guidance and are heavily affected by parental economic resources.

Therefore, this research attempts to expand earlier studies that investigate income persistence through children occupation by changing the measure of occupation from categorical to continuous. In the tradition of occupational mobility this is referred to as the "gradationalism",

contrasting the “cateogicalism”, approach (Weeden and Grusky, 2012). In indexing occupational features, common dimensions of choices include prestige (Duncan, 1961; Hauser and Warren, 1997), status (Chan and Goldthorpe, 2007), income (Nam and Powers, 1983), and desirability (Jencks et al., 1988).

I focus specifically on another continuous measure of occupational feature, that is occupational content, as defined by the activities, knowledge, and skills employed in the occupation’s daily work. The pivotal role of occupational content has been extensively demonstrated in several directions of economic and sociological studies. It serves as a significant determinant of economic return (Heckman et al., 2006; Deming, 2017; Acemoglu and Autor, 2011), a predictor of worker’s career trajectory (Deming and Noray, 2020), and also an anchor of the position of the occupation within a network (Lin and Hung, 2022). It is a strong shaping force of cross-sectional inequality (Cobb- Clark and Tan, 2011), and also a strong predictor of the flow of occupational mobility among workers (Alabdulkareem et al., 2018; Cheng and Park, 2020).

The necessity for focusing on occupational content to comprehend the role of occupations in income mobility is reinforced by the increasingly variable condition of the labor market, as discussed in an earlier section. Given the dynamically changing re- turns associated with distinct occupations and varying skill sets, parents’ perceptions of the ideal occupations for their children might also be experiencing concurrent shifts. These shifts in the focus of intergenerational status transmission can only be detected using the gradational, but not class-based, analytical approach.

THEORETICAL AND ANALYTICAL FRAMEWORK

Hypothesis: Effectively Maintained Immobility in a Dynamic Labor Market

The earlier sections reviewed the changing requirements for and returns to skills in the labor market on the one hand, and parental transfer of status via children's skill formation and children's occupational skill content on the other. A natural question thus emerges, that is, whether the persistence in income immobility over the recent decades can be attributed to the intergenerational transfer of skills and occupational skill contents, and whether and how the transformation of the labor market has facilitated the process of intergenerational inequality preservation. This section outlines the main research hypothesis in detail.

Given that shifts in the labor market can be characterized by changes in the relative value of cognitive and non-cognitive skills—with cognitive skills diminishing and non-cognitive skills increasing in importance-I split the process of income transmission within the labor market into two main conduits: the cognitive pathway and the non-cognitive pathway. For each of the cognitive and non-cognitive aspects of inter- generational persistence, I classify the transmission process into a "primary" and a "secondary" component, using concepts borrowed from (Boudon, 1974). The primary maintenance of intergenerational inequality occurs via the following channel: children with high-income parents have higher cognitive or non-cognitive skills due to factors such as parental financial investment, parenting style, cultural environment, and contextual factors like neighborhood differences. These skill advantages are then signaled to employers through educational credentials, interview performance, or other aspects of their job application. As a result, children with high cognitive or non-cognitive skills are more likely to be placed in occupations that require these skills. Since jobs with high skill requirements offer high economic returns, children who are employed in these occupations have higher incomes. The transmission process depicted here is in many ways in the same spirit as the classic Blau-Duncan model of status attainment (Blau and Duncan, 1967: p 170) where parental status affects

children's current occupational status through children's education and then the status of children's first occupation. The main difference in our approach is that we take both skills and occupational contents as multidimensional.

In addition to the primary channels of transmission, there is also a "secondary" maintenance process, whereby parents directly influence their children's occupational positions independent of their skills. Parents can have a significant impact on their children's career development and decision-making. High-SES parents may have a greater awareness of the status and economic hierarchy of the current occupational structure and may encourage their children to pursue occupations with better prospects in the current labor market (Middleton and Loughead, 1993; Nucci et al., 1996). In contrast, disadvantaged parents may be less equipped to help their children explore good occupational opportunities due to their informational disadvantage. High-income parents, who themselves are likely to be in high-status and skill-intensive occupations, may use personal and social networks developed during their own careers to help their children secure similar high-end jobs Jonsson et al. (2009). Finally, high-income parents may provide direct financial assistance, such as bridge funds, to support their children during their early career and job search process, thereby reducing their financial and psychological burden by covering their living expenses and providing them with more time to consider and better resources to secure desired occupational positions (Mitnik et al., 2016; Bukodi and Goldthorpe, 2022).

[Figure 4 about here.]

Figure 4 provides an illustration of the primary and secondary channels of maintenance given a certain skill, whether it is cognitive or non-cognitive. Each of these channels corresponds to a source of inequality. Primary maintenance occurs because of the inequality of skill by parental income. This layer of inequality can be further separated into two sources. First, parents with higher incomes often have better resources in other regards. For instance, they may have a better education, live in better neighborhoods, or themselves may have inherited more wealth, all of which may contribute to the economic success of their children. I label this channel as primary maintenance through environment, as it does not reflect the real effect of parental income but the "confounding" force of other factors. The other channel is more direct transmission through economic means, that is the inequality of skill by parental income even after controlling for other family background factors. Secondary maintenance arises due to the unequal skill intensity of occupations by parental background after controlling for children's skill level. Finally, there are channels of intergenerational transmission that are irrelevant to the labor market path and are not captured by the measures of skill and occupation.

The multidimensional nature of intergenerational income transmission, encompassing cognitive and non-cognitive skills, as well as primary and secondary channels, predicts the effective maintenance of immobility in a changing labor market environment, which constitutes the core hypothesis of the research. More specifically, intergenerational transmission shifts its mechanism as labor market demand changes, and the decline of transmission through cognitive channels is offset by the rise of transmission through non-cognitive channels, creating the stability in intergenerational mobility rate.

There are two theoretical rationales that lead to the "offsetting" hypothesis. First, children from advantaged backgrounds are typically more skilled in a wide range of areas that are

considered valuable in the labor market. Not only are they better at cognitive tasks such as reading, writing, and math (Heckman and Carneiro, 2003; Cunha and Heckman, 2007), but they also excel in communication, socialization, and emotional stability, among many other non-cognitive abilities (Heckman and Mosso, 2014; Fletcher and Wolfe, 2016; Calarco, 2011, 2014). This overall advantage in skills among children from high-SES backgrounds means that they will generally have better performance in the labor market, regardless of which skill sets are in demand at any given time. As the market demand for skills shifts, a decreasing pathway of intergenerational transmission is automatically or mechanically compensated by an increasing pathway.

Second, parents may adjust their parenting strategies and investments to focus on skills that are more valuable in the current job market. Parents from privileged back- grounds are typically better informed, better equipped, and have stronger incentives to make such adjustments. They may have more access to labor market information through their individual social networks or through professional or industry associations (Lin, 1999; Perna, 2006). With more resources at their disposal, they also have greater flexibility in terms of parenting priorities, such as selecting extracurricular activities, social activities, or choosing a college and major. They also have the strongest incentive to maintain their children's status according to the rational-action theory of social mobility (Boudon, 1974; Breen and Goldthorpe, 1997). In order to avoid relative status loss and downward mobility, wealthy parents need to continually update their strategies for child development to ensure that their children remain competitive (Lipset and Bendix, 1991; Kahneman and Tversky, 1979).

Theoretical perspectives that evaluate how intergenerational inequality is sustained in dynamic environments echo some of the well-established insights in the literature of social

stratification. As old channels of inequality close, new ones may emerge. “Effectively maintained inequality”, a perspective developed mostly as a response to the rise of horizontal stratification in the education system (Lucas, 2001), is one such example. The proponents of EMI noted how high-SES parents make flexible use of their resources to ensure an advantaged position of their children in the education system. Similarly, in the area of health inequality, the “fundamental cause theory” (FCT) has highlighted the idea of “pervasiveness” by pointing out that high-SES individuals always benefit first from advancements in health technology because of their cultural, economic, and social resources (Link and Phelan, 1995). Our theoretical perspective builds on these key insights to investigate how intergenerational inequality is sustained in a dynamic labor market environment.

Quantifying the Channels of Maintenance: Defining the Estimands

Testing the core hypothesis of the research requires quantification of these channels of transmission. Epitomized by the “status attainment model”, path analysis, or structural equation models, has been the most frequently used tool to study the mechanism of intergenerational transmission in sociology (Duncan, 1966; Sewell et al., 1969, 1970; Alwin and Hauser, 1975; Woelfel and Haller, 1971; Haller and Portes, 1973; Hauser et al., 1983). Some more ad-hoc analyses simply use intergenerational income regressions with control variables to evaluate the mechanism through those control variables. The explanandum in regression analyses, however, is individual children income, but not income immobility. Whereas a regression may tell us to what extent parental in- come, skill, and occupation explains or causes individual income, it does not tell us how skill and occupational inequality by parental income contributes to intergenerational income immobility (see Lundberg (2021) for a careful discussion of the difference between

regression analysis and disparity analysis). Furthermore, when relying on estimates derived from linear regressions, the findings are affected by inaccuracies stemming from parametric assumptions and the absence of a clearly defined estimand (Lundberg et al., 2021). Regression analysis is therefore ill-equipped to study the mechanism of intergenerational income transmission.

Decomposition analysis, like Kitagawa-Blinder-Oaxaca (KBO) decomposition, on the other hand, is tailored to study the determinants of between-group inequality. However, the group indicator in KBO decomposition is binary, whereas in this case we try to explain the inequality in income by a continuous group indicator, that is parental income. Moreover, KBO decomposition does not take the sequential order of explanatory factors into account, whereas in our setup we wish to test whether occupational content, after accounting for personal skill, still contributes to income immobility.

I develop a non-parametric sequential decomposition that estimates the contribution of skill inequality (primary maintenance) and occupational skill content inequality net of skill (secondary maintenance) to income immobility. Let A denotes parental income, M denotes skill, S denotes the occupational skill level, Y as income, and X a series of family environment associated with income, skill, and family income. First, we consider parental income A as a binary variable, $A = 1$ stands for rich and $A = 0$ stands for poor, for illustration purposes.

Simply put, I simulate interventions that sequentially remove the inequality $P(m|x, A = 1) \neq P(m|x, A = 0)$, $P(m|A = 1) \neq P(m|A = 0)$, and $P(s|m, A = 1) \neq P(s|m, A = 0)$, and estimate post-intervention income mobility. These interventions take specific parental income groups as reference groups and equalize so that all children's distribution of skills/occupational skill content follows the distribution within this parental

income group. The incremental immobility reduction achieved by these interventions reflects the magnitudes of the channels of maintenance, which are primary maintenance through economic means, primary maintenance through environment, and secondary maintenance respectively. The method can be viewed as an extension of KBO-decomposition to a non-parametric, sequential, and continuous setting. Though I interpret the results as purely descriptive summaries, they can also be interpreted as causal effects under certain assumptions. I derive identification formulas and clarify assumptions that warrant the results causal interpretations in Appendix B.1.

First, we define the estimands, the descriptive counterfactuals with which we quantify the strength of transmission mechanisms. Taking children with rich parents as a reference group, to evaluate primary maintenance through economic means, we define a counterfactual children income where the skill levels of children with rich parents are the same as the skill levels of children with poor parents, conditional on X. I denote this counterfactual as $\theta(1,0)$. Here, the first number in the parentheses means that the reference group is $A = 1$, and the second number means the skill of the reference group is as if from $A = 0$.

$$\theta(1,0) = \int\int E[Y|m,x,A=1]P(m|x,A=0)P(x|A=1)dmdx(1)$$

Then, to evaluate the full effect of primary maintenance, either through X or not, we define a descriptive counterfactual where the skill levels of children with poor parents are the same as the skill levels of children with rich parents, this time not conditional on X. I denote this counterfactual as $\theta(1,0^*)$, where the symbol $*$ denotes a marginal instead of conditional equalization.

$$\theta(1,0^*) = \int\int E[Y|m,x,A=1]P(m|A=0)P(x|A=1)dmdx(2)$$

Note that $\theta(a,a) = E[Y|A=a]$. Compare these estimands with the observed expected outcome of children with rich parents:

$$\theta(1,1) = \int\int E[Y|m,x,A=1]P(m|x,A=1)P(x|A=1)dmdx(3)$$

It is clear that compared with the observed conditional expectation, $\theta(1,0)$ reweighs the conditional expectation so that variable m is as if from the counterfactual conditional distribution where $A = 0$, $P(m|x,A=0)$. On the other hand, $\theta(1,0^*)$ is a reweighed version of the observed expectation where variable m is as if from the coun- terfactual marginal distribution where $A = 0$, $P(m|A=0)$.

The above two counterfactual estimands help evaluate the transmission through M. To evaluate the further transmission through S, we define a counterfactual where, after skill disparity is already removed, the occupational contents of children with poor parents are the same as those of children with rich parents. This counterfactual is defined as $\theta(1,0^*;0^*)$[9].

$$\theta(1,0^*;0^*) = \int\int\int E[Y|m,s,x,A=1]P(s|m,A=0)P(m|A=0)P(x|A=1)dsdmdx \quad (4)$$

Provided with these moving parts, we can decompose the observed income gap between children from rich families and children from poor families into transmission channels that can

[9] In a similar vein, $\theta(1,0^*;0) = E[Y|m,s,x,A=1]P(s|m,x,A=0)p(m|A=0)p(x|A=1)dsdmdx$. Also notice that $\theta(1,0^*) = \theta(1,0^*;0)$.

be attributed to skills and occupational skill contents:

$$
\begin{aligned}
E[Y|A=1] - E[Y|A=0] &= \theta(1,1) - \theta(0,0) \\
&= \underbrace{\theta(1,1) - \theta(1,0)}_{\text{Primary (Economic)}} + \underbrace{\theta(1,0) - \theta(1,0^*)}_{\text{Primary (Environment)}} \\
&+ \underbrace{\theta(1,0^*) - \theta(1,0^*;0^*)}_{\text{Secondary}} + \underbrace{\theta(1,0^*;0^*) - \theta(0,0)}_{\text{Unaccounted}}
\end{aligned} \tag{5}
$$

The economic component of primary maintenance is characterized by the contrast between the observed income of children from rich families and their income that would be achieved when their conditional advantage in skill is removed. The environment component of primary maintenance is quantified by the additive income disparity reduced when furthering intervening so that the skill between children from different backgrounds is completely equalized. Based on that, secondary maintenance reflects the income loss that would be achieved if the occupational skill level for children from rich families is the same as children from poor families. Finally, there may still be a class-gap in children income even after the total equalization of both children skill and occupational skill level. The decomposition takes a similar structure to the Kitagawa- Blinder-Oaxaca descriptive decomposition and the causal mediation decomposition. The relationship between descriptive, interventional, and mediational decomposition is discussed in detail in Appendix B.2.

Parental income, A, however, is only assumed to be binary for illustrative ease. In empirical analysis there is no good theoretical rationale to conceptualize parental income as a binary variable. We therefore have to transfer the decomposition to the case of a continuous A. Throughout our analyses we measure A as percentile rank from 0 to 1. We use those $A = a$ as a reference group or a target of intervention, and substitute 1 to a in the above decomposition and

iterate over the value of a sufficiently many times to get a series of decomposition results. For each value of a, we have a decomposition in the form similar to the one outlined above:

$$\begin{aligned} E[Y|A=a] - E[Y|A=0] &= \theta(a,a) - \theta(0,0) \\ &= \underbrace{\theta(a,a) - \theta(a,0)}_{\text{Primary (Economic)}} + \underbrace{\theta(a,0) - \theta(a,0^*)}_{\text{Primary (Environment)}} \\ &+ \underbrace{\theta(a,0^*) - \theta(a,0^*;0^*)}_{\text{Secondary}} + \underbrace{\theta(a,0^*;0^*) - \theta(0,0)}_{\text{Unaccounted}} \end{aligned} \tag{6}$$

Suppose, for instance, that $A = 0.5$, then the decomposition tries to decompose the income difference between children with median-income parents and children with bottom-income parents.

Since now the decomposition is a function of a, how should we summarise over the many different values of a? First, consider the total disparity $\theta(a,a) - \theta(0,0)$. We normally summarise the relation between the disparity and a with the coefficient from a linear regression, $\theta(a,a) - \theta(0,0) = \beta 0 + \beta a + \epsilon$. This is equivalent to the rank- rank regression, though the intercept is different, and β is the rank-rank slope. For all the components in the decomposition, we can also summarise that component over a with the coefficient from such a regression. For instance, I run a linear regression $\theta(a,a) - \theta(a,0) = \beta_0^{pri(econ)} + \beta^{pri(econ)} a + \epsilon$. $\beta^{pri(econ)}$ is a linear approximation of the relation between $\theta(0,a)$ and a, and is a summary of total immobility caused by the conditional inequality in M. Similarly, we can run the same regression for the other three components in the decomposition with a as the independent variable and get the other three βs. We would in the end observe that the total rank-rank slope

can be approximately decomposed into exactly the same components in the decomposition outlined above, that is $\beta \approx \beta^{pri(econ)} + \beta^{pri(envir)} + \beta^{sec} + \beta^{unaccounted}$.

The interpretation of the components is similar to the interpretation of the rank- rank slope, β. The interpretation of β is that as parental income increases one percentile, children income on average increases β percentile. Take $\beta^{pri(econ)}$ for example, the interpretation is that as children's ability increases as if their parents' income were one percentile higher, children's income on average increases by $\beta^{pri(econ)}$ percentile. If $\beta^{pri(econ)} = 0$, children's income is not associated with ability, and ability does not play a role in transmission. On the other hand, if $\beta^{pri(econ)} = \beta$, the intergeneratioanl income persistence can be fully explained by the ability advantage children with rich parents enjoy.

Though I interpret the results in a strictly descriptive way, the findings can be regarded as causal under certain assumptions (see Appendix for details). Treated as a causal decomposition, the method I propose is an extension of "the gap-closing estimand" to a setting of continuous and sequential mediators (Lundberg, 2021).

Estimation of the Decomposition

Next I discuss the specific estimation procedure of the statistical estimands, $\theta(a, 0)$, $\theta(a, 0^*)$, and $\theta(a, 0^*; 0^*)$. A weighting estimator, a regression-imputation estimator, and a double machine learning estimator can be established combining the generalized mediation functional with semi-parametric theory of continuous treatments (Kennedy et al., 2017; Zhou, 2020). This article used a "hybrid estimator" of regression-imputation and weighting estimators in sequential implementation (Zhou, 2020).

The first step is to estimate $\theta(a, 0)$. Observe the non-parametric definition of the estimand, it is noteworthy that the integral can be alternatively written as an iterative expectation:

$$\theta(a, 0) = E[E[E[Y|M, X, A = a]|X, A = 0]|A = a] (7)$$

Here, the innermost is a conditional expectation of Y given a fixed value of $A = a$ and two random variables M and X. The middle layer is an expectation with regard to the random variable M. It describes the expected innermost expectation when M is from the distribution given $A = 0$ and random variable X. Finally, the outermost expectation is with regard to X. It captures the expected middle layer expectation when the random variable X follows the conditional distribution given $A = a$.

The formula indicates an iterative regression estimator. Before staring, I sliced A, a continuous variable, into 11 bins by rounding to the first digit to avoid data sparsity, so now A consists values of 0, 0.1 ,0.2 ... 1. First, I estimate a flexible machine learning estimator of the conditional expectation, a random forest, with Y being the outcome and M ,X, and A as independent variables, $E[Y|M, X, A]$. Second, I set everyone's A to be equal a, and get the predicted values for everyone. Third, I run another flexible machine learning estimator – another random forest – with the predicted values obtained in the second step as outcome and X and A as independent variables and get predicted values setting everyone's A to be equal 0. Fourth, I average the predicted values for those $A = a$. This gets us the estimated $\theta(a, 0)$.

As the interventional decomposition is conducted in a sequential manner, the estimation of the next two components relies on the previously estimated ones. Denote the predicted values

for individual i after the third step in the estimation of $\theta(a,0)$ as $\hat{y}_i^*$, which is the imputed individual income after the conditional equalization intervention. In other words, $\hat{y}_i^* = \hat{E}[\hat{E}[Y|M, X = x_i, A = a]|X = x_i, A = 0]$. Also, because $E[\frac{p(M|X,A=0)}{p(M|X,A=a)} Y] = E[E[Y|M, X, A = a]|X, A = 0]$, the predicted values can be alternatively expressed as an estimate for reweighed observed outcome: $\hat{y}_i^* = \hat{E}[\frac{p(M=m_i|X_i,A=0)}{p(M=m_i|X_i,A=a)} Y]$.

Notice that $\theta(a,0^*) = E[\frac{p(M|A=0)}{p(M|X,A=a)} Y|A=a]] = E[\frac{p(M|A=0)}{p(M|X,A=0)} (\frac{p(M|X,A=0)}{p(M|X,A=a)} Y)|A = a]$, which suggests the following weighting estimator:

$$\hat{\theta}(a,0^*) = \frac{1}{n_{A=a}} \sum_{i:A_i=a} \frac{\hat{p}(M=m_i|A=0)}{\hat{p}(M=m_i|X_i,A=0)} \widehat{y_i^*} \qquad (8)$$

Since M is a continuous variable, I slice M into five equal-length groups to facilitate the estimation of the conditional probability. The reweighting requires estimation of two probability models, for which I use a multinomial logistics regression, gaining the predicted probabilities of each individual's own categories, constructing the weights and using them to reweigh $\widehat{y_i^*}$, and average among those $A = a$.

For the estimation of $\theta(a,0^*;0^*)$, the procedure is similar. Notice that $\theta(a,0^*;0^*) = E[\frac{p(S|A=0,M)}{p(S|A=a,M,X)} \frac{p(M|A=0)}{p(M|X,A=0)} (\frac{p(M|X,A=0)}{p(M|X,A=a)} Y)|A = a]$, which suggests the following weighting estimator:

$$\hat{\theta}(a, 0^*; 0^*) = \frac{1}{n_{A=a}} \sum_{i:A_i=a} \frac{\hat{p}(M = m_i|A = 0)}{\hat{p}(M = m_i|X_i, A = 0)} \frac{\hat{p}(S = s_i|A = 0, M_i)}{\hat{p}(S = s_i|A = a, M_i, X_i)} \widehat{y_i^*} \qquad (9)$$

Compared with the estimator of $\theta(a, 0^*)$, a further reweighing factor $\frac{\hat{p}(S = s_i|A = 0, M_i)}{\hat{p}(S = s_i|A = a, M_i, X_i)}$ is added to capture the effect of equalizing occupational skill content. I also slice S into five equal-length groups, run two additional multinomial logistic models, construct corresponding weights, and reweigh the values obtained in the last step again among those whose $A = a$.

DATA

The empirical analysis relies on two main types of data: longitudinal information on occupation-level skill requirements and income, and longitudinal individual-level survey data with income information and ability measures. The first type of data is used to demonstrate the labor market transformation, while the second type of data is used to illustrate the changing mechanism of mobility and to relate it to the transformation of labor market requirements, in conjunction with the first type of data.

Information about labor market skill demands comes from longitudinal O*NET data. O*NET is a detailed taxonomy of skills constructed based on surveys of workers from different occupations about activities and requirements of their jobs. The O*NET worker survey began in 2002 and has been updated progressively since then. The O*NET employed a similar but slightly more detailed occupational code than census. From the year 2002 to 2019, the time span I used, there are a total of 834 occupations surveyed and each occupation has been on average surveyed

3.5 times. Though still limited in some ways, O*NET is widely considered to be the best available data source for studying occupation-level skill requirements (Frank et al., 2019).

For the year before 2000, I used a longitudinal dataset constructed from newspaper job ads that follow the same occupation categorization and skill taxonomy as O*NET (Atalay et al., 2020) (henceforth referred to as APST). The APST data used the text content of about 7.8 million job ads from 1950 to 2000 in three newspapers: the Boston Globe, the New York Times, and the Wall Street Journal. The words used in the ads are then mapped to occupation titles and skill names in O*NET data using Bag of Words models that search extensively for similar words and phrases that appeared in both the ads and the O*NET. Despite best effort, APST and O*NET are not directly comparable. To validate the comparability, I compare APST 1998 and O*NET 1998, the only overlap between the two data sources. I find that the correlation (of rank) between cognitive measures between APST and O*NET is 0.38, and the correlation between two non-cognitive measures is 0.41. I employ a simple calibration procedure, assuming classical measurement error, in which the correlation between APST cognitive/non- cognitive measure and income is multiplied by 0.38/0.41. Results under alternative data choice for occupational content information before 2000 is presented in Appendix C.3.

The O*NET has several modules that ask workers detailed questions about the importance of skills, activities, and knowledge requirements in their occupations. I aggregated this granular information from O*NET to form more comprehensive measures of cognitive and non-cognitive skill contents of occupations. Cognitive skill is measured as the average importance of the following skills: reading comprehension (2A1a), writing (2A1c), mathematics (2A1e, 2C4a), science (2A1f), complex problem solving (2b2i), and creative thinking (4A2b2). Non-cognitive skill is measured as the average importance of the following skills: social perceptiveness (2B1a),

coordination (2B1b), persuasion (2B1c), negotiation (2B1d), service orientation (2B1f), assisting and caring for others (4A4a5), coordinating activities (4A4b1), developing teams (4A4b2), interpreting others (4A4a1), and communication inside and outside organizations (4A4a2, 4A4a3). These occupation-level characteristics are normalized as percentile ranks in the main analysis. Details about the wordings of these requirements can be found in Appendix A.1. A detailed discussion about the comparability and robustness of occupational skill content is in Appendix C.3.

Occupational skill requirement data need to be combined with occupation-specific trends in employment and wages to demonstrate labor market transformation. This part of the data comes from the 1980 5-percent sample, 1990 5-percent sample, and 2000–2019 ACS 1-percent samples, all extracted from the Integrated Public Use Micro-data Series (IPUMS). The occupations are classified under the Standard Occupational Classification (SOC) system, which is the one used by O*NET. When SOC codes are unavailable, I used a fuzzy match from the census code to SOC code to maximize matching between census data and occupation skill data (details in Appendix A.2).

Detailed micro-level data that contains information about personal income, parental income, and specific individual characteristics are used to carry out the study of the mechanisms of intergenerational status transmission. High-quality parental and personal income information are required for calculation of intergenerational income mobility, and detailed personal characteristics need to be controlled for in the causal analysis of intergenerational status transmission as potential confounders.

I use National Longitudinal Survey of Youth data from the 1979 cohort (NLSY79) and 1997 cohort (NLSY97) for this part. The NLSY79 and NLSY97 began with nationally

representative samples of young people aged 14–22 in 1979 and aged 12–18 in 1997 respectively. For parental income, we draw information from the first three years of NLSY79 and NLSY97 respectively. Personal income is measured as the average income between the age of 25 and 34, and those who do not have income for the whole period are excluded. Cognitive ability is measured by the reading and math components of AFQT score derived from Armed Services Vocational Aptitude Battery (ASVAB), a widely used measure of cognitive ability since (Heckman et al., 2006), and is normalized as within-birth-year percentile ranks. NLSY79 administered the ASVAB to respondents in paper form, whereas NLSY97 administered the test on computers, but an earlier evaluation shows that results from the two forms of the test have very high consistency (Moreno and Segall, 1997). A detailed discussion about the comparability of cognitive skill measures can be found in Appendix C.1. The robustness check of alternative income and ability operationalizations can be found in Appendix C.4.

For non-cognitive skill, I use four out of five personality scores in Big-Five personality as measured by Ten Item Personality Measure (TIPI), which are conscientiousness, agreeableness, neuroticism, and extraversion. These personality measures are among the most popular measurements of non-cognitive skills for studies of labor market outcomes (Humphries and Kosse, 2017). The remaining personality of the big- five, openness to new things, is more indicative of creativity and inventiveness and is conceptually closer to cognitive ability, hence not used as a measure of non-cognitive skill. With the four personality scores, I extracted the first component from a factor analysis which is then normalized as within-birth-year percentile rank. This rank is used as a measure of non-cognitive ability for the main analysis. A detailed discussion about the comparability and robustness of the choice of non-cognitive skill measures is in Appendix C.2.

I built a parallel pair of controls between NLSY79 and NLSY97 that include race, sex, number of siblings, education level, type of residence in childhood, cohabitation with parents in childhood, father's education level, father being alive at time of baseline survey, and region of residence. More details about the construction of the variables can be found in Appendix A.1. The descriptive statistics are reported in Appendix A.3.

RESULTS

[Table 1 about here.]

Before moving on to the mechanism of intergenerational transmission, I first present descriptive results about income and skills. Table 1 presents descriptive summaries of the bivariate correlations between family income/income and skills/occupational skill contents. The correlations between cognitive skill/content and both parental in- come and own income are higher compared to the correlations between non-cognitive skill/content and these income measures, especially for NLSY97. This suggests that the cognitive channel is more important than the non-cognitive channel especially for the later cohort. However, the imbalance between the correlations between the two skills and incomes is shrinking when comparing the two cohorts. NLSY97 exhibits smaller correlations between cognitive skill and both parental and own income compared to NLSY79. Conversely, NLSY97 displays larger correlations between non-cognitive skill and both parental and own income compared to NLSY79. On the other hand, the correlations between incomes and both occupational contents have increased, though the increases are larger for cognitive content. From NLSY79 to NLSY97, the correlation between

parental income and non-cognitive occupational content roughly doubled, and the correlation between non-cognitive content and income roughly tripled.

These individual-level findings align with the occupational-level results, indicating a decline in the economic return to cognitive skill and an increase in the economic re- turn to non-cognitive skill. Furthermore, the association between parental income and cognitive skill shows a decline, while the association between parental income and non- cognitive skill shows an increase. This suggests a potential adaptive parenting focus among parents; however, this interpretation should be approached with caution. Additionally, the correlations between cognitive and non-cognitive skills are more similar between the two cohorts.

[Figure 5 about here.]

Despite providing valuable insights, the raw correlations overlook two important points. First, it is uncertain to what degree the relationship between income and skills is confounded. Second, the full non-linear relationship between skills and income is not depicted. These limitations are addressed in Figure 5. Figure 5 shows the observed association (solid line) and estimated causal effect of skills on income (dashed line) in both cohorts of NLSY. In accordance with standard practice in causal inference with continuous treatment, the causal relationship is visualized using a dose-response function (DRF). The DRF illustrates the relationship between the treatment (skills) and the expected potential income that would be observed if everyone were assigned to that particular skill level. The first column of Figure 5 shows that both the observed association and the deconfounded effect of cognitive skill are smaller in NLSY97 than NLSY79, and the observed association and the DRF take a similar shape, though the slope is smaller for

DRF since it is deconfounded. By examining the non-linear relation- ship between skill and income, it becomes evident that the between-cohort difference in return to cognitive skill exists mostly at the top skill level. Unlike the NLSY79 dataset, there is no further increase in income for individuals with the highest levels of cognitive skill in the NLSY97 dataset. This is indicated by the plateau observed in the DRF after approximately the 75th percentile rank.

On the other hand, the effect of non-cognitive skill is larger in NLSY97 than in NLSY79, also most notably at the highest skill level. In NLSY79, there is a negative association between non-cognitive skill and income at the top skill level. After adjusting for confounding, the negative association diminishes substantially but does not reverse. By contrast, in NLSY97, the positive relationship between non-cognitive skill and income is the most pronounced at top skill level.

[Figure 6 about here.]

The last figure reported that the association between income and skills is non- linear, and the between-cohort difference in this association occurs only in a specific part of the skill distribution. Is it possible that the association between parental income and skills is also non-linear, and that the between-cohort difference in skill disparity by parental income is driven by specific income groups of parents? Figure 6 illustrates the observed relationship between parental income and two skills. In NLSY79, there is a higher disparity in cognitive skill based on parental income, particularly among children from the top 20 percent income group who display exceptionally high cognitive skills. However, in NLSY97, this advantage is less pronounced. Conversely, for non-cognitive skill, the disparity by parental income is lower in the older cohort

due to the lack of association between non-cognitive skill and parental income among top-income families. In summary, the between-cohort difference in skill disparity by parental income appears to be driven by parents from the top income group. In the older cohort, where the return to top cognitive skill is significant and the return to top non-cognitive skill is not, children from high-income families possess high cognitive but not non-cognitive skill. In the newer cohort, this pattern is reversed, indicating that children from high- income families have higher non-cognitive skill but not as much advantage in cognitive skill. Again, further research is needed to determine whether these concurrent trends are coincidental or driven by reactive factors.

[Figure 7 about here.]

The between-cohort comparison of causal effects of different skills on income and disparities in different skill levels by parental income provides strongly indicates a shift of transmission channel between the two cohorts. However, it does not directly address that theoretical goal. Figure 7 presents the changes in the mechanisms through which skills influence income over time. For cognitive skill, primary maintenance through eco- nomic means accounts for approximately 14 percent (0.06/0.42) of the intergenerational income association. This means that if we equalize cognitive skill by parental income within population subgroups defined by confounders, about 14 percent of income immobility could be reduced. Since there is still a net association between parental income and cognitive skill even after accounting for confounders, further equalization of cognitive skill by parental income could lead to a greater reduction in immobility, as indicated by the magnitude of primary maintenance, which accounts for about 31 percent (0.13/0.42) of total immobility. Additionally, intergenerational income transmission

occurs through the transmission of occupational skill level net of personal skill development. Therefore, secondary maintenance contributes to approximately 14 percent of immobility (0.06/0.42). Overall, the cognitive channel accounts for about 45 percent of total intergenerational income transmission. When comparing the two cohorts, the cognitive channel is larger in the older cohort than in the new cohort, primarily due to differences in primary maintenance, whereas secondary maintenance between the two cohorts is similar. Therefore, the decline in the cognitive channel is attributed to growing disparity in cognitive skill rather than an increase in direct parental influence on occupational attainment.

Similar to the case of the cognitive channel, the non-cognitive channel also substantially contributes to intergenerational income persistence. The channel is composed of both primary and secondary maintenance, and it is primary rather than secondary maintenance that drives the between-cohort change. However, there are a few notable differences between the non-cognitive and cognitive channels. Firstly, the transmission through the non-cognitive channel is generally smaller in magnitude. For example, in NLSY97, the primary maintenance and total maintenance in the non-cognitive channel are 0.11 and 0.16, respectively, whereas in the cognitive channel, they are 0.12 and 0.17 respectively. Second, unlike that through cognitive skill, the maintenance through non-cognitive skill is increasing in magnitude. In fact, the increase in primary maintenance through non-cognitive skill almost exactly offsets the decrease in primary maintenance through cognitive skill. Overall, Figure 7 provides evidence supporting the effectively maintained immobility hypothesis, that the intergenerational transmission channel through cognitive skill is decreasing and is offset by the increasing importance of the non-cognitive channel.

DISCUSSION

What accounts for the consistent intergenerational income immobility over time? To address this question, this paper examines the transformations within the labor market that influence income outcomes for workers across different skill levels, and it assesses how these skill levels are linked with parental income over time. Findings indicate that changes within the labor market have played a role in perpetuating intergenerational inequality. Although the cognitive ability gap between children from diverse socio-economic backgrounds has lessened, the economic returns to cognitive ability have witnessed a simultaneous decline. This suggests that while the convergence of cognitive ability potentially engenders an enhancement in income mobility, the impact of this convergence is diminished due to the diminishing relevance of cognitive ability in the labor market.

However, the equalization of cognitive ability could still contribute to improving income mobility, albeit to a lesser extent, if not for another parallel trend in the labor market that has counteracted this potential improvement. Non-cognitive skills have become increasingly valued and sought after in the labor market, yet parental influence on the development of non-cognitive skills has not weakened as it has for cognitive skills. In fact, the association between parental income and children's non-cognitive skills has strengthened, reinforcing the intergenerational transmission of income through non- cognitive skills. This counterbalancing effect offsets the weakening mechanism through cognitive skills, ultimately resulting in the well-recognized stable rate of income mobility over time.

The persistence of inequality is a puzzle that has captivated scholars across generations. This research contributes to the understanding of this persistence by adopting a labor market perspective to explore the maintenance of immobility. This perspective reiterates two important

points which have not received sufficient attention by the majority of existing studies. First, it emphasizes the need to consider multiple mechanisms of intergenerational status transmission simultaneously. Recognizing the complexity of intergenerational association, focusing solely on one factor can lead to an oversimplified and incomplete understanding of intergenerational mobility. Second, the specific conditions of the labor market play a crucial role in shaping income inequality. Though the labor market has been accepted as a key channel of intergenerational inequality reproduction since the advent of "status attainment model", little is known about how changes in labor market skill demand affect income transmission.

These ideas have been underexplored as they have been overshadowed by the dominant approach that primarily focuses on education and cognitive ability as the main mediator of intergenerational transmission. Indeed, for many social observes both within sociology and beyond, meritocracy is almost equated with a distribution system based on cognitive ability, and intergenerational transmission is equated with the transmission of cognition. Bukodi and Goldthorpe (2021) describes the current status of intergenerational inequality as a "virtual cognitive caste", where highly educated parents cultivate their children to possess high cognitive abilities in order to preserve their privileged position. Similarly, Goodhart (2020) argues that mobility level is low because "smart produces smart". This research does not seek to undermine or replace the education- centered perspective, which undoubtedly provides significant explanatory power. How- ever, it underscores the necessity to consider the non-cognitive channel of transmission, which has become increasingly important in the recent labor market climate. This call resonates with a growing body of recent research that reflects a "cultural turn" in the study of intergenerational mobility (Jackson et al., 2005; Jackson, 2009; Rivera, 2012;

Armstrong and Hamilton, 2013; Rivera, 2016; Laurison and Friedman, 2016; Lareau, 2015; Friedman, 2016).

The analytical framework developed in this research offers a principled approach for studying the mechanism of intergenerational transmission. While the mechanism of intergenerational immobility has long been of interest to scholars in the field of stratification, previous analyses often relied on linear regression models, which are better suited for examining individual income determinants rather than assessing the channels of income transmission. The proposed framework in this research takes a different approach by decomposing total immobility into specific channels. This decomposition can be viewed as an extension of the Kitagawa-Blinder-Oaxaca decomposition, which is typically used to decompose inequality by a binary group indicator. In this case, how- ever, the focus is on decomposing inequality by a continuous variable, namely parental income. Furthermore, it is possible to interpret this decomposition as causal under certain assumptions, a point discussed in detail in the Appendix. In this regard, it can be seen as an extension of "the gap-closing estimand" to the context of continuous treatment.

This research can be improved on in several ways by future studies. First, there is no sufficient evidence to examine whether the changes in the skill advantage of children with high-income parents are coincidental or are a result of active reorientation of parental focus. It is possible that parents may have anticipated the changing labor market conditions and thus opted to invest more in their children's non-cognitive skills and less on cognitive ones. However, testing this hypothesis is beyond the scope of this research. Future studies can take on this task with more comprehensive information about parenting practices.

Second, though the results can be interpreted as causal effects under certain assumptions, I do not intend to make such a claim due to the lack of experimental evidence. Future research

may benefit from employing this decomposition method alongside quasi-experimental evidence to yield more credible causal estimates. By doing so, we can aspire to attain not only a descriptive summary of the magnitude of differing transmission channels but also gain insight as to how mobility might improve if the mediators within these transmission channels are equalized by parental income through targeted policy intervention.

We find ourselves at a crucial juncture for examining the ramifications of labor market transformations on social inequality. Regrettably, current data constraints limit our capacity to fully grasp the extent of these changes. While my dataset effectively captures the fallout from the early 2000s bursting of the "dot-com bubble" as well as the prevalent patterns of automation and computerization from the 2000s to the 2010s, the recent rise in machine learning and artificial intelligence (AI) introduces a fresh technological shock to the labor market, the full impact of which is yet to be entirely discerned. With a plethora of routine analytic jobs being replaced by computers, and the advent of advanced AI models like ChatGPT blurring the lines between "routine" and "creative" work—thus threatening even non-routine analytical positions—a deeper understanding of how this latest surge in AI will shape the labor market and influence intergenerational inequality constitutes a worthwhile direction for future exploration.

APPENDICES

A. DATA APPENDIX

A.1 Variable Construction

Cognitive Skill:

The measures of cognitive skill, AFQT, are constructed following similar procedures in NLSY79 and NLSY97. First, a standardized achievement test that contains multiple subjects, ASVAB (Armed Services Vocational Aptitude Battery), was administered among respondents of both NLSY79 and NLSY97 in the baseline wave, when most of the respondents were in high school. Note that the tests were delivered in different forms, whereas in NLSY79 it was done in paper and pencile, in NLSY97 it was done on computer using Item Response Theory (IRT). The comparability of the two measures will be discussed in Appendix C6. Second, respondents are grouped into three-month age groups. Third, within each three-month age group and using sampling weight, NLS staff assigned percentiles for the scores for the tests on Mathematical Knowledge (MK), Arithmetic Reasoning (AR), Word Knowledge (WK), and Paragraph Comprehension (PC) based on the weighted number of respondents scoring below each score. Here, for NLSY79, percentiles are assigned to raw scores, and for NLSY97, percentiles are assigned to IRT theta scores. Percentile scores for WK and PC were added to get an aggregate Verbal score (V) for which an aggregated intra-group, internally normed percentile was then computed. The percentile scores for MK, AR and two times the aggregated percentile for V were then summed. Finally, within each group NLS staff computed a percentile score, using the weights, on this aggregate score, yielding a final value between zero and 0.99.

Non-cognitive Skill:

Non-cognitive skill in the main analyses is operationalized as the rank of the first principle component of four of the big five personalities, conscientiousness, agreeable- ness, extroversion, and (low) neuroticism. Each personality dimension is measured with two questions, which are listed as follows:

Please indicate how much you agree or disagree about the following statements (1-7). You should indicate how much the traits apply to you, even if one part of the pair is more like you than the other.

- 1. Extraverted, enthusiastic
- 2. Critical, quarrelsome
- 3. Dependable, self-disciplined
- 4. Anxious, easily upset
- 5. Open to new experiences, complex
- 6. Reserved, quiet
- 7. Sympathetic, warm
- 8. Disorganized, careless
- 9. Calm, emotionally stable
- 10. Conventional, uncreative

Here, item 1 and 6 measure extroversion. Item 2 and 7 measure agreeableness. Item 3 and 8 measure conscientiousness. Item 4 and 9 measure neuroticism. Item 5 and 10 measure openness to new experience. To construct the measure used in analysis, I 1) sum first four pairs of items with item 6, 2, 8, 4 reversely ordered, so that the higher the sore the more favorable the personality; 2) extracted the first principal component of the four personalities; 3) created the

within-birth-group percentile rank to be consistent with other key measures. The robustness of results to alternative measures of non-cognitive skill is discussed in Appendix C4.



Figure A1: Coverage of Occupations for All Waves of O*NET (2003 - 2022)

Occupational skill content:

The Occupational Information Network (O*NET) is a comprehensive database of worker attributes and job characteristics. As the replacement for the Dictionary of Occupational Titles (DOT), O*NET information is the nation's primary source of occupational information. O*NET randomly surveys workers across occupations during which they are asked batteries of questions on the contents of their jobs. Initiated in 2001, and as of May 2023, O*NET has experienced a total of 27 major releases, and longitudinal information about many of the occupations have been collected. Figure A1 shows the coverage of occupations across waves of O*NET.

The cognitive and non-cognitive requirements are extracted from three modules from the worker survey: skills, activities, and knowledge requirements. In these modules workers are required to evaluate the importance of items in their work environment. The questions are worded in the following way: "How important is XXX to the performance of your current job? (on a scale of 1 to 5)". The occupational cognitive and non- cognitive variables are calculated in the following way. First, for cognitive skill, I average the importance of the following skills: reading comprehension (2A1a), writing (2A1c), mathematics (2A1e, 2C4a), science (2A1f), complex problem solving (2b2i), and creative thinking (4A2b2). For non-cognitive skill, I average the importance of the following skills: social perceptiveness (2B1a), coordination (2B1b), persuasion (2B1c), negotiation (2B1d), service orientation (2B1f), assisting and caring for others (4A4a5), coordinating activities (4A4b1), developing teams (4A4b2), interpreting others (4A4a1), and communication inside and outside organizations (4A4a2, 4A4a3). These represent occupation-year-level cognitive and non-cognitive content.

For individual-level analyses, occupation-year level work contents from O*NET are merged with NLSY97. An individual's cognitive and non-cognitive occupational con- tent is calculated first as the average of the contents of his/her jobs from the age 25 to 34. Then, a within-birth-year percentile rank of occupational cognitive/non-cognitive content is generated for all respondents to reflect the relative position of one's job skill requirement compared with peers. The within-cohort percentile rank operationalization is also consistent with personal skill, income, and parental income measures. For NLSY79 respondents, because O*NET information is unavailable, occupational content information is extrapolated from NLSY97. Specifically, an individual's cognitive/non- cognitive occupational content is constructed as the cognitive/non-cognitive content of the individual with the same occupation in NLSY97.

A.2 Match of Occupations across Data Sources

NLSY79 uses 1970 occupation code (OCC1970). NLSY97 uses 2002 occupation code (OCC2002). O*NET uses multiple versions of Standard Occupational Classifica- tion (SOC).

Analyses of occupational content require cross-walking between NLSY79/97 and O*NET. First, occupation codes of all O*NET waves are transformed to SOC2010 using crosswalks provided by O*NET.

For NLSY79, OCC1970 is first cross-walked to OCC2010 (using crosswalking files supplied by Deming (2017)), and OCC2010 is cross-walked to SOC2010 (using R pack- age "occupationcross"). For NLSY97, OCC2002 is first cross-walked to OCC2010 using a crosswalk supplied by the census bureau, and OCC2010 is cross-walked to SOC2010 (using R package "occupationcross").

Occupational under- or over-match between data sources in this research is less harmful than other settings where precise occupational crosswalk is required, because in the eventual analysis occupation is scaled as continuous variables (cognitive/non- cognitive content), and the exact occupation variable is not used.

A.3 Descriptives

This section presents the descriptives of control variables, non-cognitive skill mea- sures, and occupational skill contents. Cognitive skill measure is in percentile rank in its raw form as acquired from NLSY. Income measures are operationalized as per- centile ranks in the analyses. Descriptive statistics about cognitive skill and income are omitted as they are percentile ranks.

	NLSY79			NLSY97		
Variable	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Race	12686			8984		
-Black	2002	16%		2335	26%	
-Hispanic	3174	25%		1901	21%	
-White	7510	59%		4748	53%	
Sex	12686	1.5	0.5	8984	1.5	0.5
Siblings	12668	3.9	2.6	7406	2.8	2.4
Residence	12646			6711		
-Urban	9970	79%		5227	78%	
-Rural	2676	21%		1484	22%	
Father alive	8179	0.6	0.49	8750	0.043	0.2
Father years of education	10880	11	3.9	7120	13	3.2
Live with both Parents	10465	0.6	0.49	8984	0.49	0.5
Region	12447			8984		
-Northeast	2550	20%		1585	18%	
-Northcentral	2934	24%		2050	23%	
-South	4568	37%		3359	37%	
-West	2395	19%		1990	22%	

Table A1: Descriptive Statistics of Control Variables for NLSY79 and NLSY97

The non-cognitive skills are measured with Ten-Item-Personality-Inventory (TIPI). For each of the five dimensions (conscientiousness, openness, agreeableness, extraver- sion, neuroticism), two questions are asked. In the analyses these variables are op- erationalized as a percentile rank. For details about questions wordings and how the variables are used in the analyses, refer to Appendix B1, Appendix C4, and main text.

	NLSY79							
Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max	
Conscientiousness 1	6985	6	1.5	1	6	7	7	
Conscientiousness 2	6989	5.5	1.8	1	4	7	7	
Openness 1	6942	4.7	1.9	1	3	6	7	
Openness 2	6970	5.2	1.6	1	4	7	7	

Agreeableness 1	6849	4.8	1.9	1	3	7	7
Agreeableness 2	6977	5.6	1.6	1	5	7	7
Extroversion 1	6883	5	1.8	1	4	7	7
Extroversion 2	6989	3.8	2.1	1	2	6	7
Neuroticism 1	6991	4.8	1.9	1	3	6	7
Neuroticism 2	6980	5.4	1.6	1	5	7	7
			NLSY97				
Conscientiousness 1	7434	6.1	1.1	1	6	7	7
Conscientiousness 2	7454	5.3	1.7	1	4	7	7
Openness 1	7404	5.2	1.6	1	4	7	7
Openness 2	7410	5.7	1.3	1	5	7	7
Agreeableness 1	7195	4.4	1.7	1	3	6	7
Agreeableness 2	7395	5.5	1.3	1	5	7	7
Extroversion 1	7445	4	2	1	3	6	7
Extroversion 2	7363	5.3	1.4	1	5	6	7
Neuroticism 1	7457	4.4	1.8	1	3	6	7
Neuroticism 2	7427	5.5	1.4	1	5	6	7

Table A2: Descriptive Statistics of Non-Cognitive Skills in NLSY79 and NLSY97

Finally, Table A3 reports descriptive statistics of occupational work contents. It reports the summaries of the detailed items used to build the cognitive and the non-cognitive measure, as well as the average cognitive and non-cognitive content. The units are occupation-year. Be noted that in the individual-level analyses the cognitive and non-cognitive contents are further transformed into within-cohort percentile ranks.

			Cognitive Skill Content				
Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
reading comprehension	2827	3.6	0.6	1.7	3.1	4	5
writing	2827	3.2	0.66	1.2	2.8	3.8	5
mathematics	2827	2.7	0.68	1	2.1	3.1	5
science	2827	2.1	0.93	1	1.4	2.6	5
complex problem solving	2827	3.2	0.54	1.4	2.9	3.6	4.8
mathematic skill	2296	3.1	0.7	1.1	2.5	3.5	5
creative thinking	2296	3.3	0.7	1.1	2.8	3.8	5
Average	2933	3	0.52	1.5	2.6	3.4	4.8
			Non-Cognitive Skill Content				
social perceptiveness	2827	3.2	0.53	1	2.9	3.6	4.9
coordination	2827	3.2	0.45	1.4	3	3.6	4.7
persuasion	2827	2.8	0.55	1	2.4	3.1	4.7
negotiation	2827	2.7	0.54	1	2.3	3	4.8
assisting and caring for others	2296	2.9	0.87	1	2.2	3.3	5
service orientation	2827	3	0.57	1	2.6	3.4	4.7

coordinating activities	2296	3.1	0.58	1.1	2.7	3.5	4.9
developing teams	2296	3	0.61	1	2.5	3.4	4.7
interpreting others	2296	3.3	0.67	1	2.8	3.8	5
communication inside organizations	2296	3.9	0.48	1.4	3.7	4.3	4.9
communication outside of organizations	2296	3.3	0.75	1	2.8	3.8	5
Average	2933	3.1	0.46	1.2	2.8	3.4	4.4

Table A3: Summary Statistics of Occupational Work Contents from O*NET 2003-2022

B. METHODOLOGY APPENDIX

B.1 Causal Identification

To begin with, let M^* denotes the post-intervention value of M. The first-step of intervention is to remove primary maintenance through economic means, that is the conditional inequality of skill by parental income. Take children from poor parents, $A = 0$, as targets of intervention, the formal expression of the intervention is $P(m^*|x, A = 0) = P(m|x, A = 1)$, where we equalize the skill level of poor children to be the same as rich children, conditional on X. I denote $Y(M_1^*)$ as the potential outcome after this conditional intervention where the distribution of M is intervened to follow those from $A = 1$ (hence the 1 in subscript) conditionally, and the post-intervention income of poor children is expressed as $E[Y(M_1^*)|A = 0]$.

The next step of the intervention is to marginally equalize skill levels between children with different parental income. The intervention takes the next formal form $P(m^{**}|A = 0) = P(m|A = 1)$. Notice that I use the double asterisk to denote the marginal intervention and use the single asterisk to denote the conditional intervention. This intervention completely removes primary maintenance. After the intervention, the expected income of poor children is expressed as $E[Y(M_1^{**})|A = 0]$.

The two causal quantities are identified with the following assumptions:

Assumption 1. *System invariance*: $Y^*(m) = Y(m)$ for all m.

Assumption 2. *SUTVA*: $Y = Y(M)$ and $Y^* = Y(M^*)$.

Assumption 3. *Ignorability* : $Y(m) \perp M|A, X$ and $Y^*(m) \perp M^*|A, X$ for all m.

Assumption 4. *Positivity*: $0 < P(m|X = x, R = r) < 1$ for all m, x, and r.

$$
\begin{aligned}
E[Y(M_a^*)|A=0] &= \int\int E[Y(M_a^*)|A=0, M^*=m, x]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{A2}} \int\int E[Y^*(m)|A=0, M^*=m, x]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{A3}} \int\int E[Y^*(m)|A=0, x]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{A1}} \int\int E[Y(m)|A=0, x]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{A3}} \int\int E[Y(m)|A=0, x, M=m]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{A2}} \int\int E[Y|A=0, x, M=m]P^*(m|x, A=0)P(x|A=0)dmdx \\
&\underbrace{=}_{\text{Specification}} \int\int E[Y|A=0, x, M=m]P(m|x, A=a)P(x|A=0)dmdx
\end{aligned}
$$

The same identification result applies for $E[Y(M_a^{**}|A = 0)]$, except for the final step where the intervention would be $P^*(m|x, A = 0) = P(m|A = a)$. The identification essentially shows the causal interpretation of the descripitive decomposition this research applied.

On the basis of the marginal equalization of M, a further step of the intervention is to equalize S, the occupational skill content, by family income. Therefore, the in- tervention contains not only $P(m^*|A = 0) = P(m|A = 1)$, but also the intervention on occupations, $P(s^{**}|m, A = 0) = P(s|m, A = 1)$. The post-intervention expected income of poor children is expressed $as\ E[Y(M_1^{**}, S_1^{**})|A = 0]$[10]. For the identification of the sequential estimand, $E[Y(M_a^{**}, S_a^{**})|A = 0]$, sequential ignorability is required.

Assumption 1*. *System invariance*: $Y^*(m, s(m)) = Y(m, s(m))$ for all m, s.

[10] Note that a more rigorous notation for the potential outcome would be $Y(M_1^{**}, S_1^{**}(M_1^{**}))$ since the value of M carries over like the conventional causal mediation setting. The notation is suppressed for readability.

Assumption 2*. *SUTVA*: $Y = Y(M, S(M))$ and $Y^* = Y(M^*, S^*M(^*))$.

Assumption 3*. *Ignorability of M* : $Y(m, s) \perp M|A, X$ and $Y^*(m, s) \perp M^*|A, X$ for all m, s.

Assumption 4*. *Ignorability of S* : $Y(m, s) \perp S|M, A, X$ and $Y^*(m, s) \perp S^*|M^*, A, X$ for all m, s.

Assumption 5*. *Positivity*: $0 < P(m|X = x, R = r) < 1$ for all m, x, and r; $0 < P(m|X = x, R = r) < 1$ for all s, m, x, and r.

Here Assumption 3* and Assumption 4* are jointly referred to as sequential ignorability. Notice that because the value of M always carries over to S and we are uninterested in estimating cross-world counterfactuals like $Y(1, S(0))$, the assumptions are less strict than the estimation of natural direct/indirect effects. With the above assumptions, the sequential decomposition estimand is estimated as follows:

$$E[Y(M_a^{**}, S_a^{**})|A = 0]$$
$$= \int\int\int E[Y(M_a^{**}, S_a^{**})|A = 0, M^* = m, S^* = s, x]P^*(m|\sim)P^*(s|\sim)P(X|A = 0)dmdsdx$$

For exposition ease I focus on the identification of $E[Y(M_a^{**}, S_a^{**})|A = 0, M^* = m, S^* = s, x]$, as the rest is simply intervention specification.

$$
\begin{aligned}
E[Y(M_a^{**}, S_a^{**})|A=0, M^*=m, S^*=s, x] &\underbrace{=}_{\text{A2}^*} E[Y^*(m,s)|A=0, M^*=m, S^*=s, x] \\
&\underbrace{=}_{\text{A4}^*} E[Y^*(m,s)|A=0, M^*=m, x] \\
&\underbrace{=}_{\text{A3}^*} E[Y^*(m,s)|A=0, x] \\
&\underbrace{=}_{\text{A1}^*} E[Y(m,s)|A=0, x] \\
&\underbrace{=}_{\text{A3}^*} E[Y(m,s)|A=0, M=m, x] \\
&\underbrace{=}_{\text{A4}^*} E[Y(m,s)|A=0, M=m, S=s, x] \\
&\underbrace{=}_{\text{A2}^*} E[Y|A=0, M=m, S=s, x]
\end{aligned}
$$

The method can be seen as a generalization of "the gap-closing estimand" to the setting of continuous and multiple mediators Lundberg (2021), and is a new development to the nascent literature on causal intervention in epidemiology and sociology (VanderWeele and Robinson, 2014; Jackson and VanderWeele, 2018; Jackson, 2018, 2021; Lundberg, 2021; Wei and Xie, 2022).

B.2 Relation between Descriptive, Interventional, and Mediational Decomposition

The decomposition I propose is related to causal interventional and causal mediation decomposition in critical ways. As a matter of fact, whether to interpret the same result as descriptive, interventional, or mediational is up to the assumptions that the researcher is willing to impose. The building block of the decomposition is $\theta(a, a') \coloneqq \int\int E[Y|m,x,a]P(m|x,a')P(x|a)dmdx$. Without any additional identification assumption, $\theta(a, a')$ can be used to construct a special type of non-parametric KBO decomposition, where total inequality between group a and group a′ can be de- composed into the "explained" and

"unexplained" component by group difference in M controlling X. Here, the term "explained" and "unexplained" are completely causal-neutral. The decomposition is a useful accounting exercise that should not be related to the potential consequence of manipulation or intervention.

On the other end of the spectrum, the exact same decomposition has a causal interpretation as the natural direct/indirect mediation effect decomposition, provided that we are willing to assume that 1) there is no unobserved confounder between M and Y; 2) there is no unobserved confounder between A and Y; 3) A does not cause X[11]; plus the usual positivity and no-interference assumption. With these assumptions, $\theta(a, a') = E\left[Y\left(a, M(a')\right)\right] \coloneqq \theta^{med}(a, a')$, where the first equation is an identification result and the second a definition. The mediation interpretation is what almost all prior analyses that try to investigate intergenerational mechanism implicitly takes. However, the assumptions 2 and 3 are highly improbable. Possible common causes of parental and children income are plentiful: grandparents' situation, genetic endowment, geographic location, etc. Social surveys hardly have a comprehensive collection of these information since it requires an extensive investigation of respondents' parental background. The assumption that no ability-income confounder is caused by parental income is also hard to hold. In fact, most of those confounders are likely results of parental income, like childhood family condition, school condition, and parental education investment. Therefore, it is highly questionable to formulate intergenerational transmission as a causal mediation problem, even if we disregard the conceptual dilemma that family origin is "immutable".

[11] There should not be treatment-induced mediator-outcome confounder -- a famous restriction that enables natural direct/indirect effect decomposition.

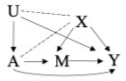
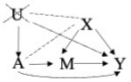
Statistical Estimand	Assumption	Causal Estimand	Decomposition
$\theta(a,a^{'}) = \int\int E[Y\|m,x,a]P(m\|x,a^{'})P(x\|a)dmdx$	NA	NA	Non-Parametric KBO Decomposition: $E[Y\|a] - E[Y\|a^{'}] =$ $\underbrace{\theta(a,a) - \theta(a,a^{'})}_{\text{Explained Component}} + \underbrace{\theta(a,a^{'}) - \theta(a^{'},a^{'})}_{\text{Unexplained Component}}$
		$\theta(a,a^{'}) =$ $E[Y(M_{a^{'}})\|a]$ $= \theta^{int}(a,a^{'})$	Interventional Decomposition: $E[Y\|a] - E[Y\|a^{'}] =$ $\underbrace{\theta^{int}(a,a) - \theta^{int}(a,a^{'})}_{\text{Interventional Effect}} + \underbrace{\theta^{int}(a,a^{'}) - \theta^{int}(a^{'},a^{'})}_{\text{Residual Disparity}}$
		$\theta(a,a^{'}) =$ $E[Y(a,M(a^{'}))]$ $= \theta^{med}(a,a^{'})$	Natural Effect Decomposition: $E[Y(a)] - E[Y(a^{'})] = E[Y\|a] - E[Y\|a^{'}] =$ $\underbrace{\theta^{med}(a,a) - \theta^{med}(a,a^{'})}_{\text{Indirect Effect}} + \underbrace{\theta^{med}(a,a^{'}) - \theta^{med}(a^{'},a^{'})}_{\text{Direct Effect}}$

Figure A2: The Relation between Descriptive KBO Decomposition, Interventional Decomposition, and Natural Direct/Indirect Effect Decomposition

Note: For the interventional decomposition to hold, one also needs to assume SUTVA. For the causal mediation decomposition to hold, one also needs to assume SUTVA and positivity. Dashed line means that the temporal order is irrelevant.

As a central ground, to interpret the results of the decomposition as interventional effect, one only needs to assume no $M - Y$ unmeasured confounder. In other words, the only assumptions required are those that we need to identify the effect of M on Y. Admittedly, in an observation setting, it is often hard to control for all potential $M - Y$ confounders. But the interventional interpretation is still better than the mediational one because it avoids the conceptual dilemma where we need to consider parental origin as manipulable treatment and the substantive conundrum where we need to figure out the causes of parental income.

C. ROBUSTNESS

C.1 Comparability of Cognitive Measure across Cohort

In NLSY79, ASVAB is administered as a pen and paper test (PP), and the AFQT score is calculated as the sum of raw test scores. In NLSY97, ASVAB is administered as an IRT-based computer-assisted testing (CAT), and AFQT is calculated as IRT theta score. Though in the analysis AFQT from both cohorts are further normed to be within-cohort percentile ranks, the concern of measurement comparability still remains. Fortunately, the validation between PP and CAT versions of AFQT has been done by earlier researchers and the comparability between the two measures appears to be astonishingly high (Chapter 17 of Moreno and Segall (1997)).

Data collection happened in 1989-1990. Two thousand ninety male Navy recruits were randomly assigned to either the PP or the CAT version of ASVAB.

Test	Correlation	SE	p
Mathematical Knowledge	1.00	.015	.975
Arithmetic Reasoning	1.02	.021	.287
Word Knowledge	1.02	.017	.370
Paragraph Comprehension	1.11	.082	.145

Table A4: Disattenuated Correlations Between CAT and PP ASVAB Note: Adapted from Table 17-4 in Moreno and Segall (1997), page 173

Table A4 shows the disattenuated correlations between the two implementations of ASVAB. The correlations are unconstrained so they may exceed 1. From the Table it is clear that all tests used to construct AFQT are highly correlated (close to 1), and there is no evidence to reject the hypotheses that the correlations are equal to 1.

C.2 Comparability of the Non-Cognitive Measure across Cohort

The non-cognitive skill measure is based on TIPI, which was administered in both NLSY79 and NLSY97. However, in NLSY79, the question was asked in the year when respondents were between the age 50 and 57 (respondents were born between 1957 and 1964, and question was asked in 2014); in NLSY97, the question was asked in the year when the respondents were between the age 24 and 28 (respondents were born between 1980 and 1984, question collected in Round 12 in 2008). If there is significant life-course variation of personality measures, the comparison between NLSY79 and NLSY97 may be questionable.

However, only some forms of life-course variation would endanger the finding. The main result is that compared with NLSY79, the correlations between non-cognitive skill and both parental income and income are higher in NLSY97. To nullify the finding, it needs to be shown that the skill-income correlation in NLSY97, when skill was measured at a young age, is overestimated. Earlier studies on the life-course variation of non- cognitive skill suggest otherwise, which is that skill measurement at young age likely causes underestimation of skill-income correlation.

First, earlier studies established the observation that the correlation between non-cognitive skill and permanent income becomes higher as one ages. Soldz and Vaillant (1999) reported that compared with Big-Five traits measured at around age 22, traits measured at around age 67 are much more predictive of life income (pp223-224). This effectively suggests that the difference between skill-income correlations in NLSY79 and NLSY97 might be even higher than observed, because the correlation in NLSY97 might be underestimated as measurement happened at a young age.

Second, earlier studies established that non-cognitive skill is more stable or contains less measurement error at later ages, and the most stable measurement happens around age 50 (Specht et al., 2011). This means that skill-income correlation observed at a young age is likely to be attenuated downwards due to higher measurement error, so the correlation in NLSY97 is underestimated rather than overestimated. The true difference between 97 and 79 should be larger.

C.3 Comparability of Occupational Content across Data Sources

For respondents in NLSY79, I collected information about their occupations be- tween age 25 and 34, and take the average skill content of these occupations as the skill content measure. These respondents were born between 1957 and 1964, which means that the occupational information needed are from 1982 to 1997.

In the main analyses, for job content information before 2000, I used information from job ads as provided by (Atalay et al., 2020). However, despite best effort, the dataset is not strictly comparable with O*NET. Here I test the robustness of results under alternative choices of job content information before 2000.

First, I use O*NET 98, the earliest wave of O*NET, to measure the occupational skill contents of respondents in both NLSY79 and NLSY97. Since O*NET 98 relies on expert evaluation rather than worker self-report, for the main analyses O*NET 98 is not used due to concerns about its comparability with longitudinal O*NET after 2003. O*NET 98 can be a more reliable measure of work contents for the period 1982-1998. Table A5 shows that the association between parental income and both occupational contents have increased, the association between cognitive content and income has slightly decreased, and the association between non-

cognitive content and income has slightly increased. This result is overall consistent with the main finding.

	NLSY79		NLSY97	
	Parental Income	Income	Parental Income	Income
Cognitive Content	0.24	0.41	0.31	0.40
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Content	0.18	0.28	0.22	0.29
	(0.02)	(0.02)	(0.02)	(0.01)
N	6,064		5,511	

Table A5: Correlations between Parental Income/Income and Occupational Skill Content Measured by O*NET98

Second, I used the average skill content as observed in the multiple waves of O*NET to measure the skill content before 2000. The result is shown in Table A6. The correlations between cognitive skill/content and both parental income and own income are notably higher compared to the correlations between non-cognitive skill/content and these income measures. This suggests that the cognitive channel is important than the non-cognitive channel for both cohorts. However, the imbalance between the two channels is shrinking when comparing the two cohorts. NLSY97 exhibits smaller correlations between cognitive skill and both parental and own income compared to NLSY79. Conversely, NLSY97 displays larger correlations between noncognitive skill and both parental and own income compared to NLSY79. A similar trend is observed for the correlation between income and the non-cognitive content of occupation. From NLSY79 to NLSY97, the correlation between parental income and non-

cognitive occupational content more than doubled, and the correlation between non-cognitive content and income tripled.

	NLSY79		NLSY97	
	Parental Income	Income	Parental Income	Income
Cognitive Content	0.19	0.31	0.25	0.32
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Content	0.06	0.06	0.14	0.18
	(0.02)	(0.02)	(0.02)	(0.01)
N	3,968		4,140	

Table A6: Correlations between Parental Income/Income and Occupational Skill Content, Measured by calibrated APST for NLSY79, longitudinal O*NET for NLSY97

C.4 Alternative Operationalize of Income and Ability

There are possible alternative specifications of income and ability. This subsection reports basic results under these alternative specifications. Specifically, I test alternative measures of income and non-cognitive ability. Cognitive ability and occupational contents are left out. Cognitive ability in the analysis is the default measure provided by NLSY and is well-validated, so I do not test the robustness to alternative specification. The measure of occupational contents is straightforward and there is no clear alternative measure.

Table A7 reports the basic results under income specified as logged income instead of percentile ranks. The difference with the result under percentile measure is minimal.

	NLSY79		NLSY97	
	Parental Income	Income	Parental Income	Income
Cognitive Skill	0.40	0.41	0.35	0.36
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Skill	0.06	0.14	0.10	0.18

	(0.02)	(0.02)	(0.02)	(0.01)
Cognitive Content	0.19	0.28	0.22	0.31
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Content	0.06	0.06	0.12	0.17
	(0.02)	(0.02)	(0.02)	(0.01)
N	3,968		4,140	

Table A7: Correlations between Parental Income/Income and Skills/Occupational Skill Content (Alternative Income Measure)

Note: Replication of Table 1 but incomes are log incomes instead of percentile ranks.

Table A8 reports the correlations between income and personality measures. To ensure comparability the personalities are normalized to be within-cohort rank. Results using raw scores are almost identical and are omitted. There is heterogeneity of the trends of personality measures income. The correlations between emotional stability (low neuroticism) and both parental and own income have increased. So is conscientiousness. The correlations between extroversion and income, on the other hand, have decreased. The correlations between agreeableness and income changed little. Openness to experience has become much less predictive of income, whereas it is slightly more correlated with parental income.

As far as the skill-income association is concerned, two types of non-cognitive skills seem to show diverging trend. The lesson to be drawn is that intergenerational transmission is becoming more through training children to be stable and responsible workers, rather than training children to be outgoing or warm workers.

	NLSY79		NLSY97	
	Parental Income	Income	Parental Income	Income
Neuroticism (lack of)	0.03	0.10	0.08	0.12
	(0.02)	(0.02)	(0.01)	(0.01)

Extroversion	0.13	0.16	0.10	0.15
	(0.02)	(0.02)	(0.02)	(0.02)
Agreeableness	0.06	0.05	0.05	0.06
	(0.02)	(0.02)	(0.02)	(0.02)
Conscientiousness	-0.04	0.09	-0.01	0.10
	(0.02)	(0.02)	(0.02)	(0.02)
Openness	0.03	0.05	0.04	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
N	3,968		4,140	

Table A8: Correlations between parental Income/Income and Specific Personality Note: Replication of Table 1 but with five specific personality measures.

In the main analysis non-cognitive skill is operationalized as the first principal component as there is no natural way to weigh the importance of the multiple skill measures, and also because there is a long tradition of using factor analysis in the study of cognitive/non-cognitive skills. Another potential issue of the operationalization in the main analysis is that the personality "openness to experience" is dropped as it reflects more cognitive rather than non-cognitive ability.

To address the concerns, in Table A9, I report the correlations between income and non-cognitive skill incorporating "openness", and non-cognitive skill based on raw score. The raw-score based measure is again normalized as within-cohort percentile rank to be consistent with the measure of parental and own income. The results are similar between raw-score based and principle component based measure, and similar whether incorporating "openness" or not.

C.5 Missingness

In the main analyses missing covariates are multiply imputed using the "mice" package in R. The results are almost identical without imputing missing values.

There are two places in the analyses where covariates are used, which are the demonstration of the effect of skill on income, and the decomposition of transmission channel. Figure A3 show the estimated causal effect of skill on income but without imputing missing values in covariates. The observed non-linearity still holds using the non-imputed dataset.

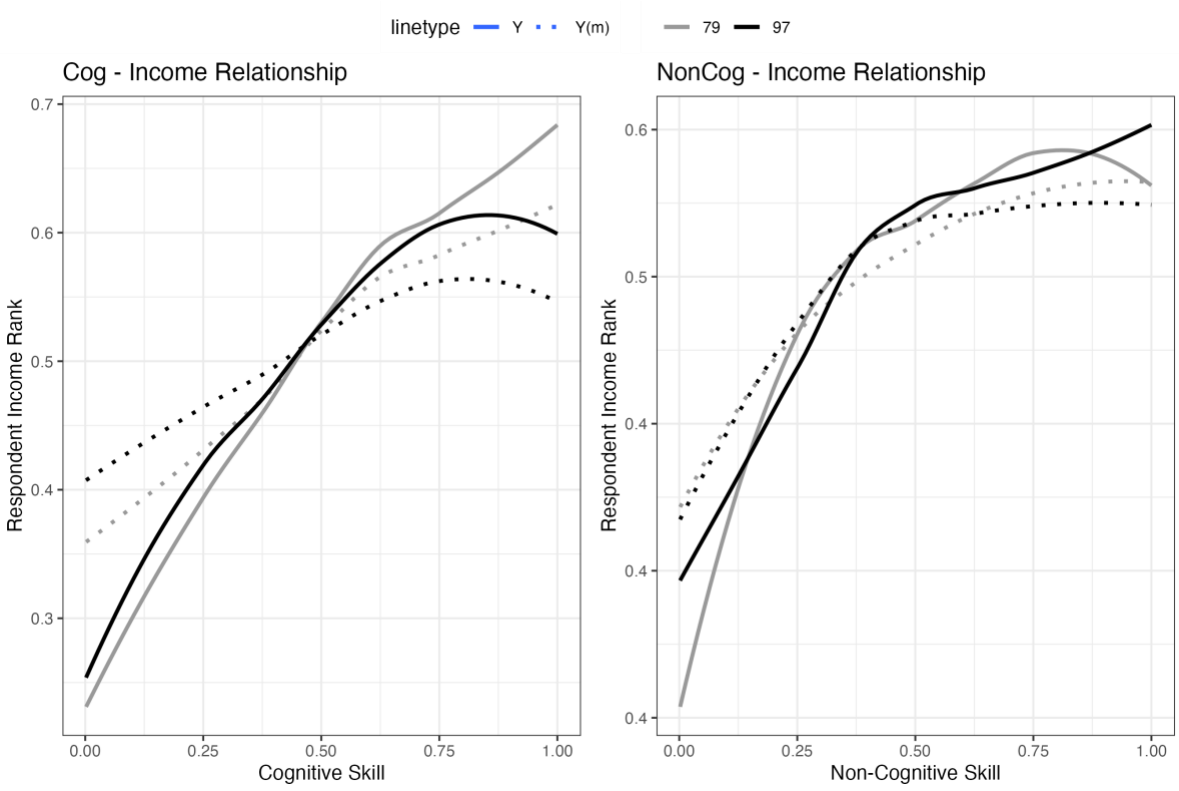


Figure A3: Observed Association and Causal Effect of Personal Skills on Income in NLSY79 and NLSY97

Note: Replication of Figure 5 without imputing missing values in control variables.

FIGURES AND TABLES

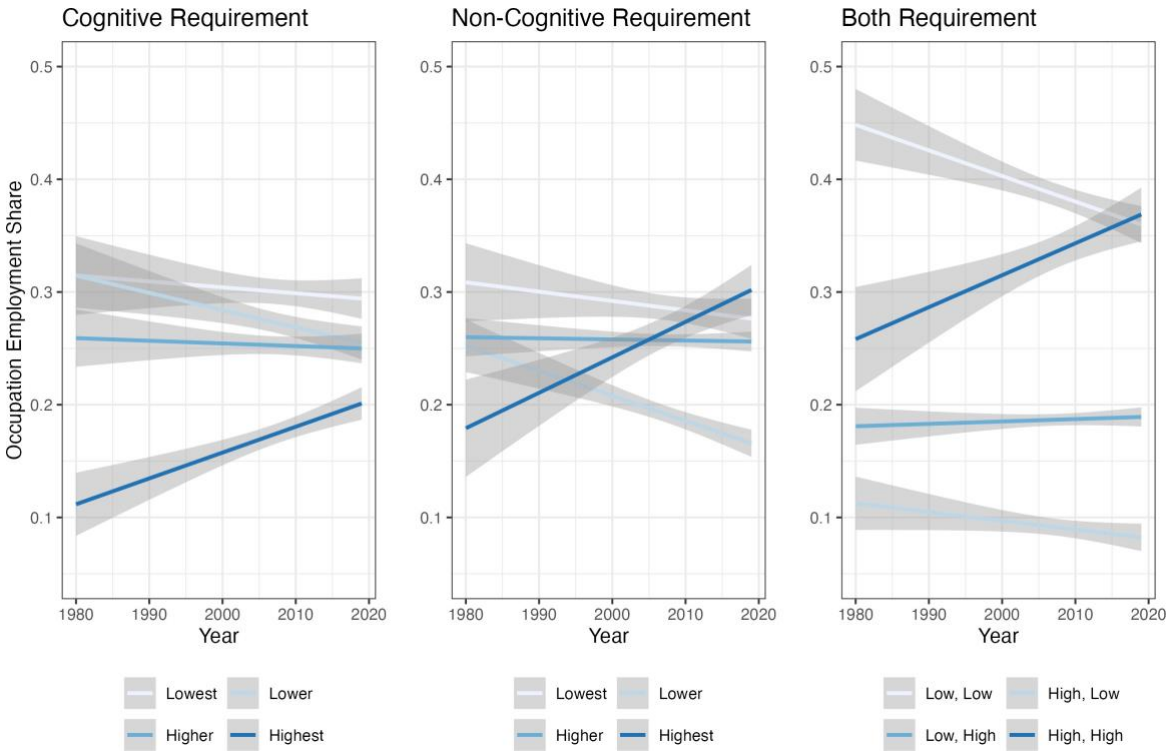


Figure 1: Change of Employment Share by Occupational Skill Content in the US

Note: Data from 1980 5 percent samples, 1990 5 percent samples, and 2000–2019 ACS 1 percent samples, all extracted from the Integrated Public Use Micro-data Series (IPUMS). Skills are standardized as mean 0 and sd 1. Whether an occupation is in ”Lowest”,”Lower”,”Higher”, or ”Highest” categories is determined by the quartiles of skill requirement the occupation is in among all occupations, and the skill requirement is measured as the average from all O*NET waves. In the third column, whether an occupation is in ”Low” or ”High” in is defined by whether the occupation’s skill requirement is in the first or second half. The third column resembles Figure IV in (Deming, 2017) closely. Shaded areas are 95 percent intervals of predicted values.

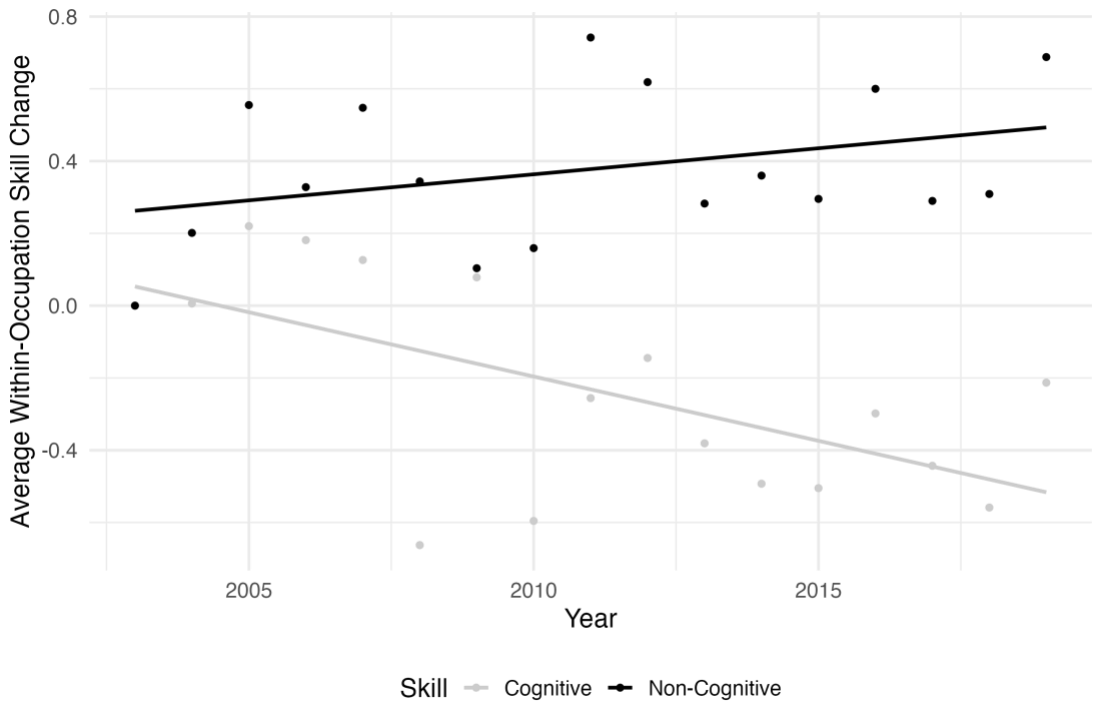


Figure 2: Within-Occupation Change of Skill Content from 2003 to 2019

Note: Data from O*NET and 2003–2019 ACS 1 percent samples. Skills are standardized as mean 0 and sd 1. Each point represents a year fixed effect from the regression "Skill = Year FE + Occupation FE" weighted by the number of people in the occupation as from ACS. Values for year 2003 are set to 0. Lines are linear approximations of the year FEs with year.

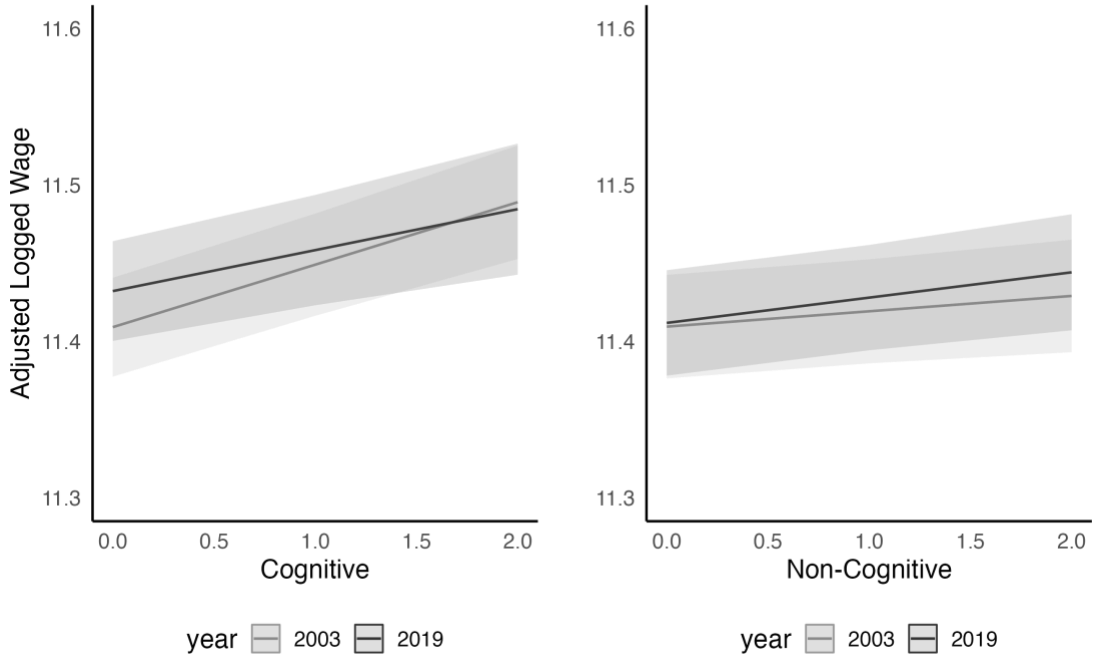


Figure 3: Observed Return to Cognitive and Non-cognitive Skills 2003-2019

Note: Data from O*NET and 2003–2019 ACS 1 percent samples. Skills are standardized as mean 0 and sd 1. Wage is adjusted for inflation. The lines represent the conditional effect of skills on wage from the regression "Wage = Year*Skill + Occupation FE" weighted by the number of people in the occupation as from ACS. Shaded areas are 95 percent confidence bands.

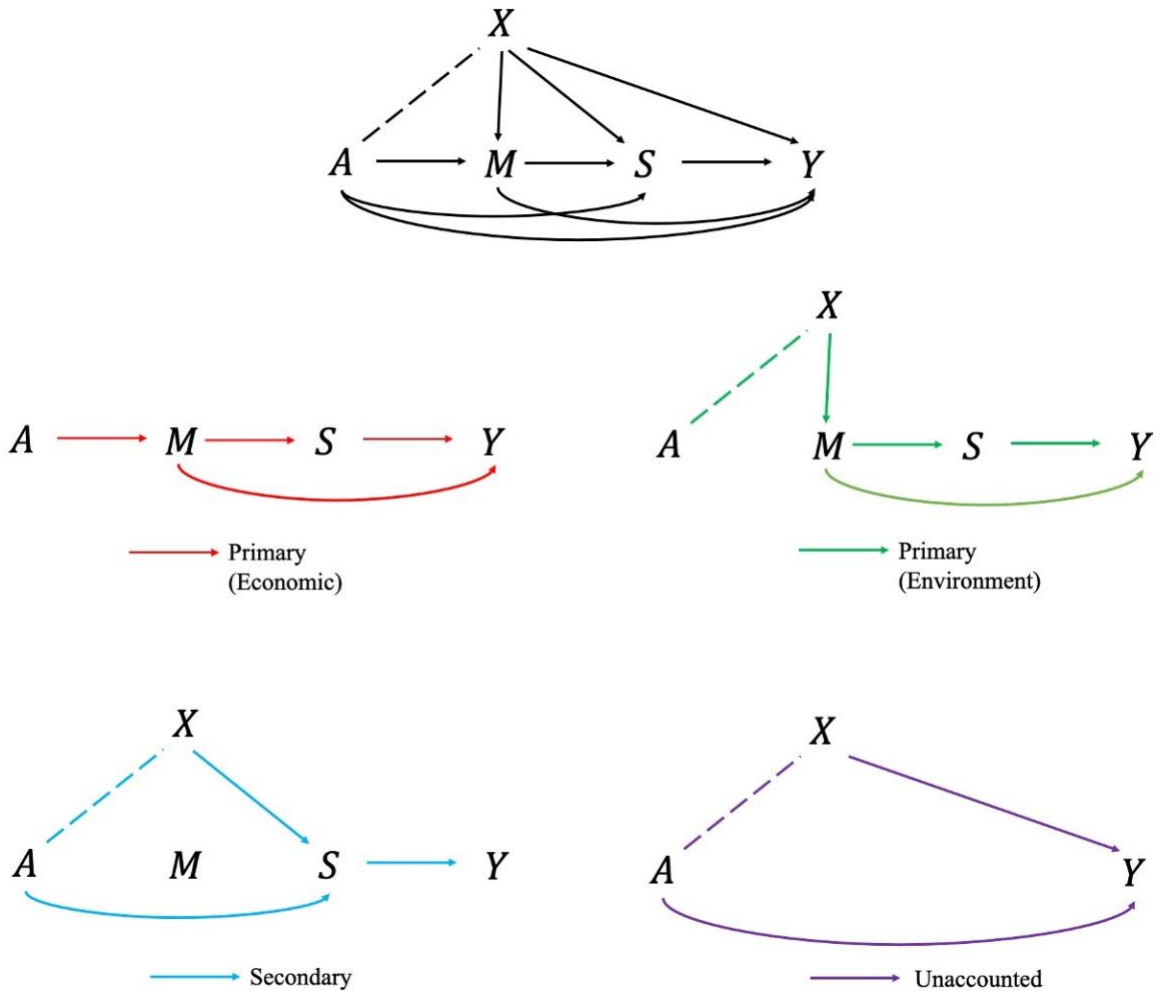


Figure 4: The Channels of Intergenerational Income Transmission through the Labor Market Path

Note: X denotes covariates. A denotes parental income. M denotes children skill. S denotes children occupational skill content. Y denotes Children income.

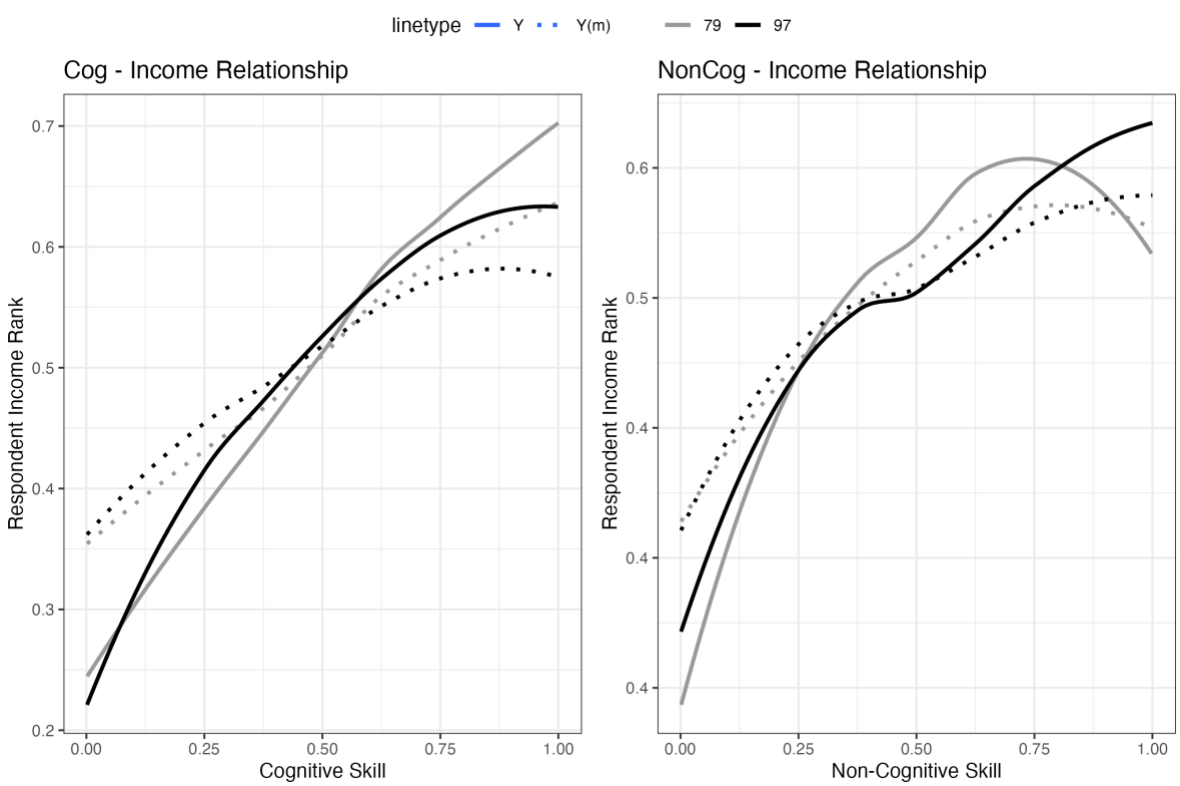


Figure 5: Observed Association and Causal Effect of Personal Skills on Income in NLSY79 and NLSY97

Note: Solid lines are observed association between skill and income. Dashed lines are association between skill and potential income given the skill, also known as the dose-response curve. The smoothing is done using “loess”. The potential incomes, that is the values on the dashed line, are predicted using random forest. Skill and income are percentile ranks.

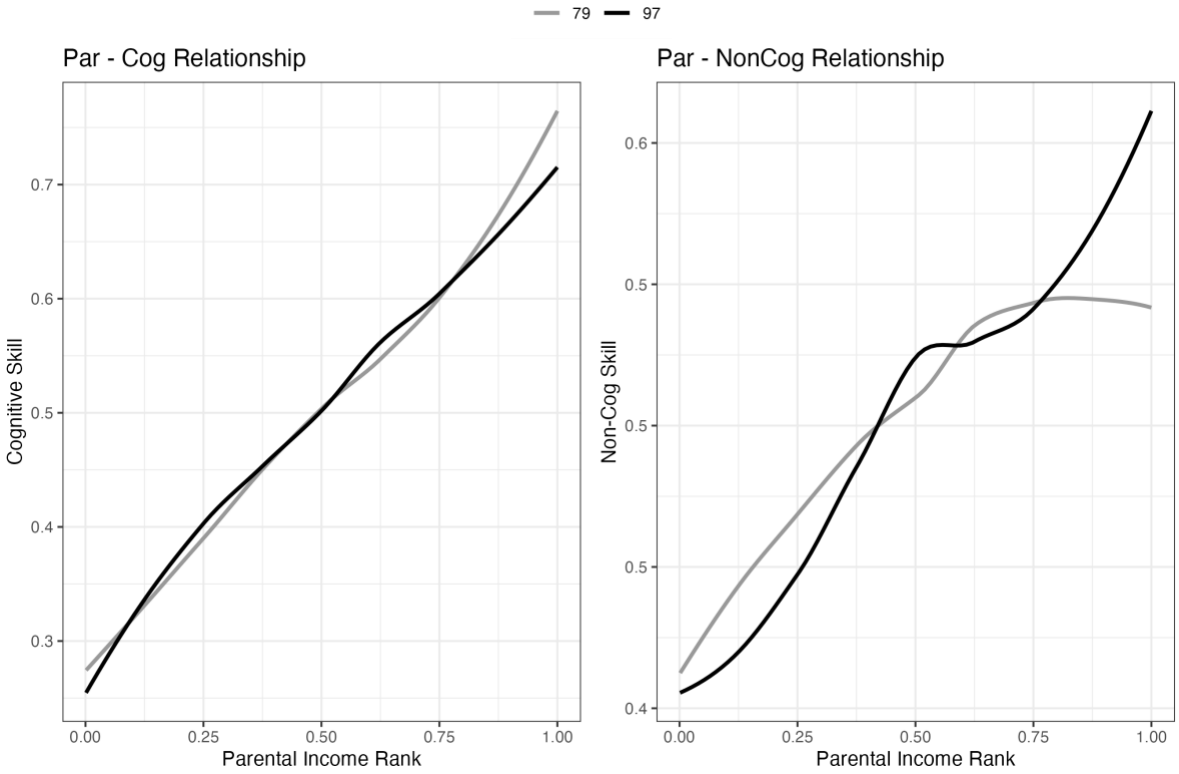


Figure 6: Observed Association between Parental Income and Personal Skills in NLSY79 and NLSY97

Note: The smoothing is done using “loess”. Skill and income are percentile ranks.

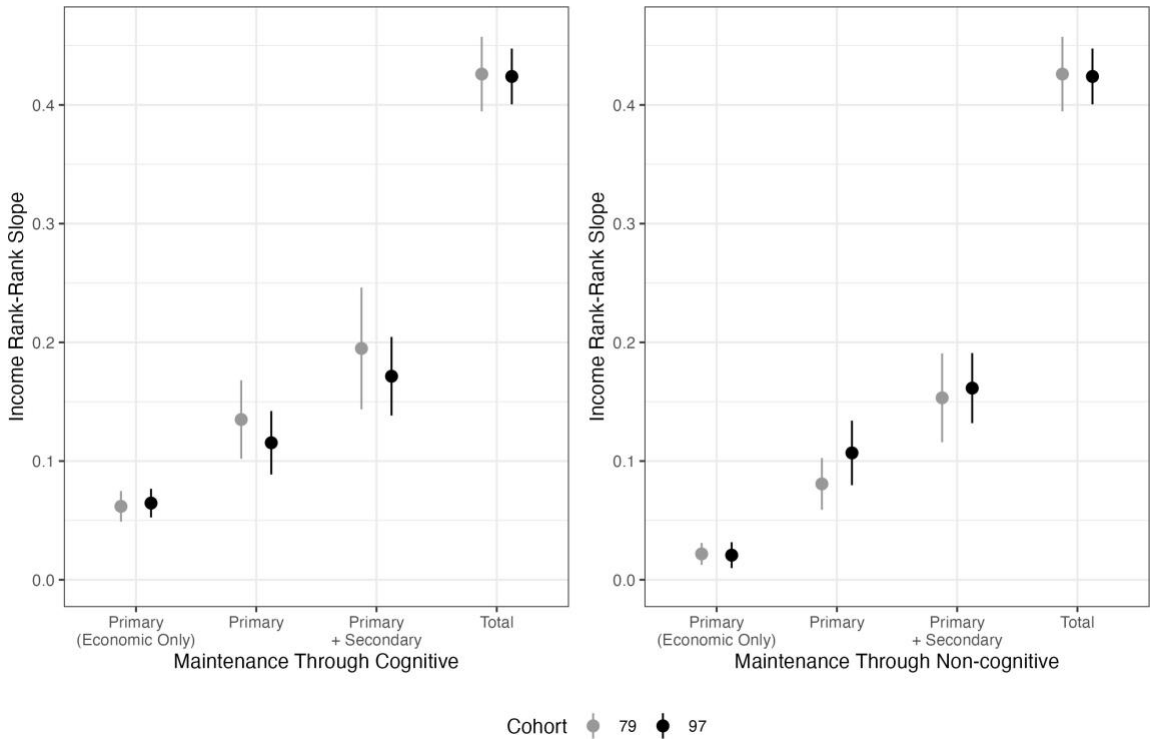


Figure 7: Intergenerational Transmission of Income through Skills in NLSY79 and NLSY97

Note: Error bars indicate 95 percent confidence interval. Method for estimation laid out in Section “Theoretical and Analytic Framework”.

	NLSY79		NLSY97	
	Parental Income	Income	Parental Income	Income
Cognitive Skill	0.41	0.43	0.37	0.35
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Skill	0.06	0.16	0.11	0.18
	(0.02)	(0.02)	(0.02)	(0.01)
Cognitive Content	0.06	0.11	0.25	0.32
	(0.01)	(0.01)	(0.01)	(0.01)
Non-Cognitive Content	0.06	0.10	0.14	0.18
	(0.02)	(0.02)	(0.02)	(0.01)
N	3,968		4,140	

Table 1: Correlations between Parental Income/Income and Skills/Occupational Skill Content

Note: Data from NLSY79 and NLSY97. All estimates are survey weighted. Standard errors in parentheses.

CHAPTER 3 THE IMPACT OF PARENTAL EDUCATION ON INTERGENERATIONAL INCOME TRANSMISSION

Abstract

How college education affects social mobility has been extensively investigated. Earlier studies of intergenerational mobility, however, predominantly concentrate on college graduates as children, but not on college graduates as parents. Neglecting to consider the family processes between college graduates and their offspring risks underestimating the overall effects of college education on mobility. The income association of college-educated parents with their children can be either higher (the complement hypothesis) or lower (the substitute hypothesis) compared with non-college-educated parents. I derive and examine the mechanisms underlying this pair of competing hypotheses, drawing from theoretical perspectives in sociology and economics. The results from PSID-CDS overwhelmingly support the substitute hypothesis. Parents who completed college display significantly lower income association with their children. The finding is robust after potential confounders are adjusted using a novel technique. This pattern is attributed to two mechanisms. First, children of college-educated parents have lower income returns to education. Second, college-educated parents have uniformly high educational expectations and supportive parenting style, irrespective of their income levels. The main potential mechanism behind the complement hypothesis, the investment efficiency theory, is also supported, but is insufficient to offset the earlier two mechanisms that produce the observed substitutivity.

INTRODUCTION

The rapid technological progress that happened in the second half of the last century dramatically improved the returns to college and professional education (Goldin and Katz 2008). Coupled with the expansion of the college student pool, the increasing return to education created a new “meritocratic class” (Markovitz 2019; Goodhart 2020; Bukodi and Goldthorpe 2021). These meritocrats, who excelled in the educational realm, work in occupations that require specific training and credentials, so the influence of their family background on their earnings and career paths is limited (Hout 1988; Torche 2011).

This phenomenon prompted a group of scholars and social observers to hail college education as "the great equalizer", an honor established by the empirical fact that, compared with non-college graduates, college graduates have lower income association with their parents. An alternative interpretation of this fact is that own education and parental income are substitutes, instead of complements, in income determination. If one is from a rich family, she does not need to be highly achieved, and if one is highly achieved, she does not need to rely on parents. Robust as this finding is, whether it is appropriate to interpret this observational pattern causally has been hotly contested (Zhou 2019; Karlson 2019; Fiel 2020). Should the descriptive difference reflect a causal effect, expanding college access could contribute to an increase in income mobility.

These previous studies provide valuable insights into the impact of college on intergenerational mobility. However, the story of college’s meritocratic force should not be considered complete unless we fully consider another related, but much less studied, empirical question, that is compared with non-college graduates, do college-graduated parents have higher or lower income association with their children? Earlier studies that explore the relation between

college and mobility mostly concentrate on college graduates as children, but not college graduates as parents.

The influence of parental education and parental income on children income may be complementary or substitutive. In other words, income association with children may be higher or lower for college-educated parents than their non-college-educated counterparts.

If parental education and parental income are complements, the claim of college as equalizer becomes dubious. While college created an equal arena for children from different backgrounds to compete in, it paradoxically exacerbated the inequality resulted from the competition, as college-educated winners are more willing and better able to secure their prize and transmit it to the next generation (Breen and Goldthorpe 1997).

If parental education and parental income are substitutes, the claim of college as equalizer becomes reinforced. Indeed, college education will be a "double equalizer", not only benefiting the individuals who receive a college education but also extending its positive effects to subsequent generations. In this scenario, both college-educated children and children of college-educated parents are expected to experience a reduced adverse impact if their families are financially disadvantaged. By obtaining a college education, individuals from lower-income families can potentially break the cycle of intergenerational income inequality. They not only improve their own prospects but also increase the likelihood of providing better opportunities and resources for their own children.

Previous literature offers conflicting predictions regarding whether parental income and parental education act as complements or substitutes in influencing children's income. Whereas theoretical perspectives from sociology imply substitutivity, those from economics suggest complementarity. I derive a series of competing hypotheses from these past theories, focusing on

two types of social processes that may produce the interactional effect between parental income and education. The first is heterogeneous educational attainment, that is how the influence of parental income on children achievement differs by parental education. The second is heterogeneous educational return, which examine the extent to which children's income returns on education differ based on their parents' educational backgrounds.

Using data from Panel Studies of Income Dynamics and its Children Development Supplement, the empirical analyses in this study aim at achieving three goals. First, I examine descriptively whether college-educated parents display higher or lower income association with their children, or in other words, whether parental education and parental income are complements or substitutes. Second, I test whether the descriptive finding is robust to selection issues. Third, I explore the potential mechanisms that underlie the observed interactional effect of parental education and income. I pay particular attention to the roles of educational expectation, parenting style, and educational expenditure in the interactive effects of parental education and income.

DOES COLLEGE IMPROVE SOCIAL MOBILITY?

The impact of college education on social mobility has been extensively discussed. Hout (1984 1988) found that the association between parental class and children class is much lower when the child completes college. This empirical pattern proves to be robust when income mobility rather than class mobility is measured, and has been corroborated in more recent studies (Chetty et al. 2017; Torche 2012). The observation of heightened social mobility among college graduates has led many to believe that higher education is a powerful equalizer. Consequently, there's a popular notion that expanding college access or increasing the proportion of college

graduates within a society would increase social mobility and foster greater equality (Treiman, 1970; Pfeffer & Hertel, 2015; Breen, 2010). Scholars have sought to understand why college graduates experience greater mobility. One hypothesis posited by earlier researchers is that college-educated individuals often secure high-skilled jobs where credentials and skills hold greater significance, thereby limiting the influence of parental socioeconomic status (Breen and Jonsson, 2007; Torche, 2011).

However, this canonical argument has recently been hotly contested. An influential article highlighted that the apparent heightened mobility among college graduates might not signify a causal equalizing effect, but rather stem from self-selection (Zhou 2019). College attendants represent a distinct group within the youth demographic, potentially diverging from non-college graduates in multiple aspects. Their enrollment in college implies their personal capabilities, which help them overcome the obstacles linked with coming from a lower socioeconomic status (SES) background. It is this unobserved ability, rather than attending college itself, that facilitates their upward mobility. Although the extent of this self-selection remains a subject of debate (Karlson, 2019; Fiel, 2020), it's widely acknowledged that the issue of selection cannot be disregarded when assessing the equalizing effect of college education. Furthermore, alternative theoretical perspectives suggest that college is not a venue for inequality reduction but for inequality generation. Children from advantaged backgrounds may benefit more from going to college. They are more accustomed to the educational system and can skillfully utilize their college experience to develop their cultural and social capital (Armstrong and Hamilton, 2013), consequently securing themselves higher-status positions in the labor market (Rivera, 2016). Children from disadvantaged backgrounds, in contrast, often lack the knowledge to effectively leverage their college degree. Limited returns to college, coupled with the prohibitively high cost

of college education in the United States, further undermines the notion of college serving as a great equalizer.

Despite the extensive discussions surrounding the impact of college on social mobility, these discussions have primarily centered on the scenario where college graduates are viewed as children rather than as parents. While we now possess significant knowledge regarding how people's college education affects their income correlation with their parents, we know considerably less about how obtaining a college education impacts their income association with their children. This gap leaves the conversation about the role of college education in mobility incomplete because it fails to consider the longer-term or cross-generation effects of college on mobility. One can envision scenarios where college initially promotes equality by boosting social mobility among graduates, but over time, when these graduates become parents, it may actually contribute to inequality if social mobility is lower among children of highly educated parents. Conversely, in the short term, college attendance may have minimal or even negative effects on mobility due to processes that reproduce inequality within college settings, but in the long run, it could enhance mobility if the children of college-educated parents experience higher levels of social mobility.

PARENTAL INCOME AND EDUCATION: OVERLAPPING, ADDITIVE OR INTERACTIVE EFFECTS?

Does going to college influences one's income association with future children? This question can be translated to whether the effects of parental income and parental education are interactive in determining children income. If parental income and parental education are complementary, meaning that the effect of parental income increases with parental education,

then attending college would lead to an increase in income association with children. Conversely, if parental income and parental education are substitutive, meaning that the effect of income decreases with education, then attending college would result in a decrease in income association with children.

To understand why the effects of parental income and education may interact, it's essential to first understand their separate effects. It's widely acknowledged that parental education, income, and social class all correlate with children's achievement. Traditionally, research on social mobility utilizes one single comprehensive measure of socioeconomic status, like occupation, education, and income, to calculate the correlation between parental and children's status measures. This approach operates under the assumption that the effects of various status measures are overlapping and inseparable. For instance, Hauser (2010) noted that "occupational socioeconomic status could be seen as roughly equivalent to permanent income in studies of intergenerational mobility". However, this assumption has faced increasing scrutiny in recent years (Jaeger, 2007; Bukodi and Goldthorpe, 2013; Mood, 2017), with a growing understanding that different parental attributes impact children through distinct social mechanisms.

The influence of parental income on children's income can be understood through two theoretical perspectives. From an economic standpoint, income transmission is emphasized through the development of children's human capital, as proposed by Becker (1962). Wealthier parents have greater financial resources, enabling them to invest more in their children's skill development, leading to higher academic achievement, subsequent success in the labor market, and ultimately higher income. In essence, parents engage in a monetized opportunity market where they purchase human capital for their children. Alternatively, a sociological perspective

underscores the significance of cultural capital in the intergenerational transmission process. Children from affluent families often benefit from access to social networks, mentorship opportunities, and high-status extracurricular activities, which not only enhance academic performance but also secure pathways to prestigious employment opportunities (Bourdieu and Passeron, 1990; DiMaggio, 1982; Dimmagio and Mohr 1985; Rivera 2016).

The impact of parental education on children income operates through mechanisms that partly overlaps but are also distinctive from those of parental income in significant ways. Parental education's effects can be roughly categorized into two mechanisms: parenting style and expectations. First, from the perspective of cultural capital, parents with college education often adopt a parenting style and foster a family environment characterized by "concerted cultivation" (Lareau, 2011). This style emphasizes the development of non-cognitive skills such as leadership, coordination, and cultural and personality traits associated with the upper class, traits that the school system tends to value and reward (Bourdieu, 1977; DiMaggio, 1982; Lareau, 2002). These signals of elite status not only contribute to educational achievement but also exert a direct influence on labor market outcomes, particularly in the attainment of high-end jobs (Rivera, 2016). Second, parental education influences mechanisms such as motivation, expectations, and aspirations. Out of fear of status loss, parents with higher levels of education typically hold high educational expectations for their children, irrespective of their children's circumstances (Breen and Goldthorpe 1997; Breen and Yaish, 2006). Parental prioritization of children education nurtures their children's ambitions for accomplishment. This dynamic between parental expectations and children's aspirations serves as an important channel driving academic and career success. Dating back to the status attainment model, research by Rosen (1956), Sewell et al. (1970), and Sewell and Hauser (1975) underscore the pivotal roles of

expectation and aspiration in shaping children's trajectories. More recently, Davis-Kean (2005) illuminates how these psychological factors play a crucial role in determining children's outcomes, emphasizing the critical mediating role of parental expectations on intergenerational status transmission.

Though the distinct mechanisms through which parental education and income impact children's success have been documented in great detail, our understanding of whether these effects interact remains limited. Interactions between parental class and income have been discussed, driven by the some-time disparate findings from class mobility and income mobility. Blanden et al. (2013) and Breen et al. (2016) have proposed theoretical frameworks and decomposition methods for jointly analyzing the effects of parental class and income. Similarly, using administrative data, Mood (2017) analyzed how the income effect differs by parental class. In terms of the effects of parental income and parental education, numerous studies have examined the effects of them simultaneously and contrasted them with each other (Haveman and Wolff, 1995; Davis-Kean, 2005; Bukodi and Goldthorpe, 2013; Chevalier et al., 2013). However, these studies have provided scant information about the potential interaction between income and education. The research most closely aligned with our paper is that of Connolly et al. (2022), which utilizes Canadian data to explore how parental education moderates intergenerational income mobility. However, this study did not delve into the potential mechanisms underlying this moderation.

It is becoming increasingly urgent for us to understand the interaction between parental income and education, as the two parental features are becoming increasingly aligned (Reardon 2011). It is well-known that the process of skill-biased technological change has increased the payoff to college, hence strengthening the association between income and college education

(Goldin and Katz 2009). Little attention has been paid to the intergenerational consequence of this transformation. The increasing convergence between education and income means that more resources are being concentrated within “doubly advantaged” households. Compared with previous times, rich parents are more likely to be highly educated, and highly educated parents are more likely to be rich. If the effects of parental education and income are substitutes, then status crystallization has a limited impact on exacerbating inequality, because having money makes education less relevant, and having a college degree makes money less relevant. Conversely, if parental education and income have additive or complementary effects, it suggests that skill-biased technological change has the unintended consequence of worsening intergenerational inequality. This is because children from high-income, high-education families may benefit from a "double advantage," or even a "reinforced advantage," as the effects of parental income and education amplify each other.

RESEARCH HYPOTHESES: COMPLEMENTARITY OR SUBSTITUTIVITY?

The preceding sections have discussed the motivations for investigating the potential interactive effects of parental education and income on children income. This section outlines the theoretical framework underpinning this research. Building upon this framework, I formulate competing hypotheses regarding the potential interactive effects of parental education and income and the mechanisms underlying this interaction.

Figure 1 illustrates the theoretical framework of a two-generation status transmission. Following the conventions of social stratification research, children's education, particularly college attainment, is incorporated as a crucial mediator in intergenerational income transmission (Blau and Duncan, 1967; Breen, 2010). The causal chain unfolds as follows: parental college

attainment influences parental income, which in turn impacts children's college attainment, ultimately influencing their income. The upper panel of Figure 1 depicts the core hypotheses of this article, with the dashed line symbolizing moderation. The primary competing hypotheses concerning the interaction between parental education and income are expressed as follows:

H1a Substitutivity: For parents who completed higher-education, their income association with children is lower compared with non-college-graduated parents.

H2a Complementarity: For parents who completed higher-education, their income association with children is higher compared with non-college-graduated parents.

[Figure 1 about here.]

Why might there be an interaction between parental education effect and parental income effect? From the perspective of the canonical origin-education-destination (OED) framework, income transmission can be separated into two processes, college attainment, that is the effect of parental income on children college completion, and college return, that is the effect of children college completion on children income. We formulate the hypothetical mechanisms through which parental education moderates income mobility by examining how it moderates college attainment and college return. The subsidiary hypotheses that support substitutivity will be labelled as H1, and the hypotheses that support complementarity will be labelled as H2.

First, regarding how parental education moderates college return, children from various parental backgrounds may experience differing economic returns from college education. The question of which children benefit most from college is a subject of long debate. From a rational-

choice perspective, assuming perfect information and certainty, those most likely to pursue college are those who stand to gain the most from it (Willis and Rosen, 1979; Heckman et al., 2006). Given that children from advantaged backgrounds are disproportionately represented in college, the economics perspective suggests that children whose parents have higher levels of education enjoy greater benefits from college. This implies that parental education and children's education are complementary in determining children's income, supporting H2a, the complementarity hypothesis. Conversely, an alternative theoretical perspective posits that it is the most disadvantaged children who benefits the most from college. This is because those from disadvantaged backgrounds face significant adversity in the labor market without a college degree. Whereas children from affluent families are protected by the family safe net which can mitigate the impact of educational setbacks (Brand and Xie, 2010). This perspective suggests that parental education and children's education are substitutes, providing support for H1a. These competing hypotheses are expressed as H1b and H2b.

H1b: For children with parents who completed higher education, their return to education is lower than those with non-college-educated parents.

H2b: For children with parents who completed higher education, their return to education is higher than those with non-college-educated parents[12].

Next, we turn to how parental education may moderate college attainment. The process of differential college attainment can be further separated two different mechanisms. As previously discussed, parental education influences children's achievement through two crucial channels:

[12] The relation between descriptive difference and causal effect is discussed later.

parenting style and educational expectations. Given that parental education positively influences these two mediators, it follows a similar logic that parental education may also weaken the link between parental income and these mediators. While income may contribute to fostering a conducive home environment for children's academic development, it's conceivable that these influences can be substituted by parental educational attainment. Similarly, parental college attainment may also attenuate the link between parental income and educational expectations. When parents have attained a college education, it's culturally anticipated that their children will also pursue higher education, regardless of the family's financial resources. These mechanisms imply that parental education and parental income are substitutes in determining parenting style and educational expectation for children. Therefore, this hypothesis is in support of the substitutivity hypothesis, H1a. These hypotheses are expressed as H1c.

H1c: For parents who completed higher education, the association between their income and parenting style, and the association between their income and educational expectation, are lower.

Parental education may moderate children's college attainment through an alternative mechanism. Recent literature suggests that parental human capital and parental income may act as complements in developing children's human capital, and consequently, their income (Becker et al., 2018; Durlauf and Seshadri, 2018). The rationale behind this is that the effectiveness of parental investment increases with parental human capital. In other words, given a fixed budget, highly educated parents are better equipped to make informed decisions about how to allocate resources towards their children's education. This implies that parental education and educational

investment in children are complementary factors in determining children's education: for parents with a college degree, the relationship between educational investment and children's achievement is stronger. If this hypothesis holds true, it would support H2a, the complementarity hypothesis. This hypothesis is expressed as follows.

H2c: For parents who completed higher education, the association between their educational expenditure and children's college attainment is higher.

To summarize, the core competing hypotheses of this research revolve around whether the effects of parental education and parental income are complementary (H2a) or substitutive (H1a). Another way to express these hypotheses is whether parental education positively or negatively moderates income mobility. To explore the potential mechanisms underlying this moderation, we focus on how parental education may moderate college return and college attainment. Regarding college return, if children of college-educated parents benefit more from college (H2b), then H2a is supported; conversely, if children of college-educated parents benefit less from college (H1b), then H1a is supported. In terms of college attainment, H1c suggests that parental education decreases the association between parental income and parenting style/college expectations, thereby supporting H1a. On the other hand, H2c posits that parental education increases the association between educational investment and children's college attainment, thus supporting H2a.

DATA

Here, this study makes use of the Panel Study of Income Dynamics (PSID), along with its Child Development Supplement (CDS). The PSID is a nationally representative survey that has been tracking over 5,000 families in the United States since 1968. Beginning in 1997, an additional data collection effort started that focused on children aged 0-12 within the sample, known as CDS-I. Subsequent follow-up interviews were conducted in 2002 and 2007 with the same group of children, referred to as CDS-II and CDS-III, respectively. This research leverages data from both the PSID main survey and CDS I-III. The unique strengths of PSID-CDS dataset are manifold. First, it offers a wealth of child-related information, encompassing educational expectations, expenditures, and parenting styles. Second, by administering questionnaires to all family members, including parents and children, the dataset provides a comprehensive array of parental background characteristics. Controlling for these parental features is essential for the causal analysis conducted in this study. Third, PSID features a rare multigenerational structure, offering detailed information spanning across multiple generations. This enables the undertaking of a straightforward three-generational descriptive analysis, which assesses how college education influences both the income association with parents and the income association with children.

The key independent and dependent variables are as follows. Both parental and children income are the average family income between ages 25-35, inflation-adjusted to 2020 US dollars. In analyses they are both operationalized as percentile ranks. We use father's college status as a measure of parental education. Father college and college refer to an approximation of college completion, that is 16 years of schooling or above. College expectation is the answer to

the question "How much schooling do you expect that child will complete?", with "graduate from 4-year college" as 1 and others as 0. Expenditure is the sum of three child-specific annual expenditure items, tuition, tutoring, and school supplies, consistent with the educational expenditure measure use by earlier scholars (Hao and Yeung 2015). Parenting style is measured by the age-standardized Home Observation for Measurement of the Environment-Short Form (HOME-SF) scale. HOME-SF is a measure of family environment that incorporate more than 50 items from primary-giver report and interviewer observation[13]. HOME-SF can be roughly split into two categories: cognitive stimulation and emotional support. Originally proposed in Caldwell and Bradley (1984), HOME-SF is a very popular measure of parenting style and home environment used by sociologists (for a review of studies using HOME-SF, see Mott 2004). Expectation is asked in CDS I-III, expenditure in CDS I-II, and parenting style in CDS I. When the variables are measured in multiple waves, I take the average.

For the causal analysis, a series of confounders that predate parents' educational attainments are also used in the analyses. The confounders adjusted include race, youth health conditions, youth psychological condition, region of residence, type of residence, family economic status during childhood, cohabitation with parents. Missing values in control variables were imputed using the R package "mice". Specifically, a missing covariate was imputed using a model with all other control variables as predictors. Continuous, binary, and categorical variables were imputed using, respectively, predictive mean matching, logistic regression imputation, and polytomous regression imputation.

Two samples are used in the analyses. The first is a three-generation sample with known educational and income for all three generations. This sample serves a simple data description

[13] See https://www.nlsinfo.org/content/cohorts/nlsy79-children/other-documentation/codebook-supplement/appendix-home-sf-scales for a summary of items included in HOME-SF.

that shows whether the income associations between the middle generation and the older generation, and between the middle generation and the younger generation, differ by the education of the middle generation. This analysis provides an intuitive picture for the possible cross-generation equalizing spillover of education. The linking is completed across generations using the PSID Family Identification Mapping System 3-generation prospective linkage file. There are a total of 2,730 three-generation families with available income and education information in the dataset.

The overall PSID sample, however, does not have the rich child-relevant information we need for understanding the mechanisms. Thus, for the main analyses, we restrict our sample to two-generation data from PSID-CDS. As mentioned above, these include children aged 0-12 in 1997 and their parents. After removing missing values in parental income, children income, and father education, there remain 2,360 father-child pairs in the dataset. After further removing missing values in parenting style, educational expectation, and educational expenditure, 981 father-child pairs remain for analysis. Due to the relatively limited sample size, pair-wise deletion is employed in our analyses, except for control variables where missing values being imputed. Descriptive statistics of these variables are reported in Table 1.

[Table 1 about here.]

DESCRIPTIVE RESULTS

As an intuitive demonstration of the potential cross-generation equalizing effect of college education, I leverage the multigenerational nature of the PSID dataset and conduct a straightforward three-generation descriptive analysis. The results of this analysis are presented in

Table 2. Consistent with numerous earlier findings, individuals who complete college exhibit lower income associations with their parents. Furthermore, when this middle-generation cohort becomes parents themselves, another notable pattern emerges: their income associations with their own children are also lower. This observation suggests a higher rate of mobility among children with college-educated parents. Taken together, these findings indicate that parental education and income are substitutes, aligning with the hypothesis H1a.

[Table 2 about here.]

The three-generation data unfortunately lacks sufficient child information as well as background information of parents, which are essential for our causal and mechanism analyses. In the main analyses we stick to the PSID-CDS sample. Table 3 reports results that provide tests for Hypotheses 1a, 1b, 2a, and 2b using the main sample. Column 1 shows that compared with non-college-educated parents, the income association of college-educated parents with their children is about 32 percent lower (0.15/0.47). H1a is supported and college education and income seem to be strong substitutes in determining children income. The second column shows heterogeneous returns to college education by father education. This shows that, descriptively, the income returns to college for those whose parents go to college is only two thirds that of those whose parents have not finished college. This supports H1b, and heterogeneous college return could be a mechanism that explains substitutivity.

[Table 3 about here.]

Table 4 presents the tests for Hypotheses 1c and 2c, which investigate the roles of expectation, parenting style, and expenditure in the interactive effect of parental education and income on children's income. The first two columns indicate that paternal college attendance weakens the association between income and both expectation and supportive and stimulative parenting. Specifically, among college-educated fathers, the association between income and expecting children to complete college is reduced by over 50 percent (0.15/0.29). For fathers without a college degree, a one percentile increase in income correlates with approximately a 3 percentage point rise in the expectation of their children completing college. This value decreases to about 1.5 percentage points among fathers who have completed college. A similar pattern is observed in the association between parenting style and parental income. In families where fathers have attained a college education, the relationship between parental income and parenting style is approximately 60 percent (0.83/1.4) weaker compared to families with fathers who have not completed college. H1c is therefore supported. The substitutive effects of education and income on educational expectation and parenting style are strong explanations for the substitutivity of education and income in determining children income.

The rightmost column of Table 4 examines Hypothesis 2c, the investment efficiency hypothesis. The results suggest that educational expenditure is more effective when the father is college-educated. Among fathers without a college degree, a one percent increase in educational expenditure is associated with a 2 percentage point increase in the probability of children finishing college. When the father is college-educated, this effect substantially increases to approximately 6 percentage points, nearly tripled. This finding provide support for Hypothesis 4, indicating that parental education and expenditure on children are indeed complementary factors influencing children's achievement. However, despite the support for H2c, the overall trend leans

towards substitutivity rather than complementarity due to more pronounced mechanisms in the opposite direction.

[Table 4 about here.]

ADDRESSING SELECTION: A NEW ESTIMAND

Apart from simple descriptive analyses of associations of key variables, this study identifies and estimates a novel estimand that adjusts for confounding in the observational effect moderation between parental education and income. Such a procedure to address selection issues is necessary as a descriptive pattern does not necessarily warrant causal interpretation. Given that college graduated parents are a selected group of the population, they may be distinct in ways that would impact income transmission. For instance, Asian Americans, who are overrepresented in college, also happen to exhibit higher intergenerational mobility (Liu and Xie 2016). Therefore, the fact that college graduates have lower income association with parents could simply be that Asian Americans constitute a larger proportion of college graduates than non-college graduates.

Two existing causal approaches are possible to evaluate the interactive effect of parental income and education on children income in this setup. First, we can estimate the conditional average treatment effect (CATE) of parental income on children income by parental education group. Second, we can estimate the controlled direct effect (CDE) of parental education on children income, controlling for parental income. Though both estimands reflect how the effects of parental education and parental income depend on each other, they do not directly target our research goal, that is the causal effect of parental education on parent-child income association.

In other words, we are interested in the treatment effect on the association of two outcome, henceforth referred to as TEA.

Let D denotes parental college completion, X parental income, and Y children income. Let Z denotes the confounders behind the interaction of parental education and parental income. Let $X(1)$ and $Y(1)$ denotes the potential parental income and children income if parents have completed college, that is $D = 1$. Similarly, $X(0)$ and $Y(0)$ are the potential parental and children income if parents have not completed college. Notice that these potential outcomes are unobserved counterfactuals and can only be estimated from data with identifying assumptions.

First, we define the observed income mobility. Let β denotes observed income persistence (or immobility), $\beta_{|1}$ observed income persistence among college-graduated parents, and $\beta_{|0}$ observed income persistence among non-college-graduated parents.

$$\beta = \frac{cov(X,Y)}{var(X)}, \quad \beta_{|1} = \frac{cov(X,Y|D=1)}{var(X|D=1)}, \quad \beta_{|0} = \frac{cov(X,Y|D=0)}{var(X|D=0)}$$

To evaluate whether the results are robust to confounding, we focus on the following counterfactual estimand:

$$\beta_{*1} = \frac{cov\big(X(1),Y(1)\big)}{var\big(X(1)\big)}$$

This estimand is the counterfactual income mobility we would observe if all parents completed college. Similarly, we can define the counterfactual where no parent completed college, that is:

$$\beta_{*0} = \frac{cov\big(X(0), Y(0)\big)}{var\big(X(0)\big)}$$

The difference between the two represents a causal effect of D on the association between X and Y:

$$TEA = \beta_{*1} - \beta_{*0}$$

TEA, however, evaluates the difference in results between two hypothetical interventions, one setting all people, and the other setting zero people, to receive the treatment. Given that those are extreme scenarios, TEA may not be the most theoretically- or policy-relevant quantities in a lot of research settings. Some other quantities may be more informative given our research purpose. First, we can define treatment effect among the treated and the untreated. Let us define $\beta_{*1|0} = \frac{cov(X(1),Y(1)|D=0)}{var(X(1)|D=0)}$ and $\beta_{*0|1} = \frac{cov(X(0),Y(0)|D=1)}{var(X(0)|D=1))}$, which are the counterfactual associations if those $D = 1$ or $D = 0$ receive the opposite treatment. For instance, $\beta_{*1|0}$ is the income mobility among parents with a college degree had they not completed college. Obviously, $\beta_{*1|1} = \beta_{|1}$, and $\beta_{*0|0} = \beta_{|0}$. Then, treatment effect among the treated and the untreated are defined as follows:

$$TEAT = \beta_{*1|1} - \beta_{*0|1} = \beta_{|1} - \beta_{*0|1}$$

$$TEAUT = \beta_{*1|0} - \beta_{*0|0} = \beta_{*1|0} - \beta_{|0}$$

An additional quantity that may be of interest is how the intervention of giving everyone treatment would change income mobility compared with the status quo. I refer to this as realizable TEA, or RTEA.

$$RTEA = \beta_{*1} - \beta$$

Since β_*s are unobserved values, we can only identify them with assumptions. Following conventions in causal inference using observational data, we impose three identification assumptions: ignorability, SUTVA, and positivity.

Assumption 1 *Ignorability:* $X(d), Y(d) \perp D|Z$ *for* $d = 1$ *and* $d = 0$.

Assumption 2 *SUTVA:* $X(d) = X$ *and* $Y(d) = Y$ *for respondents whose* $D = d$.

Assumption 3 *Positivity:* $0 < p(D = d|Z) < 1$ *for* $d = 1$ *and* $d = 0$ *and the support of* Z.

Here, Assumption 1 states that there is no unobserved confounder between the causal relation of D and X, and the causal relation between D and Y. That is, Z incorporated all common causes of parental education and parental income, and parental income and children income. This assumption would be violated in the presence of unobserved confounders. Notice the critical

difference between TEA and causal mediation is that we do not assume sequential ignorability. That is, we are completely agnostic to the causal structure between X and Y. Assumption 2 is also oftentimes understood as system invariance or no spillover. It states that the counterfactual outcome given a treatment value equals the observed outcome among those with this treatment value. This assumption would be violated if the hypothetical intervention, i.e., all parents or no parent completed college, fundamentally changed the effect of education on income such that we cannot infer post-intervention income based on observed incomes of the population. Assumption 3 states that for all respondents and all possible treatment values, the probability of receiving this treatment values are strictly non-zero. It means that no population subgroup is bound to complete or not complete college.

Causal interventional effects normally have two interpretations, a global interpretation, and a local interpretation (Lundberg 2022). In our case, the global interpretation of TEA would be the difference in income persistence between two dramatic hypothetical interventions, one sending all parents to college and the other removing all parents from college. In this article we stick to the local interpretation, which is a lot less ambitious. The local interpretation of TEA is the difference in income persistence between two hypothetical interventions on a small random sample of the population of parents, one sending this sample to college, and the other removing all sample from college.

The choice of the local interpretation has two advantages. First, the hypothetical sample is small in contrast to the full population, so it is safe to assume that this mild intervention would not violate system invariance. Second, when using income percentile ranks as measures of income, the limited size of the sample ensures that any alterations in ranks within it have a negligible effect on the overall population rankings.

With these assumptions, $\beta^*_{|1}$, $\beta^*_{|0}$ and TEA can be non-parametrically identified. In this section we focus on the identification and estimation of counterfactual covariances. The counterfactual variances, $var\big(X(d)\big)$ and $var\big(Y(d)\big)$, can be identified and estimated using similar techniques to that on covariance. The details are skipped, and similar methods can be found in Chapter 1 of the dissertation.

Take $cov\big(X(1), Y(1)\big)$ for example, its identification result is as follows (proof in Appendix):

$$cov\big(X(1), Y(1)\big) = \int cov(X, Y|D = 1, z)p(z)dZ + cov(E[X|D = 1, Z], E[Y|D = 1, Z])$$

By contrasting it with the observed mobility among college parents, we can have a more intuitive understanding about the source of confounding:

$$cov(X, Y|D = 1) = \int cov(X, Y|D = 1, z)p(z|D = 1)dZ + cov(E[X|D = 1, Z], E[Y|D = 1, Z]|D = 1)$$

We focus on the difference in the first term. The difference between the two quantities is how the conditional income covariance, $cov(X, Y|D = 1, z)$, is aggregated based on background characteristics Z. While observed mobility among college-graduated parents, naturally, reweighs the conditional covariance with the characteristics among college graduates, $p(z|D = 1)$, the counterfactual quantity reweighs with the background distribution of the general population $p(z)$. This observation provides an intuitive understanding about why $\big(\beta^*_{|1} - \beta_{|1}\big)$ reflects

confounding or self-selection: the difference between the two results from the difference between the background characteristics between college graduates and the general population.

Next, we discuss how to estimate the quantity, still taking $cov\big(X(1), Y(1)\big)$ as an example. Notice from the identification result that $cov\big(X(1), Y(1)\big)$ contains two terms, and I will discuss the estimation of them separately. For the first term, given the definition of covariance, it can be expressed as $\int cov(X, Y|D = 1, z)p(z)dZ = \int (E[XY|D = 1, z] - E[X|P = 1, z]E[Y|P = 1, z])p(z)dZ$. This expression suggests a regression-based imputation estimator and a weighting estimator.

We first discuss the regression imputation estimator. First, we estimate three conditional expectation functions, using OLS or flexible machine learning estimators. For the three functions, the outcomes are $X * Y$, X, and Y respectively, and the predictors are D and the covariates Z. Second, we set parental education, D, of all respondents to be 1, and get the predicted values of XY, X, and Y using the three functions estimated in the first step. Denote the predicted values for individual i as $\widehat{\tau_i}(XY) = \hat{E}[XY|D = 1, z_i]$, $\widehat{\tau_i}(X) = \hat{E}[X|D = 1, z_i]$, and $\widehat{\tau_i}(Y) = \hat{E}[Y|D = 1, z_i]$ respectively. Third, an estimator for the first term of $cov\big(X(1), Y(1)\big)$ is calculated as $\sum_{i=1}^{N} \frac{\widehat{\tau_i}(FY) - \widehat{\tau_i}(F)\widehat{\tau_i}(Y)}{N}$.

Now we look at the second term in the identification result of $cov\big(X(1), Y(1)\big)$, that is $cov(E[X|D = 1, Z], E[Y|D = 1, Z])$. Notice that this can be conveniently estimated with the sample covariance of two conditional expectations $cov(\widehat{\tau_i}(X), \widehat{\tau_i}(Y))$. Therefore, we can write the full expression of the regression-imputation estimator of the counterfactual covariance as:

$$\widehat{cov_{RI}}\big(X(1), Y(1)\big) = \frac{1}{n}\sum_{i=1}^{n} \widehat{\tau_i}(XY) - \widehat{\tau_i}(X)\widehat{\tau_i}(Y) + cov(\widehat{\tau_i}(X), \widehat{\tau_i}(Y))$$

A weighting estimator can also be calculated from the identification result of $cov\big(F(1), Y(1)\big)$. Still, start from the first term of the result $\int cov(X, Y|D = 1, z)p(z)dZ = \int (E[XY|D = 1, z] - E[X|D = 1, z]E[Y|D = 1, z])p(z)dZ$. Note that, for the first term of the identification result, if the conditional functions can be aggregated by $p(z|P = 1)$ rather than $p(z)$, we can estimate it with simple sample averages among those $P = 1$. Therefore, we reweigh the samples whose $P = 1$ with $\frac{p(z)}{p(z|D = 1)}$. After applying the Bayes rule, $\frac{p(z)}{p(z|D = 1)} = \frac{p(D=1)}{p(D = 1|z)}$. This result implies the following weighting estimator. First, we run a propensity score model, using logistic regressions or other estimators used for binary outcomes, with D as the outcome and Z the predictors. Second, for all respondents whose $D = 1$, we calculate the weight $\frac{\hat{p}(D=1)}{\hat{p}(D = 1|z_i)}$ where $\hat{p}(D = 1)$ is simply the proportion of parents who complete college in the sample, and $\hat{p}(D = 1|z_i)$ is the propensity score of completing college for parent i, predicted using the model in the first step. We estimate the second term of the counterfactual covariance using the same way as in the regression-imputation estimator. Therefore, the weighting estimator for the counterfactual covariance can be expressed as:

$$\widehat{cov_W}\big(F(1), Y(1)\big) = \frac{1}{n}\sum_{i=1}^{n}\frac{X_i Y_i I(D_i=1)}{\hat{p}(D = 1|z_i)} - \frac{1}{n}\sum_{i=1}^{n}\frac{X_i I(D_i=1)}{\hat{p}(D = 1|z_i)}\frac{1}{n}\sum_{i=1}^{n}\frac{Y_i I(D_i=1)}{\hat{p}(D = 1|z_i)} + cov(\hat{\tau}_i(X), \hat{\tau}_i(Y))$$

In the results section, I report estimates from regression-imputation estimators with OLS as the conditional expectation function. All standard errors are bootstrapped.

The approach outlined here is related to, but also distinct from, several earlier literatures on causal inference. First, within the parametric framework, some previous studies have developed covariance regression models to examine how an independent variable influences the covariance of two outcomes (Hoff and Niu, 2012; Bloome and Schrage, 2021). In contrast, this article offers non-parametric identification results and more generalized estimation procedures that are not confined to a specific type of parametric model. Second, our proposed treatment effect on the association (TEA) concept shares similarities with the gap-closing estimand (GCE), which focuses on how a causal intervention on one variable affects the association between two other variables (VanderWeele and Robinson 2014; Jackson and VanderWeele 2018; Lundberg, 2022). The difference between the two is that while TEA focuses on the treatment effect on the association between two outcomes, GCE studies the treatment effect on the association between one outcome and one pre-treatment covariate. Third, our statistical objective shares common ground with that of causal mediation or, more specifically, controlled direct effect (Pearl, 2001; Robins, 2003). While causal mediation analyzes the relationships between a treatment, a mediator, and an outcome, the analyses in this article examine a treatment variable and two outcomes. Importantly, we remain agnostic to the causal relation between parental and children's income. Instead, our focus is on how parental education causally impacts the association between the two incomes.

CAUSAL ANALYSES

Next, I present results of the causal analysis. In the earlier descriptive analyses, I show evidence supporting H1a, the substitutivity hypothesis. It is possible that it is not college education per se that reduces intergenerational income association, but some other attributes of

college-educated parents that are also linked to their college completion. Here, we test whether the finding of substitutivity is robust to confounders.

Table 5 reports the observed and counterfactual income associations by father educational background, as well as causal quantities of interest. Descriptive, income mobility is 32 percent ((0.47-0.32)/0.47) higher for children with college-educated father compared with those with non-college-educated father. The causal quantities are reported in the lower panel of Table 5. It is estimated that, for fathers who completed college, their income association with children would rise from 0.32 to 0.44 had they not completed college. In other words, TEAT equals -0.13 (0.32-0.44 with rounding). Similarly, TEAUT is -0.03. This means that, for the fathers who did not complete college, had they graduated from college, their children would have higher income mobility, as the income association drops from 0.43 to 0.47. Therefore, though the mobility-enhancing effect is not as large as among college-graduated fathers, they can also potentially benefit from completing college. These results mean that completing college indeed causes income association with children to decrease, both for those who completed college and for those who did not. RTEA is the most policy-relevant measure as it expresses the expected mobility increase if all parents completed college. The result shows that observed income association in the society would be expected to decrease by 0.12, approximately 24 percent lower, if all parents completed college.

[Table 5 about here.]

Figure 2 provides a visual representation of the mobility disparity between parents who completed college and the general population. It illustrates that income mobility is higher among

fathers who completed college because the negative impact of low income on children is mitigated for college-educated fathers. For instance, in the general population, if parents belong to the lowest income quintile, the expected income percentile for their children is approximately 0.3. However, for children of college-educated parents from the lowest income quintile, the expected income percentile is closer to 0.4. Economic upward mobility is considerably more common for children of college-educated but financially disadvantaged parents. Therefore, the observed high mobility among children of college-educated parents does not stem from children of high-income parents being more likely to move downward; instead, it arises from children of low-income parents being more likely to move upward.

Comparing the income mobility among college-educated parents with the counterfactual mobility if all parents completed college, we observe a general alignment between the two patterns. This suggests that the role of confounding is limited, indicating that the observed mobility trends among college-educated parents are likely attributable to college completion. The only difference is for parents in the lowest-income quintile: the expected income of children of parents from the lowest income quintile does not change significantly, even if all parents were to attend college. This implies that expanding college access may not offer substantial benefits for children of the most disadvantaged parents, although it could positively impact most other children in relatively disadvantaged positions. The high upward mobility observed among children of college-educated parents from the lowest income quintile may be influenced by confounding factors associated with these parents, rather than solely due to the college completion of parents.

[Figure 2 about here.]

In the descriptive analyses, we find evidence in support of H1c. Similar to the case of H1a, these observations should not be interpreted causally without further investigation. It is plausible that the reduction in the association between income and the expectation of children completing college may not solely be attributed to attending college itself. Other attributes associated with parental college completion, such as grandparental SES, could also weaken the association between income and expectation on children. We can also conduct the same causal exercise we did to H1a on how parental college and income interact to affect educational expectation and parenting style (H1c), and how parental college completion and educational expenditure interacts to affect children's college completion (H2c).

Table 6 shows the treatment effects of father completion on covariances between parental income and expectation, parental income and parenting style, and expenditure and college completion. The results after adjusting for confounding generally agree with the observational pattern as reported in Table 4. For instance, consider the association between educational expectation and parental income. Descriptively, this association is 0.016 lower for fathers with a college degree. Causally, estimates suggest that without college completion, the association between income and educational expectation would be 0.022 higher. If all parents completed college, we'd expect the income-expectation association to be 0.011 lower.

Similarly, the causal interaction between college completion and income on parenting style, in particularly TEAT, also aligns with observational patterns. However, the interaction of college and expenditure on children's college completion becomes statistically insignificant after adjusting for confounders. These results suggest that H1c is robust to the influence of confounders and that the observed interaction is primarily attributed to the moderation effect of

college completion. H2c, on the other hand, while descriptively supported, appears to be sensitive to selection issues.

[Table 6 about here.]

CONCLUSION

Most earlier studies on the relation between college education and social mobility focuses on college graduates as children. The cross-generation effect of college on mobility, or how going to college affects mobility between college graduates and their children has not received sufficient attention. In this research, I tested whether parental college education and parental income are complements or substitutes in determining children income, addressed the selection issue in this analysis, and explored the mechanisms for the education-income interaction.

The results from PSID-CDS overwhelmingly support the substitute hypothesis. Parents who completed college display significantly lower income association with their children. The finding is robust after potential confounders are adjusted using a novel technique. This pattern is attributed to two mechanisms. First, children of college-educated parents have lower income returns to education. Second, college-educated parents have uniformly high educational expectations and also uniformly supportive and stimulative parenting style, irrespective of their income levels. The main potential mechanism behind the complement hypothesis, the investment efficiency theory, is also supported, but is insufficient to offset the earlier two mechanisms that produce the observe substitutivity.

It has been widely recognized that different family background measures, like income, education, or class, have distinct influences on children through different social mechanisms.

Few research, however, has investigated how family backgrounds interact to affect children. In an era when parental SES measures are becoming more aligned (Reardon 2011), it is ever-more pressing to understand the intergenerational consequence of status crystallization. If parental backgrounds are complements, then the increasing alignment between income and education would no doubt exacerbate intergenerational inequality, with resources more concentrated among families with advantages across different forms of capital. This study, however, discovers that parental income and education are substitutes rather than complements, which significantly ameliorates the threat from status crystallization on intergenerational equality. The substitutivity of SES may explain why intergenerational mobility in the US has remained stable over time despite increasing levels of status crystallization.

The findings presented in this article prompt us to reconsider the roles of well-known mediums of intergenerational status maintenance, such as cultural capital, parental expectations, and parenting style. Traditionally, these factors have been depicted as channels that perpetuate inequality, serving as pathways through which affluent parents pass on their wealth and advantages to their offspring. However, this research highlights that these mechanisms are equally crucial for well-educated parents to offset their economic disadvantage. Through these channels, children of well-educated but economically disadvantaged parents demonstrate equal capability in the realms of education and the labor market, comparable to children from affluent backgrounds.

Methodologically, this research makes a contribution to social stratification studies by offering a clear definition, identification results, and estimators for an estimand that specifically targets the treatment effect on the association between outcomes. This methodological approach holds potential for broader applications beyond the current setup. Frequently, social stratification

scholars are interested not in the effect of a treatment on an outcome but in its effect on inequality, often measured by the association between outcomes. While causal mediation analysis can provide insights into such substantive questions, it typically addresses them indirectly. By comparison, the approach introduced in this research directly targets the effect on inequality.

In recent years, the equalizing role of college education has come under serious scrutiny (Zhou 2019). It is argued that, after accounting for selection, going to college does not reduce income association with parents. This challenges the conventional policy perspective that expanding access to college would naturally promote equality of opportunity. Our article provides a response to that challenge by directing attention to a previously unexplored aspect: the cross-generational equalizing spillover effect of college education. The findings suggest that educational policies aimed at boosting the educational attainment of today's youth may not only influence intergenerational mobility as they enter the labor market in the coming years, but also contribute to enhancing equality of opportunity in society when they become parents in subsequent decades. Therefore, while college expansion may not yield immediate equalizing effects, it holds promise for fostering long-term equality.

From the perspective of parents, this research highlights the "buffer" effect of attending college on the well-being of their children. Equipped with both cognitive and non-cognitive skills acquired during their college experience, parents are better positioned to provide a supportive developmental environment for their children, even in the face of financial challenges. This buffering benefit holds particular significance for parents who themselves come from disadvantaged backgrounds. Even if these parents encounter difficulties in the labor market

due to their own disadvantages, their children are still able to fare nearly as well as the children of their more successful peers, thus breaking the intergenerational cycle of poverty.

APPENDICES

A. PROOF OF THE IDENTIFICATION RESULT OF TEA

According to the law of total covariance:

$$cov\big(X(1), Y(1)\big) = E[cov(X(1), Y(1)|Z)] + cov(E[X(1)|Z], E[Y(1)|Z]) \coloneqq A + B$$

Where:

$$\begin{aligned} A &= \int cov(X(1), Y(1)|z)p(z)dz \\ &= \int cov(X(1), Y(1)|z, D = 1)p(z)dz \quad \leftarrow Ignorability \\ &= \int cov(X, Y|z, D = 1)p(z)dZ \quad \leftarrow SUTVA \\ &= \int (E[XY|z, D = 1] - E[X|z, D = 1]E[Y|z, D = 1])p(z)dZ \end{aligned}$$

And:

$$B = cov(E[X|D = 1, Z], E[Y|D = 1, Z])$$

FIGURES AND TABLES

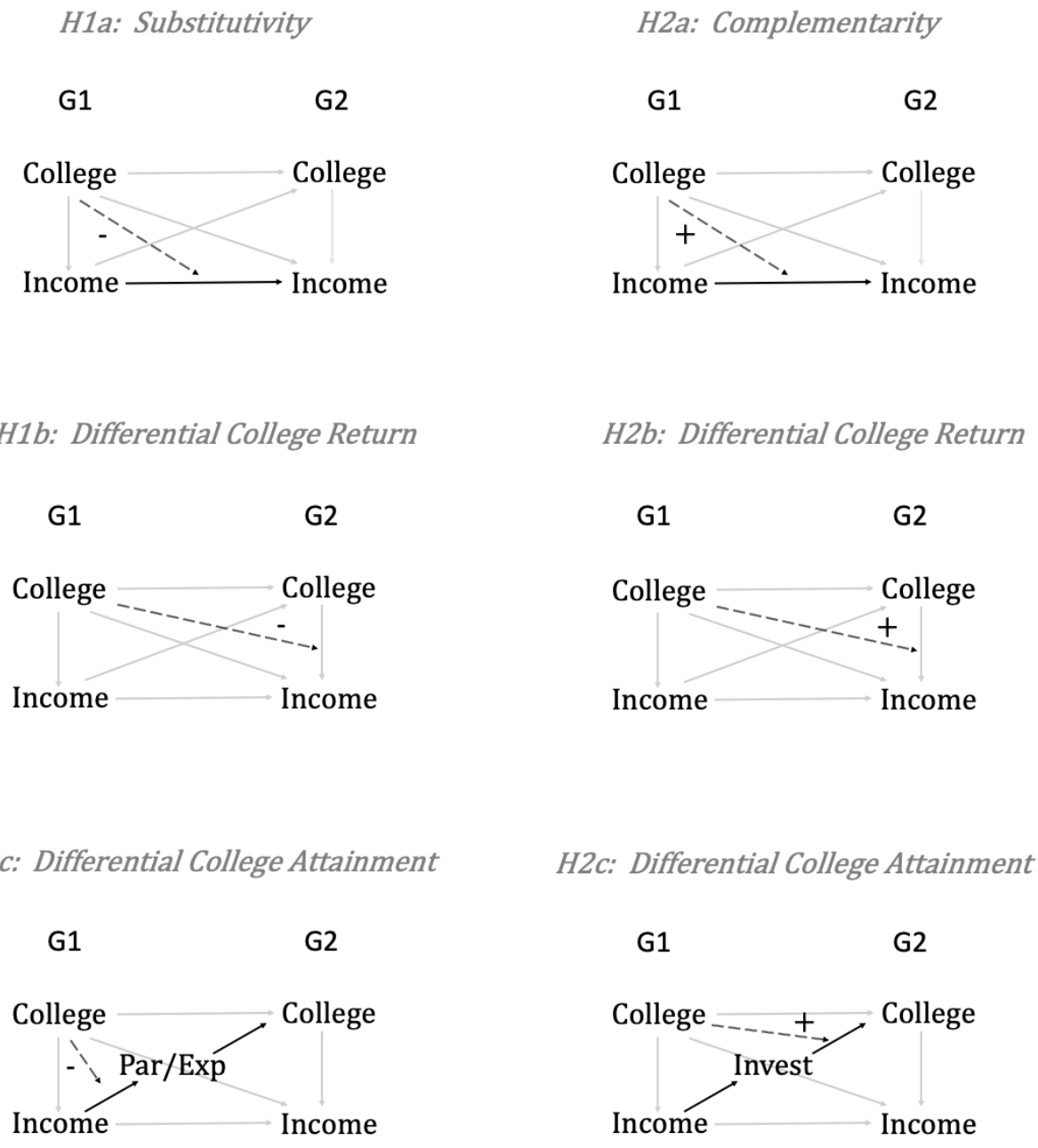


Figure 1 Theoretical Framework and Core Hypotheses

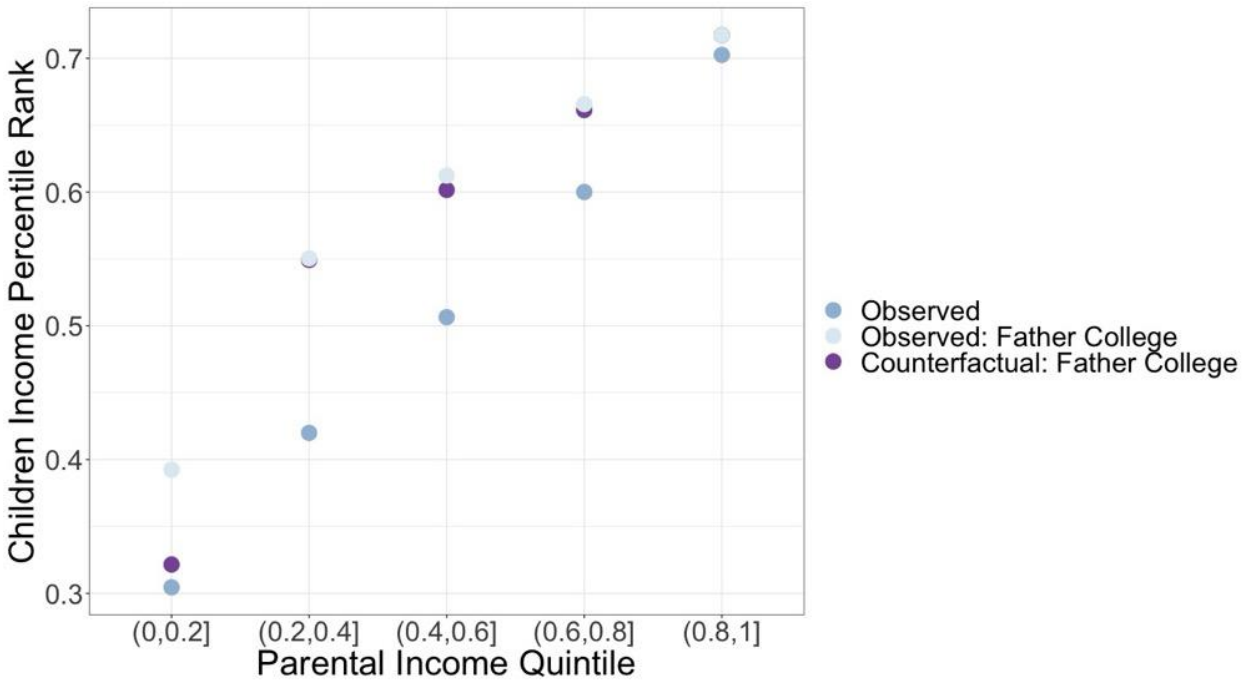


Figure 2 Observed and Counterfactual Children Income by Parental Income Quintile

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Key Variables							
Parental Income	7746	0.5	0.29	0	0.25	0.75	1
Children Income	2521	0.5	0.29	0	0.25	0.75	1
Father College	10177	0.23	0.42	0	0	0	1
College Expectation	10169	0.38	0.42	0	0	1	1
Parenting Style	3563	0	1	-4	-0.6	0.69	3.5
Logged Educational Expense	1437	6	1.7	-4.6	5.3	6.5	12
Controls							
Race	10187						
... White	5395	53%					
... Black	4181	41%					
... Other	611	6%					
Depression in Youth	10187						
... Yes	530	5%					
... No	9657	95%					
Drug Problem in Youth	10187						
... Yes	172	2%					
... No	10015	98%					
Self-Rated Health in Youth	10187	1.7	0.9	1	1	2	5
Psychological Issue in Youth	10187						
... Yes	190	2%					
... No	9997	98%					
Region Growing Up	10187						
... Northeast	1220	12%					
... North Central	2710	27%					
... South	4017	39%					
... West	1437	14%					
... Alaska, Hawaii	7	0%					
... Foreign	796	8%					
Residence Growing Up	10187						
... Country	814	8%					
... Small Town	4939	48%					
... Large City	4191	41%					
... Other	243	2%					
Self-Rated SES in Youth	10187	3.1	1.5	1	1	5	5
Live with Both Parents in Youth	10187						
... Yes	6781	67%					

... No	3406	33%

Table 1 Descriptive Statistics of Key Analytic Variables and Control Variables: PSID & CDS I-III

	G1-G2 Mobility			G2-G3 Mobility		
	Overall	G2 College	G2 Non-College	Overall	G2 College	G2 Non-College
Intercept	0.30	0.54	0.30	0.30	0.39	0.30
	(0.01)	(0.03)	(0.01)	(0.01)	(0.04)	(0.01)
Parental Income	0.41	0.30	0.40	0.40	0.32	0.38
	(0.02)	(0.04)	(0.02)	(0.02)	(0.06)	(0.02)
N	2,730	446	2,284	2,730	446	2,284

Table 2 PSID 1968-2019: Intergenerational Income Rank-Rank Correlation for Three Generations by G2 College Completion Status

	Income	Income	College
Intercept	0.26	0.41	0.02
	(0.01)	(0.01)	(0.01)
Parental Income	0.47		0.28
	(0.02)		(0.03)
Father College	0.18	0.17	0.13
	(0.04)	(0.02)	(0.05)
College		0.21	
		(0.01)	
Parental Income * Father College	-0.15		0.004
	(0.06)		(0.07)
College*Father College		-0.07	
		(0.03)	
N	2,335	2517	3705

Table 3 OLS Results for the Associations between Income, Parental Income, College, and Father College (PSID-CDS I-III)

	Expectation	Parenting Style	College
Intercept	0.24	-0.74	0.04
	(0.01)	(0.03)	(0.04)
Parental Income	0.29	1.40	
	(0.02)	(0.07)	
Father College	0.15	0.94	-0.02
	(0.04)	(0.13)	(0.10)
Parental Income * Father College	-0.15	-0.83	
	(0.05)	(0.18)	
Expenditure			0.02
			(0.01)
Expenditure*Father College			0.04
			(0.02)
N	7,726	3,177	1,390

Table 4 OLS Results for the Associations between Expectation, Parenting Style, Expenditure, Parental Income, College, and Father College (PSID-CDS 1997-2021)

		Condition		
		College	Non-College	Total
Counterfactual	College	0.32	0.43	0.38
		(0.01)	(0.03)	(0.03)
	Non-College	0.44	0.47	0.39
		(0.03)	(0.02)	(0.02)
	Observed	0.32	0.47	0.49
		(0.01)	(0.02)	(0.02)

TEAT	TEAUT	RTEA	TEA
-0.13	-0.03	-0.12	-0.01
(0.04)	(0.02)	(0.01)	(0.04)

Table 5 Conditional and Counterfactual Income Mobility and Causal Quantities of Interest

	Observed	TEA	TEAT	TEAUT	RTEA
Expectation ~ Father College * Parental Income	-0.016 (0.004)	-0.009 (0.005)	-0.022 (0.004)	0.006 (0.002)	-0.011 (0.004)
Parenting Style ~ Father College * Parental Income	-0.066 (0.010)	0.008 (0.013)	-0.075 (0.011)	0.009 (0.006)	-0.011 (0.010)
College ~ Father College* Expenditure	0.059 (0.056)	0.098 (0.069)	0.051 (0.067)	0.009 (0.028)	0.068 (0.052)

Table 6 Observed Covariance Difference by Father College Completion and Treatment Effects of Father Completion on Covariances between Parental Income and Expectation, Parental Income and Parenting Style, and Expenditure and College Completion

CONCLUSION

Intergenerational mobility has long been a cornerstone topic in sociology, with a tradition of research spanning at least half a century. This tradition has yielded numerous studies that offer precise descriptions of reality, shed light on theoretical perspectives, and inform policy discussions. However, the emergence of the "causal revolution" in recent decades has brought to light some limitations inherent in the description-centered research paradigm. While policy implications are often discussed in these studies, they are rarely explicitly modeled within research articles.

Motivated by this methodological gap, this dissertation introduces and applies novel causal inference methods for studying social stratification. It contributes to the recent trend in sociological methodology that focuses on developing causal inference techniques specifically designed for social stratification research (Lundberg et al., 2021; Lundberg, 2022). Chapter 1 delves into the consequences of social mobility by redefining a classic concept, the mobility effect, to establish a more clearly and precisely defined causal quantity. Chapter 2 explores the mechanisms underlying social mobility using a novel causal decomposition technique extended from the Kitagawa-Blinder-Oaxaca decomposition. This method focuses on the transmission mechanism, capturing the immobility that would be eliminated if inequality in a mediator were removed. Chapter 3 investigates the causes of mobility. I examine how parental education may influence intergenerational mobility by evaluating a new estimand that captures the causal effect of college attendance on the income association with children.

The methodology proposed in this dissertation extends beyond the study of social mobility and offers broader contributions to the field of social stratification research. Existing causal inference techniques primarily focus on assessing the effect of a treatment on individual

outcomes. However, social stratification research is inherently concerned with understanding the impact of social changes on inequality, often characterized by associations between variables rather than singular outcomes. This methodological disparity presents a challenge, as it limits researchers' ability to derive theoretically and policy-relevant causal insights from their analyses. The proposals put forth in this dissertation aim to address this challenge by reorienting causal inference techniques from individual treatment effects to macro-level interventional effects. Drawing heavily from the concept of "causal intervention" (VanderWeele and Robinson, 2014; Jackson and VanderWeele, 2018; Lundberg, 2022), I explore methods for evaluating the causes, consequences, and mechanisms of social stratification—specifically, the associations between two outcomes.

In addition to their methodological contributions, the papers in this dissertation offer substantive insights into the study of intergenerational mobility, particularly within the context of the United States. Chapter 1 reassesses a classic proposal that links increasing social mobility during industrialization to decreasing fertility rates observed during the second demographic transition. Contrary to the famous argument, this article finds no evidence supporting the notion that increasing social mobility leads to higher fertility rates.

Chapter 2 addresses another fundamental issue in mobility research: the mechanisms of intergenerational status transmission. This chapter explores a mechanism not systematically evaluated in prior literature: the roles of labor market demand and skills. The findings reveal a shift in the channel of income transmission in recent decades, with non-cognitive skills gaining importance while cognitive skills diminish in their impact on intergenerational transmission. This brings new perspective to the theory of effective inequality maintenance by highlighting the importance of the labor market.

Chapter 3 revisits the debate surrounding the impact of college education on intergenerational mobility. This debate pits two schools of thought against each other: one contends that college serves as a crucial infrastructure for promoting upward mobility, while the other argues that the equalizing role of college is overstated. This chapter contributes to the discussion by evaluating the potential cross-generation equalizing effect of college, a dimension rarely explored by earlier scholars. The findings suggest that the children of college-educated parents indeed experience higher social mobility, indicating a positive cross-generation spillover on social inequality. This long-term benefit aligns with the argument advocating for college as an equalizer in society.

While each individual chapter of the dissertation may suggest different substantive directions for future research, there are common methodological areas that could benefit all three chapters. Here are some potential avenues for further methodological research.

First, the technique of sensitivity analysis needs to be developed for causal interventional study. Sensitivity analysis plays a crucial role in understanding the robustness of results derived from observational data to potential unobserved confounders. As of now, there is no systematic discussion of sensitivity analysis for causal intervention.

Second, it is imperative to compare observational-based results with experiment-based results to understand the precision of observational findings in these research contexts. Such comparisons can provide valuable insights into the reliability and validity of observational studies in capturing causal effects.

Third, novel methods are required to address general equilibrium effects. Intervening in complex social systems poses challenges that can compromise causal estimates. For instance, assessing the effect of a hypothetical intervention like sending everyone to college on social

mobility requires assumptions about the stability of returns to college. If the intervention is so dramatic that alters returns to college, then we cannot know post-intervention return with our observed data, and naturally cannot have any valid causal estimate. Procedures to evaluate sensitivity to general equilibrium effects, akin to sensitivity analysis for ignorability, are currently underdeveloped and warrant further exploration.

BIBLIOGRAPHY

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, 4:1043–1171. Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2019. "Automation and New Tasks: How Technology Displaces and Reinstates Labor." *Journal of Economic Perspectives* 33 (2): 3–30.

Alabdulkareem, Ahmad, Morgan R Frank, Lijun Sun, Bedoor AlShebli, César Hidalgo, and Iyad Rahwan. 2018. "Unpacking the Polarization of Workplace Skills." *Science Advances* 4 (7): eaao6030.

Alwin, Duane F, and Robert M Hauser. 1975. "The Decomposition of Effects in Path Analysis." *American Sociological Review*, 37–47.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.

Armstrong, Elizabeth A., and Laura T. Hamilton. 2013. *Paying for the Party: How College Maintains Inequality*. Harvard University Press.

Atalay, Enghin, Phai Phongthiengtham, Sebastian Sotelo, and Daniel Tannenbaum. 2020. "The Evolution of Work in the United States." *American Economic Journal: Applied Economics* 12 (2): 1–34.

Attewell, Paul, David Lavin, Thurston Domina, and Tania Levey. 2007. *Passing the Torch: Does Higher Education for the Disadvantaged Pay off across the Generations?* Russell Sage Foundation.

Autor, David. 2015. "Why Are There Still so Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30.

Autor, David H, Frank Levy, and Richard J Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118 (4): 1279–1333.

Barthelemy, Johan, and Thomas Suesse. 2018. "Mipfp: An R Package for Multidimensional Array Fitting and Simulating Multivariate Bernoulli Distributions." *Journal of Statistical Software* 86 (20): 10–18637.

Bartley, Mel, and Ian Plewis. 1997. "Does Health-Selective Mobility Account for Socioeconomic Differences in Health? Evidence from England and Wales, 1971 to 1991." *Journal of Health and Social Behavior*, 376–86.

———. 2007. "Increasing Social Mobility: An Effective Policy to Reduce Health Inequalities." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 170 (2): 469–81.

Basse, Guillaume, and Avi Feller. 2018. "Analyzing Two-Stage Experiments in the Presence of Interference." *Journal of the American Statistical Association* 113 (521): 41–55.

Bassok, Daphna, Jenna E Finch, RaeHyuck Lee, Sean F Reardon, and Jane Waldfogel. 2016. "Socioeconomic Gaps in Early Childhood Experiences: 1998 to 2010." *Aera Open* 2 (3): 2332858416653924.

Basu, Debabrata. 2011. "On the Elimination of Nuisance Parameters." In *Selected Works of Debabrata Basu*, 279–90. Springer.

Beaudry, Paul, David A Green, and Benjamin M Sand. 2016. "The Great Reversal in the Demand for Skill and Cognitive Tasks." *Journal of Labor Economics* 34 (S1): S199–247.

Becker, Gary S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70 (5, Part 2): 9–49.

Becker, Gary S. 2010. *The Economics of Discrimination*. University of Chicago press.

Becker, Gary S., Scott Duke Kominers, Kevin M. Murphy, and Jörg L. Spenkuch. 2018. "A Theory of Intergenerational Mobility." *Journal of Political Economy* 126 (S1): S7–25.

Bell, Daniel. 1976. "The Coming of the Post-Industrial Society." In *The Educational Forum*, 40:574–79. Taylor & Francis.

Bessen, James. 2015. "Toil and Technology: Innovative Technology Is Displacing Workers to New Jobs Rather than Replacing Them Entirely." *Finance & Development* 52 (001).

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2003. "Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital." National Bureau of Economic Research.

Blalock Jr, Hubert M. 1966. "The Identification Problem and Theory Building: The Case of Status Inconsistency." *American Sociological Review*, 52–61.

Blanden, Jo, Paul Gregg, and Lindsey Macmillan. 2013. "Intergenerational Persistence in Income and Social Class: The Effect of Within-Group Inequality." *Journal of the Royal Statistical Society Series A: Statistics in Society* 176 (2): 541–63.

———. 2007. "Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education," *The Economic Journal* 117, no. 519 (2007): C43-C60.

Blanden, Jo, Robert Haveman, Timothy Smeeding, and Kathryn Wilson. 2014. "Intergenerational Mobility in the United States and Great Britain: A Comparative Study of Parent–Child Pathways." *Review of Income and Wealth* 60 (3): 425–49.

Blane, David, Seeromanie Harding, and Michael Rosato. 1999. "Does Social Mobility Affect the Size of the Socioeconomic Mortality Differential?: Evidence from the Office for National Statistics Longitudinal Study." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 162 (1): 59–70.

Blau, Peter M, and Otis Dudley Duncan. 1967. *The American Occupational Structure*. Wiley.

Blau, Peter Michael. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. Vol. 7. Free Press New York.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources*, 436–55.

Bloome, Deirdre, Shauna Dyer, and Xiang Zhou. 2018. "Educational Inequality, Educational Expansion, and Intergenerational Income Persistence in the United States." *American Sociological Review* 83 (6): 1215–53.

Bloome, Deirdre, and Daniel Schrage. 2021. "Covariance Regression Models for Studying Treatment Effect Heterogeneity Across One or More Outcomes: Understanding How Treatments Shape Inequality." *Sociological Methods & Research* 50 (3): 1034–72.

Bol, Thijs, Christina Ciocca Eller, Herman G Van De Werfhorst, and Thomas A DiPrete. 2019. "School-to-Work Linkages, Educational Mismatches, and Labor Market Outcomes." *American Sociological Review* 84 (2): 275–307.

Bonvini, Matteo, Alec McClean, Zach Branson, and Edward H Kennedy. 2021. "Incremental Causal Effects: An Introduction and Review." *arXiv Preprint arXiv:2110.10532*.

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. Wiley.

Bound, John, and George E Johnson. 1989. "Changes in the Structure of Wages during the 1980's: An Evaluation of Alternative Explanations." National Bureau of Economic Research.

Bourdieu, Pierre. 1977. "The Economics of Linguistic Exchanges." *Social Science Information* 16 (6): 645–68.

Bourdieu, Pierre, and Jean-Claude Passeron. 1990. *Reproduction in Education, Society and Culture*. SAGE.

Bourne, Mollie, Erzsébet Bukodi, Bastian Betthäuser, and John H. Goldthorpe. 2018. "'Persistence of the Social': The Role of Cognitive Ability in Mediating the Effects of Social Origins on Educational Attainment in Britain." *Research in Social Stratification and Mobility* 58 (December): 11–21.

Bowen, William G, and Derek Bok. 2016. *The Shape of the River*. Princeton University Press.

Boyle, Paul J, Paul Norman, and Frank Popham. 2009. "Social Mobility: Evidence That It Can Widen Health Inequalities." *Social Science & Medicine* 68 (10): 1835–42.

Bradley, Robert H., and Bettye M. Caldwell. 1984. "The HOME Inventory and Family Demographics." *Developmental Psychology* 20 (2): 315–20.

Brand, Jennie E., and Yu Xie. 2010. "Who Benefits Most from College?: Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75 (2): 273–302.

Breen, Richard. 2010. "Educational Expansion and Social Mobility in the 20th Century." *Social Forces* 89 (2): 365–88.

Breen, Richard, and John H Goldthorpe. 1997. "Explaining Educational Differentials: Towards a Formal Rational Action Theory." *Rationality and Society* 9 (3): 275–305.

Breen, Richard, and Jan O. Jonsson. 2007. "Explaining Change in Social Fluidity: Educational Equalization and Educational Expansion in Twentieth-Century Sweden." *American Journal of Sociology* 112 (6): 1775–1810.

Breen, Richard, Carina Mood, and Jan O Jonsson. 2016. "How Much Scope for a Mobility Paradox? The Relationship between Social and Income Mobility in Sweden." *Sociological Science* 3: 39.

Breen, Richard, and Walter Müller. 2020. *Education and Intergenerational Social Mobility in Europe and the United States*. Stanford University Press.

Breen, Richard, and Christopher T Whelan. 1994. "Modelling Trends in Social Fluidity: The Core Model and a Measured-Variable Approach Compared." *European Sociological Review* 10 (3): 259–72.

Breen, Richard, and Meir Yaish. 2006. "Chapter Nine. Testing the Breen-Goldthorpe Model of Educational Decision Making." In *Chapter Nine. Testing the Breen-Goldthorpe Model of Educational Decision Making*, 232–58. Stanford University Press.

Brumback, Babette A, Miguel A Hernán, Sebastien JPA Haneuse, and James M Robins. 2004. "Sensitivity Analyses for Unmeasured Confounding Assuming a Marginal Structural Model for Repeated Measures." *Statistics in Medicine* 23 (5): 749–67.

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. WW Norton & Company.

Bukodi, E., and J. H. Goldthorpe. 2013. "Decomposing 'Social Origins': The Effects of Parents' Class, Status, and Education on the Educational Attainment of Their Children." *European Sociological Review* 29 (5): 1024–39.

Bukodi, Erzsebet, and John H. Goldthorpe. 2021. "Meritocracy and Populism – Is There a Connection?" OSF. https://doi.org/10.31235/osf.io/qgrkf.

Bukodi, Erzsébet, and John H Goldthorpe. 2022. “Intergenerational Class Mobility in Industrial and Post-Industrial Societies: Towards a General Theory.” *Rationality and Society* 34 (3): 271–301.

Bukodi, Erzsébet, John H Goldthorpe, and Jouni Kuha. 2017. “The Pattern of Social Fluidity within the British Class Structure: A Topological Model.” *Journal of the Royal Statistical Society. Series A: Statistics in Society* 180 (3): 841–62.

Calarco, Jessica McCrory. 2011. “‘I Need Help!’ Social Class and Children’s Help-Seeking in Elementary School.” *American Sociological Review* 76 (6): 862–82.

———. 2014. “Coached for the Classroom: Parents’ Cultural Transmission and Children’s Reproduction of Educational Inequalities.” *American Sociological Review* 79 (5): 1015–37.

Callaway, Brantly, Weige Huang, and others. 2019. “Local Intergenerational Elasticities.” *Economics Bulletin* 39 (2): 919–28.

Carneiro, Pedro, Karsten T Hansen, and James J Heckman. 2003. “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College.” National Bureau of Economic Research

Carneiro, Pedro, James J Heckman, and Edward Vytlacil. 2001. “Estimating the Return to Education When It Varies among Individuals.” Unpublished Manuscript.

———. 2010. “Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin.” *Econometrica* 78 (1): 377–94.

Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2013. “Maternal Education, Home Environments, and the Development of Children and Adolescents.” *Journal of the European Economic Association* 11 (suppl_1): 123–60.

Case, Anne, and Angus Deaton. 2020. *Deaths of Despair and the Future of Capitalism*. Princeton University Press.

Castex, Gonzalo, and Evgenia Kogan Dechter. 2014. "The Changing Roles of Education and Ability in Wage Determination." *Journal of Labor Economics* 32 (4): 685–710.

Chan, Tak Wing. 2018. "Social Mobility and the Well-Being of Individuals." *The British Journal of Sociology* 69 (1): 183–206.

Chan, Tak Wing, and John H Goldthorpe. 2007. "Class and Status: The Conceptual Distinction and Its Empirical Relevance." *American Sociological Review* 72 (4): 512–32.

Cheng, Siwei, Jennie E Brand, Xiang Zhou, Yu Xie, and Michael Hout. 2021. "Heterogeneous Returns to College over the Life Course." *Science Advances* 7 (51): eabg7641.

Cheng, Siwei, and Barum Park. 2020. "Flows and Boundaries: A Network Approach to Studying Occupational Mobility in the Labor Market." *American Journal of Sociology* 126 (3): 577–631.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *The Econometrics Journal*, 21(1).

Chetty, Raj, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell. 2022. *Measuring Distribution and Mobility of Income and Wealth*. University of Chicago Press.

Chetty, Raj, John N. Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2017. "Mobility Report Cards: The Role of Colleges in Intergenerational Mobility." National Bureau of Economic Research.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility across Colleges in the United States." *The Quarterly Journal of Economics* 135 (3): 1567–1633.

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang. 2017. "The Fading American Dream: Trends in Absolute Income Mobility since 1940." *Science* 356 (6336): 398–406.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review* 104 (5): 141–47.

Chevalier, Arnaud, Colm Harmon, Vincent O' Sullivan, and Ian Walker. 2013. "The Impact of Parental Income and Education on the Schooling of Their Children." *IZA Journal of Labor Economics* 2 (1): 8.

Chevalier, Arnaud, Colm Harmon, Vincent O'Sullivan, Ian Walker, David B Grusky, and Kim A Weeden. 2006. "Does the Sociological Approach to Studying Social Mobility Have a Future?" *IZA Journal of Labor Economics* 2 (1): 85–108.

Chmielewski, Anna K. 2019. "The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015." *American Sociological Review* 84 (3): 517–44.

Clark, Terry Nichols, Seymour Martin Lipset, and Michael Rempel. 1993. "The Declining Political Significance of Social Class." *International Sociology* 8 (3): 293–316.

Claussen, Bjorgulf, Jeroen Smits, Oyvind Naess, and George Davey Smith. 2005. "Intragenerational Mobility and Mortality in Oslo: Social Selection versus Social Causation." *Social Science & Medicine* 61 (12): 2513–20.

Cobb-Clark, Deborah A, and Michelle Tan. 2011. “Noncognitive Skills, Occupational Attainment, and Relative Wages.” *Labour Economics* 18 (1): 1–13.

Covay, Elizabeth, and William Carbonaro. 2010. “After the Bell: Participation in Extracurricular Activities, Classroom Behavior, and Academic Achievement.” *Sociology of Education* 83 (1): 20–45.

Cunha, Flavio, and James Heckman. 2007. “The Technology of Skill Formation.” *American Economic Review* 97 (2): 31–47.

Dale, Stacy Berg, and Alan B Krueger. 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.

Davis-Kean, Pamela E. 2005. “The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment.” *Journal of Family Psychology* 19 (2): 294–304.

Deming, David J. 2017. “The Growing Importance of Social Skills in the Labor Market.” *The Quarterly Journal of Economics* 132 (4): 1593–1640.

Deming, David J, and Kadeem Noray. 2020. “Earnings Dynamics, Changing Job Skills, and STEM Careers.” *The Quarterly Journal of Economics* 135 (4): 1965–2005.

Díaz, Iván, and Mark J van der Laan. 2013. “Assessing the Causal Effect of Policies: An Example Using Stochastic Interventions.” *The International Journal of Biostatistics* 9 (2): 161–74.

DiMaggio, Paul. 1982. “Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students.” *American Sociological Review* 47 (2): 189–201.

DiMaggio, Paul, and John Mohr. 1985. “Cultural Capital, Educational Attainment, and Marital Selection.” *American Journal of Sociology* 90 (6): 1231–61.

DiPrete, Thomas A. 2020. “The Impact of Inequality on Intergenerational Mobility.” *Annual Review of Sociology* 46: 379–98.

Doren, Catherine, and Eric Grodsky. 2016. “What Skills Can Buy: Transmission of Advantage through Cognitive and Noncognitive Skills.” *Sociology of Education* 89 (4): 321–42.

Dorie, Vincent, Jennifer Hill, Uri Shalit, Marc Scott, and Dan Cervone. 2019. “Automated versus Do-It-Yourself Methods for Causal Inference: Lessons Learned from a Data Analysis Competition.” *Statistical Science* 34 (1): 43–68.

Duncan, Greg J., and Richard J. Murnane. 2011. *Whither Opportunity?: Rising Inequality, Schools, and Children’s Life Chances*. Russell Sage Foundation.

Duncan, Otis. 1966. “Methodological Issues in the Analysis of Social Mobility.” In *Social Structure and Mobility in Economic Development*, edited by Neil J Smelser and Seymour Martin Lipset. Transaction Publishers.

Duncan, Otis Dudley. 1961. “A Socioeconomic Index for All Occupations.” *Occupations and Social Status*.

———. 1966. “Path Analysis: Sociological Examples.” *American Journal of Sociology* 72 (1): 1–16.

———. 1979. “How Destination Depends on Origin in the Occupational Mobility Table.” *American Journal of Sociology* 84 (4): 793–803.

Durlauf, Steven N., and Ananth Seshadri. 2018. “Understanding the Great Gatsby Curve.” *NBER Macroeconomics Annual* 32 (April): 333–93.

Edin, Per-Anders, Peter Fredriksson, Martin Nybom, and Björn Öckert. 2022. “The Rising Return to Noncognitive Skill.” *American Economic Journal: Applied Economics* 14 (2): 78–100.

Ellis, Robert A, and W Clayton Lane. 1963. “Structural Supports for Upward Mobility.” *American Sociological Review*, 743–56.

———. 1967. “Social Mobility and Social Isolation: A Test of Sorokin’s Dissociative Hypothesis.” *American Sociological Review*, 237–53.

Erikson, Robert, and John H Goldthorpe. 1992. *The Constant Flux: A Study of Class Mobility in Industrial Societies*. Oxford University Press.

Erikson, Robert, John H Goldthorpe, and Lucienne Portocarero. 1979. “Intergenerational Class Mobility in Three Western European Societies: England, France and Sweden.” *The British Journal of Sociology* 30 (4): 415–41.

Esping-Andersen, Gøsta, and Jorge Cimentada. 2018. “Ability and Mobility: The Relative Influence of Skills and Social Origin on Social Mobility.” *Social Science Research* 75 (September): 13–31.

Fan, Jianqing, and Qiwei Yao. 1998. “Efficient Estimation of Conditional Variance Functions in Stochastic Regression.” *Biometrika* 85 (3): 645–60.

Fiel, Jeremy E. 2020. “Great Equalizer or Great Selector? Reconsidering Education as a Moderator of Intergenerational Transmissions.” *Sociology of Education* 93 (4): 353–71.

Fienberg, Stephen E. 1970. “An Iterative Procedure for Estimation in Contingency Tables.” *The Annals of Mathematical Statistics* 41 (3): 907–17.

Fletcher, Jason M, and Barbara Wolfe. 2016. “The Importance of Family Income in the Formation and Evolution of Non-Cognitive Skills in Childhood.” *Economics of Education Review* 54: 143–54.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. “Decomposition Methods in Economics.” In *Handbook of Labor Economics*, 4:1–102. Elsevier.

Fosse, Ethan, and Fabian Pfeffer. 2019. “The Individual-Level Consequences of Social Mobility.” Working Paper

Fosse, Ethan, and Christopher Winship. 2019. “Analyzing Age-Period-Cohort Data: A Review and Critique.” *Annual Review of Sociology* 45: 467–92.

Frank, Morgan R, David Autor, James E Bessen, Erik Brynjolfsson, Manuel Cebrian, David J Deming, Maryann Feldman, et al. 2019. “Toward Understanding the Impact of Artificial Intelligence on Labor.” *Proceedings of the National Academy of Sciences* 116 (14): 6531–39.

Friedman, Sam. 2016. “Habitus Clivé and the Emotional Imprint of Social Mobility.” *The Sociological Review* 64 (1): 129–47.

Goldin, Claudia, and Lawrence F. Katz. 2009. *The Race between Education and Technology*. Harvard University Press.

Goldthorpe, John H. 2013. “Understanding – and Misunderstanding – Social Mobility in Britain: The Entry of the Economists, the Confusion of Politicians and the Limits of Educational Policy.” *Journal of Social Policy* 42 (3): 431–50.

Goodhart, David. 2020. *Head Hand Heart: The Struggle for Dignity and Status in the 21st Century*. Penguin UK.

Goodman, Leo A, and William H Kruskal. 1979. *Measures of Association for Cross Classifications*. Springer.

Gordon, Robert J. 2012. "Is US Economic Growth over? Faltering Innovation Confronts the Six Headwinds." National Bureau of Economic Research.

Grace, Katja, John Salvatier, Allan Dafoe, Baobao Zhang, and Owain Evans. 2018. "When Will AI Exceed Human Performance? Evidence from AI Experts." *Journal of Artificial Intelligence Research* 62: 729–54.

Grusky, David B, and Jesper B Sorensen. 1998. "Can Class Analysis Be Salvaged?" *American Journal of Sociology* 103 (5): 1187–1234.

Grusky, David B, and Kim A Weeden. 2001. "Decomposition without Death: A Research Agenda for a New Class Analysis." *Acta Sociologica* 44 (3): 203–18.

Hadjar, Andreas, and Robin Samuel. 2015. "Does Upward Social Mobility Increase Life Satisfaction? A Longitudinal Analysis Using British and Swiss Panel Data." *Research in Social Stratification and Mobility* 39: 48–58.

Haller, Archibald O., and Alejandro Portes. 1973. "Status Attainment Processes." *Sociology of Education* 46 (1): 51–91.

Hamilton, Laura, Josipa Roksa, and Kelly Nielsen. 2018. "Providing a '"Leg up"': Parental Involvement and Opportunity Hoarding in College." *Sociology of Education* 91 (2): 111–31.

Haneuse, Sebastian, and Andrea Rotnitzky. 2013. "Estimation of the Effect of Interventions That Modify the Received Treatment." *Statistics in Medicine* 32 (30): 5260–77.

Hao, Lingxin, and Wei-Jun Jean Yeung. 2015. "Parental Spending on School-Age Children: Structural Stratification and Parental Expectation." *Demography* 52 (3): 835–60.

Hauser, Robert M. 1980. “Some Exploratory Methods for Modeling Mobility Tables and Other Cross-Classified Data.” *Sociological Methodology* 11: 413–58.

———. 2010. “Intergenerational Economic Mobility in the United States Measures, Differentials and Trends.” Unpublished Manuscript

Hauser, Robert M, Shu-Ling Tsai, and William H Sewell. 1983. “A Model of Stratification with Response Error in Social and Psychological Variables.” *Sociology of Education*, 20–46.

Hauser, Robert M, and John Robert Warren. 1997. “Socioeconomic Indexes for Occupations: A Review, Update, and Critique.” *Sociological Methodology* 27 (1): 177–298.

Haveman, Robert, and Barbara Wolfe. 1995. “The Determinants of Children’s Attainments: A Review of Methods and Findings.” *Journal of Economic Literature* 33 (4): 1829–78.

Heckman, James J, and Pedro Carneiro. 2003. “Human Capital Policy.” National Bureau of Economic Research.

Heckman, James J, Hidehiko Ichimura, Jeffrey A Smith, and Petra E Todd. 1998. “Characterizing Selection Bias Using Experimental Data.” National Bureau of Economic Research.

Heckman, James J, Lance Lochner, and Christopher Taber. 1998. “Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents.” *Review of Economic Dynamics* 1 (1): 1–58.

Heckman, James J, and Stefano Mosso. 2014. “The Economics of Human Development and Social Mobility.” *Annual Review of Economics*. 6, no. 1 (2014): 689-733.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior.” *Journal of Labor Economics* 24 (3): 411–82.

Heckman, James J, Sergio Urzua, and Edward Vytlacil. 2006. “Understanding Instrumental Variables in Models with Essential Heterogeneity.” *The Review of Economics and Statistics* 88 (3): 389–432.

Heckman, James J, and Edward Vytlacil. 2001. “Policy-Relevant Treatment Effects.” *American Economic Review* 91 (2): 107–11.

———. 2005. “Structural Equations, Treatment Effects, and Econometric Policy Evaluation 1.” *Econometrica* 73 (3): 669–738.

Hendrickx, John, Nan Dirk De Graaf, Jan Lammers, and Wout Ultee. 1993. “Models for Status Inconsistency and Mobility: A Comparison of the Approaches by Hope and Sobel with the Mainstream Square Additive Model.” *Quality and Quantity* 27 (4): 335–52.

Hines, Oliver, Oliver Dukes, Karla Diaz-Ordaz, and Stijn Vansteelandt. 2021. “Demystifying Statistical Learning Based on Efficient Influence Functions.” *The American Statistician*, no. just-accepted: 1–48.

Hoff, Peter D., and Xiaoyue Niu. 2012. “A Covariance Regression Model.” *Statistica Sinica* 22 (2): 729–53.

Holland, Paul W. 1986. “Statistics and Causal Inference.” *Journal of the American Statistical Association* 81 (396): 945–60.

Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. “The Causal Effect of Parents’ Schooling on Children’s Schooling: A Comparison of Estimation Methods.” *Journal of Economic Literature* 49 (3): 615–51.

Holtmann, Anne Christine, Laura Menze, and Heike Solga. 2021. “Intergenerational Transmission of Educational Attainment: How Important Are Children’s Personality Characteristics?” *American Behavioral Scientist* 65 (11): 1531–54.

Hope, Keith. 1975. “Models of Status Inconsistency and Social Mobility Effects.” *American Sociological Review*, 322–43.

Hopkins, Daniel J, and Gary King. 2010. “A Method of Automated Nonparametric Content Analysis for Social Science.” *American Journal of Political Science* 54 (1): 229–47.

Horowitz, Jonathan. 2018. “Relative Education and the Advantage of a College Degree.” *American Sociological Review* 83 (4): 771–801.

Houle, Jason N, and Molly A Martin. 2011. “Does Intergenerational Mobility Shape Psychological Distress? Sorokin Revisited.” *Research in Social Stratification and Mobility* 29 (2): 193–203.

Hout, Michael. 1984. “Status, Autonomy, and Training in Occupational Mobility.” *American Journal of Sociology* 89 (6): 1379–1409.

———. 1988. “More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s.” *American Journal of Sociology* 93 (6): 1358–1400.

———. 2018. “Americans’ Occupational Status Reflects the Status of Both of Their Parents.” *Proceedings of the National Academy of Sciences* 115 (38): 9527–32.

Hout, Michael, and Thomas A DiPrete. 2006. “What We Have Learned: RC28’s Contributions to Knowledge about Social Stratification.” *Research in Social Stratification and Mobility* 24 (1): 1–20.

Hout, Michael, and Robert M Hauser. 1992. “Symmetry and Hierarchy in Social Mobility: A Methodological Analysis of the CASMIN Model of Class Mobility.” *European Sociological Review* 8 (3): 239–66.

Hsin, Amy, and Yu Xie. 2017. “Life-Course Changes in the Mediation of Cognitive and Non-Cognitive Skills for Parental Effects on Children’s Academic Achievement.” *Social Science Research* 63 (March): 150–65.

Huang, Qing, Xi Song, and Yu Xie. 2022. “Dual Pathways of Intergenerational Influence over Multiple Generations.” *Chinese Sociological Review*, 1–40.

Hudgens, Michael G, and M Elizabeth Halloran. 2008. “Toward Causal Inference with Interference.” *Journal of the American Statistical Association* 103 (482): 832–42.

Humphries, John Eric, and Fabian Kosse. 2017. “On the Interpretation of Non-Cognitive Skills–What Is Being Measured and Why It Matters.” *Journal of Economic Behavior & Organization* 136: 174–85.

Irwin, Sarah, and Sharon Elley. 2013. “Parents’ Hopes and Expectations for Their Children’s Future Occupations.” *The Sociological Review* 61 (1): 111–30.

Jackson, John W. 2018. “On the Interpretation of Path-Specific Effects in Health Disparities Research.” *Epidemiology* 29 (4): 517–20.

———. 2021. “Meaningful Causal Decompositions in Health Equity Research: Definition, Identification, and Estimation through a Weighting Framework.” *Epidemiology* 32 (2): 282.

Jackson, John W., and Tyler J. VanderWeele. 2018. “Decomposition Analysis to Identify Intervention Targets for Reducing Disparities.” *Epidemiology* 29 (6): 825.

Jackson, Michelle. 2009. “Disadvantaged through Discrimination? The Role of Employers in Social Stratification 1.” *The British Journal of Sociology* 60 (4): 669–92.

Jackson, Michelle, John H Goldthorpe, and Colin Mills. 2005. “Education, Employers and Class Mobility.” *Research in Social Stratification and Mobility* 23: 3–33.

Jackson, Michelle, and David B Grusky. 2018. “A Post-Liberal Theory of Stratification.” *The British Journal of Sociology* 69 (4): 1096–1133.

Jæger, Mads Meier. 2007. “Educational Mobility across Three Generations: The Changing Impact of Parental Social Class, Economics, Cultural and Social Capital.” *European Societies* 9 (4): 527–50.

Jencks, Christopher and others. 1972. *Inequality: A Reassessment of the Effect of Family and Schooling in America.* Basic Books.

Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. “What Is a Good Job? A New Measure of Labor-Market Success.” *American Journal of Sociology* 93 (6): 1322–57.

Jonsson, Jan O, David B Grusky, Matthew Di Carlo, Reinhard Pollak, and Mary C Brinton. 2009. “Microclass Mobility: Social Reproduction in Four Countries.” *American Journal of Sociology* 114 (4): 977–1036.

Juhn, Chinhui, Kevin Murphy, and Brooks Pierce. 1993. “Wage Inequality and the Rise in Returns to Skill.” *Journal of Political Economy*, no. 3: 410–42.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47 (2): 363–91.

Kalil, Ariel, Kathleen M Ziol-Guest, Rebecca M Ryan, and Anna J Markowitz. 2016. “Changes in Income-Based Gaps in Parent Activities with Young Children from 1988 to 2012.” *AERA Open* 2 (3): 2332858416653732.

Karlson, Kristian Bernt. 2019. “College as Equalizer? Testing the Selectivity Hypothesis.” *Social Science Research* 80 (May): 216–29.

Kennedy, Edward H. 2019. “Nonparametric Causal Effects Based on Incremental Propensity Score Interventions.” *Journal of the American Statistical Association* 114 (526): 645–56.

Kennedy, Edward H, Zongming Ma, Matthew D McHugh, and Dylan S Small. 2017. "Non-Parametric Methods for Doubly Robust Estimation of Continuous Treatment Effects." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79 (4): 1229–45.

King, Gary, and Ying Lu. 2008. "Verbal Autopsy Methods with Multiple Causes of Death." *Statistical Science* 23, no. 1 (2008): 78.

Kitagawa, Evelyn M. 1955. "Components of a Difference between Two Rates." *Journal of the American Statistical Association* 50 (272): 1168–94.

Koopmans, Tjalling C. 1949. "Identification Problems in Economic Model Construction." *Econometrica, Journal of the Econometric Society*, 125–44.

Krueger, Alan. 2012. "The Rise and Consequences of Inequality." *Presentation Made to the Center for American Progress, January 12th*.

Lareau, Annette. 2002. "Invisible Inequality: Social Class and Childrearing in Black Families and White Families." *American Sociological Review* 67 (5): 747–76.

———. 2011. *Unequal Childhoods: Class, Race, and Family Life*. University of California Press.

———. 2015. "Cultural Knowledge and Social Inequality." *American Sociological Review* 80 (1): 1–27.

Laurison, Daniel, and Sam Friedman. 2016. "The Class Pay Gap in Higher Professional and Managerial Occupations." *American Sociological Review* 81 (4): 668–95.

Lefgren, Lars, Matthew J Lindquist, and David Sims. 2007. "Rich Dad, Smart Dad: Decomposing the Intergenerational Transmission of Income." *Journal of Political Economy*, 36.

Lenski, Gerhard E. 1954. “Status Crystallization: A Non-Vertical Dimension of Social Status.” *American Sociological Review* 19 (4): 405–13.

———. 1956. “Social Participation and Status Crystallization.” *American Sociological Review* 21 (4): 458–64.

Levy, Frank, and Richard J Murnane. 2012. “The New Division of Labor.” In *The New Division of Labor*. Princeton University Press.

Lin, Ken-Hou, and Koit Hung. 2022. “The Network Structure of Occupations: Fragmentation, Differentiation, and Contagion.” *American Journal of Sociology* 127 (5): 1551–1601.

Lin, Nan. 1999. “Social Networks and Status Attainment.” *Annual Review of Sociology* 25 (1): 467–87.

Link, Bruce G, and Jo Phelan. 1995. “Social Conditions as Fundamental Causes of Disease.” *Journal of Health and Social Behavior*, 80–94.

Lipset, Seymour Martin, and Reinhard Bendix. 1991. *Social Mobility in Industrial Society*. Transaction Publishers.

Little, Roderick JA, and Mei-Miau Wu. 1991. “Models for Contingency Tables with Known Margins When Target and Sampled Populations Differ.” *Journal of the American Statistical Association* 86 (413): 87–95.

Liu, Airan. 2020. “Non-Cognitive Skills and the Growing Achievement Gap⋆.” *Research in Social Stratification and Mobility* 69 (October): 100546.

Liu, Airan, and Yu Xie. 2016. “Why Do Asian Americans Academically Outperform Whites? – The Cultural Explanation Revisited.” *Social Science Research* 58 (July): 210–26.

Lucas, Samuel R. 2001. “Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects.” *American Journal of Sociology* 106 (6): 1642–90.

Luce, R Duncan. 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

Lundberg, Ian. 2021. “Quantifying the Contribution of Occupational Segregation to Racial Disparities in Health: A Gap-Closing Perspective.” Working Paper

———. 2022. “The Gap-Closing Estimand: A Causal Approach to Study Interventions That Close Disparities Across Social Categories.” *Sociological Methods & Research*, January, 00491241211055769.

Lundberg, Ian, Rebecca Johnson, and Brandon M Stewart. 2021. “What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory.” *American Sociological Review* 86 (3): 532–65.

Luo, Liying. 2021. “Heterogeneous Effects of Intergenerational Social Mobility: An Improved Method and New Evidence.” *American Sociological Review*.

Manski, Charles F. 2009. *Identification for Prediction and Decision*. Harvard University Press.

———. 2013. “Identification of Treatment Response with Social Interactions.” *The Econometrics Journal* 16 (1): S1–23.

Markovits, Daniel. 2020. *The Meritocracy Trap: How America’s Foundational Myth Feeds Inequality, Dismantles the Middle Class, and Devours the Elite*. Penguin Publishing Group.

Mcfadden, Daniel. 1973. “Conditional Logit Analysis of Qualitative Choice Behavior.” In *Frontiers in Econometrics*, edited by Paul Zarembka. Academic Press.

McFadden, Daniel. 2001. “Economic Choices.” *American Economic Review* 91 (3): 351–78.

McNeal, Ralph B., Jr. 1999. “Parental Involvement as Social Capital: Differential Effectiveness on Science Achievement, Truancy, and Dropping Out.” *Social Forces* 78 (1): 117–44.

Michelle, Jackson., and David. B. Grusky. 2018. “A Post-Liberal Theory of Stratification.” *The British Journal of Sociology* 69: 1096–1133.

Middleton, Eric B, and Teri A Loughead. 1993. “Parental Influence on Career Development: An Integrative Framework for Adolescent Career Counseling.” *Journal of Career Development* 19 (3): 161–73.

Mills, C Wright. 2000. *The Sociological Imagination*. Oxford University Press.

Mitnik, Pablo A, Erin Cumberworth, and David B Grusky. 2016. “Social Mobility in a High-Inequality Regime.” *The ANNALS of the American Academy of Political and Social Science* 663 (1): 140–84.

Mood, Carina. 2017. “More than Money: Social Class, Income, and the Intergenerational Persistence of Advantage.” *Sociological Science* 4: 263–87.

Mood, Carina, Jan O. Jonsson, and Erik Bihagen. 2012. “Socioeconomic Persistence across Generations : Cognitive and Noncognitive Processes.” In *From Parents to Children: the Intergenerational Transmission of Advantage*, 53–84. Russell Sage Foundation.

Moreno, K. E., and D. O. Segall. 1997. “Reliability and Construct Validity of CAT-ASVAB.” In *Computerized Adaptive Testing: From Inquiry to Operation*, edited by B. K. Waters W. A. Sands and J. R. McBride, 169–80. American Psychological Association.

Mott, Frank L. 2004. “The Utility of the HOME-SF Scale for Child Development Research in a Large National Longitudinal Survey: The National Longitudinal Survey of Youth 1979 Cohort.” *Parenting* 4 (2–3): 259–70.

Muñoz, Iván Díaz, and Mark Van Der Laan. 2012. “Population Intervention Causal Effects Based on Stochastic Interventions.” *Biometrics* 68 (2): 541–49.

Naimi, Ashley I, Jay S Kaufman, and Richard F MacLehose. 2014. “Mediation Misgivings: Ambiguous Clinical and Public Health Interpretations of Natural Direct and Indirect Effects.” *International Journal of Epidemiology* 43 (5): 1656–61.

Nam, Charles B, and Mary G Powers. 1983. *The Socioeconomic Approach to Status Measurement: With a Guide to Occupational and Socioeconomic Status Scores*. Cap and Gown Press

Nucci, Larry, Cleanice Camino, and Clary Milnitsky Sapiro. 1996. “Social Class Effects on Northeastern Brazilian Children’s Conceptions of Areas of Personal Choice and Social Regulation.” *Child Development* 67 (3): 1223–42.

Oaxaca, Ronald. 1973. “Male-Female Wage Differentials in Urban Labor Markets.” *International Economic Review*, 693–709.

Ogburn, Elizabeth L, and Tyler J VanderWeele. 2014. “Causal Diagrams for Interference.” *Statistical Science* 29 (4): 559–78.

Pakulski, Jan, and Malcolm Waters. 1996. *The Death of Class*. Sage.

Pearl, Judea. 2001. “Direct and Indirect Effects.” *Probabilistic and Causal Inference: The Works of Judea Pearl*, 373.

———. 2009. *Causality*. Cambridge university press.

Perna, Laura W. 2006. “Studying College Access and Choice: A Proposed Conceptual Model.” *Higher Education: Handbook of Theory and Research*, 99–157.

Pfeffer, Fabian T. 2018. “Growing Wealth Gaps in Education.” *Demography* 55 (3): 1033–68.

Pfeffer, Fabian T., and Florian R. Hertel. 2015. “How Has Educational Expansion Shaped Social Mobility Trends in the United States?” *Social Forces* 94 (1): 143–80.

Raftery, Adrian E, and Michael Hout. 1993. “Maximally Maintained Inequality: Expansion, Reform, and Opportunity in Irish Education, 1921-75.” *Sociology of Education*, 41–62.

Ramey, Garey, and Valerie A Ramey. 2009. “The Rug Rat Race.” National Bureau of Economic Research.

Reardon, Sean F. 2011. “The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations.” *Whither Opportunity* 1 (1): 91–116.

Richey, Jeremiah, and Alicia Rosburg. 2018. “Decomposing Economic Mobility Transition Matrices.” *Journal of Applied Econometrics* 33 (1): 91–108.

Rivera, Lauren A. 2012. “Hiring as Cultural Matching: The Case of Elite Professional Service Firms.” *American Sociological Review* 77 (6): 999–1022.

Rivera, Lauren A. 2016. *Pedigree: How Elite Students Get Elite Jobs*. Princeton University Press.

Robins, James M. 1999. “Association, Causation, and Marginal Structural Models.” *Synthese*, 151–79.

Robins, James M. 2003. “Semantics of Causal DAG Models and the Identification of Direct and Indirect Effects.” *Highly Structured Stochastic Systems*, May, 70–82.

Robins, James M, and Sander Greenland. 1992. “Identifiability and Exchangeability for Direct and Indirect Effects.” *Epidemiology* 3 (2): 143–55.

Robinson, Peter M. 1988. “Root-N-Consistent Semiparametric Regression.” *Econometrica: Journal of the Econometric Society*, 931–54.

Rodgers, Willard L. 1982. “Estimable Functions of Age, Period, and Cohort Effects.” *American Sociological Review*, 774–87.

Rosen, Bernard C. 1956. “The Achievement Syndrome: A Psychocultural Dimension of Social Stratification.” *American Sociological Review* 21 (2): 203–11.

Rubin, Donald B. 1974. “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies.” *Journal of Educational Psychology* 66 (5): 688.

Ryan, Rebecca M, Ariel Kalil, Caitlin Hines, and Kathleen Ziol-Guest. 2020. “Trends in Parental Values in a Period of US Labor Market Change.” *Journal of Marriage and Family* 82 (5): 1495–1514.

Scharfstein, Daniel O, Andrea Rotnitzky, and James M Robins. 1999. “Adjusting for Nonignorable Drop-out Using Semiparametric Nonresponse Models.” *Journal of the American Statistical Association* 94 (448): 1096–1120.

Scheier, Michael F, Charles S Carver, and Michael W Bridges. 1994. “Distinguishing Optimism from Neuroticism (and Trait Anxiety, Self-Mastery, and Self-Esteem): A Reevaluation of the Life Orientation Test.” *Journal of Personality and Social Psychology* 67 (6): 1063.

Sewell, William H., Archibald O. Haller, and George W. Ohlendorf. 1970. “The Educational and Early Occupational Status Attainment Process: Replication and Revision.” *American Sociological Review* 35 (6): 1014–27.

Sewell, William H., and Robert M. Hauser. 1975. *Education, Occupation, and Earnings. Achievement in the Early Career*. Academic Press Inc.

Shavit, Yossi, and Hans-Peter Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries.* Westview Press.

Sobel, Michael E. 1981. “Diagonal Mobility Models: A Substantively Motivated Class of Designs for the Analysis of Mobility Effects.” *American Sociological Review*, 893–906.

———. 1985. “Social Mobility and Fertility Revisited: Some New Models for the Analysis of the Mobility Effects Hypothesis.” *American Sociological Review*, 699–712.

Soldz, Stephen, and George E. Vaillant. 1999. “The Big Five Personality Traits and the Life Course: A 45-Year Longitudinal Study.” *Journal of Research in Personality* 33 (2): 208–32.

Song, Xi, Catherine G Massey, Karen A Rolf, Joseph P Ferrie, Jonathan L Rothbaum, and Yu Xie. 2020. “Long-Term Decline in Intergenerational Mobility in the United States since the 1850s.” *Proceedings of the National Academy of Sciences* 117 (1): 251–58.

Sorokin, Pitirim Aleksandrovich. 1927. *Social Mobility*. New York: Harper and Brothers.

Specht, Jule, Boris Egloff, and Stefan C. Schmukle. 2011. “Stability and Change of Personality across the Life Course: The Impact of Age and Major Life Events on Mean-Level and Rank-Order Stability of the Big Five.” *SOEP papers on Multidisciplinary Panel Data Research* 101 (4): 862–82.

Stryker, Sheldon, and Anne Statham Macke. 1978. “Status Inconsistency and Role Conflict.” *Annual Review of Sociology* 4 (1): 57–90.

Susskind, Richard E, and Daniel Susskind. 2015. *The Future of the Professions: How Technology Will Transform the Work of Human Experts*. Oxford University Press, USA.

Tchetgen, Eric J Tchetgen, and Tyler J VanderWeele. 2012. “On Causal Inference in the Presence of Interference.” *Statistical Methods in Medical Research* 21 (1): 55–75.

Thurow, Lester C. 1972. “Education and Economic Equality.” *The Public Interest* 28: 66.

Torche, Florencia. 2011. “Is a College Degree Still the Great Equalizer? Intergenerational Mobility across Levels of Schooling in the United States.” *American Journal of Sociology* 117 (3): 763–807.

Treiman, Donald J. 1970. “Industrialization and Social Stratification.” *Sociological Inquiry* 40 (2): 207–34.

Van der Laan, Mark J, and Sherri Rose. 2011. *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer Science & Business Media.

VanderWeele, Tyler J., and Whitney R. Robinson. 2014. “On the Causal Interpretation of Race in Regressions Adjusting for Confounding and Mediating Variables.” *Epidemiology* 25 (4): 473.

Weeden, Kim A, and David B Grusky. 2005. “The Case for a New Class Map.” *American Journal of Sociology* 111 (1): 141–212.

———. 2012. “The Three Worlds of Inequality.” *American Journal of Sociology* 117 (6): 1723–85.

Wei, Lai, and Yu Xie. 2022. “Social Mobility as Causal Intervention.” Working Paper

Westoff, Charles F. 1953. “The Changing Focus of Differential Fertility Research: The Social Mobility Hypothesis.” *The Milbank Memorial Fund Quarterly* 31 (1): 24–38.

Willis, Robert J., and Sherwin Rosen. 1979. “Education and Self-Selection.” *Journal of Political Economy* 87 (5, Part 2): S7–36.

Woelfel, Joseph, and Archibald O Haller. 1971. “Significant Others, the Self-Reflexive Act and the Attitude Formation Process.” *American Sociological Review*, 74–87.

Xie, Yu. 1992. “The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables.” *American Sociological Review*, 380–95.

Xie, Yu, Jennie E Brand, and Ben Jann. 2012. “Estimating Heterogeneous Treatment Effects with Observational Data.” *Sociological Methodology* 42 (1): 314–47.

Yao, Qiwei, and Howell Tong. 1994. “Quantifying the Influence of Initial Values on Non-Linear Prediction.” *Journal of the Royal Statistical Society. Series B (Methodological)* 56 (4): 701–25.

Young, Jessica G, Lauren E Cain, James M Robins, Eilis J O’Reilly, and Miguel A Hernán. 2011. “Comparative Effectiveness of Dynamic Treatment Regimes: An Application of the Parametric g-Formula.” *Statistics in Biosciences* 3 (1): 119.

Yu, Ang, and Felix Elwert. 2021. “Causal Decomposition of Group Disparities.” Working Paper

Zhou, Xiang. 2019. “Equalization or Selection? Reassessing the ‘Meritocratic Power’ of a College Degree in Intergenerational Income Mobility.” *American Sociological Review* 84 (3): 459–85.

———. 2022. “Semiparametric Estimation for Causal Mediation Analysis with Multiple Causally Ordered Mediators.” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(3), 794-821.

Zhou, Xiang, and Guanghui Pan. 2023. “Higher Education and the Black-White Earnings Gap.” *American Sociological Review* 88 (1): 154–88.

Zhou, Xiang, and Yu Xie. 2019a. “Marginal Treatment Effects from a Propensity Score Perspective.” *Journal of Political Economy* 127 (6): 3070–84.

———. 2019b. “Market Transition, Industrialization, and Social Mobility Trends in Postrevolution China.” *American Journal of Sociology* 124 (6): 1810–47.

Zhou, Xiang, and Teppei Yamamoto. 2022. “Tracing Causal Paths from Experimental and Observational Data.” *The Journal of Politics* 0 (0): 000–000.