Decomposing economic mobility transition matrices

Richey, Jeremiah and Rosburg, Alicia

Kyungpook National University, University of Northern Iowa

15 August 2015

Decomposing Economic Mobility Transition Matrices

Jeremiah Richey
Kyungpook National University
jarichey80@gmail.com

Alicia Rosburg
University of Northern Iowa
alicia.rosburg@uni.edu

September 7, 2015

Abstract

The intergenerational mobility literature has consistently found that the distribution of adult economic outcomes differ markedly depending on parental economic status, yet much remains to be understood about the drivers or determinants of this relationship. Existing literature on potential drivers focuses primarily on mean effects. To help provide a more complete picture of the potential forces driving economic persistence, we propose a method to decompose transition matrices and related indices. Specifically, we decompose differences between an estimated transition matrix and a benchmark transition matrix into portions attributable to differences in characteristics between individuals from different households (a composition effect) and portions attributable to differing returns to these characteristics between individuals from different households (a structure effect). We also incorporate a detailed decomposition, based on copula theory, that decomposes the composition effect into portions attributable to specific covariates and their interactions. We illustrate our method using data on white men from the 1979 National Longitudinal Survey of Youth. Estimation is based on an extended Mincer equation that includes cognitive and non-cognitive measures. To address the potential endogeneity of education, we implement an IV strategy that allows us to estimate causal effects and investigate the role of unobserved ability.

JEL Classification: C14, C20, J31, J62

Key Words: Intergenerational mobility, Transition matrices, Decomposition methods

1 Introduction

Since the seminal papers of Becker and Tomes (1979, 1986), much effort has been made in the field of economics to understand the persistence of incomes across generations.[1] And while great strides have been made in estimating income persistence (see Solon, 1999 for a review and Mazumder, 2005 for a recent reassessment), much remains to be understood about the drivers or determinants of this persistence (Black and Devereux, 2011). The literature that has sought to understand the driving mechanisms has mainly been focused on mean effects, such as empirical estimates of the intergenerational elasticity of income (Björklund et al., 2006; Björklund et al., 2012; Blanden et al., 2007; Bowles and Gintis, 2002; Cardak et al., 2013; Lefgren et al., 2012; Liu and Zeng, 2009; Mayer and Lopoo, 2008; Richey and Rosburg, 2015; Sacerdote, 2002; Shea, 2000). Alternative measures, such as transition matrices and related summary indices, provide a more 'complete' picture of intergenerational persistence by looking across the entire income distribution (Bhattacharya and Mazumder, 2011; Black and Devereux, 2011; Jäntti et al., 2006). However, the literature on these measures has focused solely on point estimation rather than understanding the driving mechanisms behind these estimates. The reason for the confinement of the transition matrix and related literature to point estimates is not, we believe, a lack of interest, but rather a lack of framework to explore further.

In this paper, we seek to fill this gap by proposing a framework to 'decompose' transition matrices and related indices. This framework provides a path to improve our understanding of the causal forces behind intergenerational income persistence. While our method is not a decomposition in the traditional sense, it builds directly on the decomposition literature. Specifically, we start with a benchmark transition matrix to which our actual (empirically estimated) transition matrix is compared. We then explain differences between the actual and benchmark matrices by identifying the portion attributable to differing characteristics between children from different households (a composition effect) and the portion attributable to differing returns to characteristics between children from different households (a structure effect). We also incorporate a detailed decomposition on the composition effect, based on copula theory (Rothe 2015), which assigns effects to specific variables (e.g., education, experience) and their interactions.

Our method begins by recognizing that transition matrices are representations of multiple conditional distributions (i.e., distributions of children's incomes conditional on parental income grouping). Furthermore, we define our benchmark matrix as a representation of multiple, but identical, 'baseline' distributions. There-

[1] Of course, interest in intergenerational income persistence predates these papers, especially in the field of sociology (see Blau and Duncan 1967). However, these publications and the formal framework within, initiated a large body of research within the field of economics.

fore, understanding the actual-benchmark differences can be characterized by understanding the differences in these underlying distributions. As a result, our decomposition is based on an array of counterfactual experiments that ask how outcomes for children from different households would change if we varied certain characteristics and/or wage structures. The counterfactuals are used to recast transition matrices and to attribute portions of the actual-benchmark differences to specific effects.

A better understanding of the underlying determinants of mobility will provide a better understanding of the likely effects of different policy interventions. For example, if the main driving component is a composition effect (i.e., differences in characteristics), then interventions aimed at promoting skill formation for children from disadvantaged households may prove effective at increasing mobility. However, if the main driving component is a structure effect, these types of interventions are unlikely to yield much mobility improvements. Further, by decomposing transition matrices, our approach provides more information on how these effects may differ across the distribution. For example, we may find that a particular characteristic, say self-esteem, has relatively large effects on a particular portion of the distribution but smaller effects on other portions of the distribution. This would be valuable information when evaluating potential policy programs that aim to improve non-cognitive ability in children.

We apply our method to the intergenerational income mobility of white men surveyed by the 1979 National Longitudinal Survey of Youth (NLSY79). We base our wage structure on an extended “Mincer equation” that includes education, experience, experience squared, and cognitive and non-cognitive ability measures. To address endogeneity concerns regarding education decisions in our detailed decomposition, we reestimate our decomposition using an instrumental variables approach. The IV approach allows us to estimate causal effects as well as investigate the role of unobserved ability. While the general method we propose extends to any function of transition matrices, we focus our application on four summary indices and two sets of specific entries in the transition matrix.

To preview our empirical results, we find that the overall composition-structure split varies between 35-65 to 65-35, depending on the index or matrix entry of interest. These results have two key implications. First, they highlight the ability of transition matrices to present asymmetric patterns, not only in mobility, but in explaining the driving forces behind mobility. Second, the variation in decomposition results show the flexibility and importance in choosing a summary index that appropriately reflects how one wishes to measure mobility. In the baseline detailed decomposition (i.e., no IV), education generally plays the largest role in the composition effect, although cognitive ability and experience play significant roles as well. Just as in the overall decomposition, the detailed decomposition exhibits substantial heterogeneity by index and matrix entry of interest. Reestimation with an IV approach results in a reduced role for education and, in general, an increased role of cognitive ability. And while the point estimates of the role of unobserved ability are at times at a magnitudes similar to or larger than that of the causal effect of education, these effects are not statistically significant.

2 Methods

2.1 Transition Matrices and Mobility Indices

Transition matrices are a popular way to document intergenerational mobility.[2] A transition matrix depicts the probability a child will have adult earnings (Y^c) in a specific income bracket given his/her parents' income (Y^p) was in a certain income bracket. Transition matrices have several advantages over mobility measures that focus only on the average degree of transition (e.g., intergenerational elasticities or correlation coefficients). Transition matrices provide more information about mobility across the *entire* distribution, allow for asymmetric patterns across the distribution (e.g., more mobility at top than bottom), and allow subgroup comparisons across the entire distribution (Black and Devereux, 2011; Jäntti et al., 2006).

Let there be m income brackets (defined as equal percentile groups) with boundaries $0 < \zeta_1 < \zeta_2 < ... < \zeta_{m-1} < \infty$ for the parental distribution and $0 < \xi_1 < \xi_2 < ... < \xi_{m-1} < \infty$ for the children's distribution. A transition matrix (P) is a $m \times m$ matrix with elements p_{ij} that represent the conditional probability that a child is in income bracket j given his/her parents were in income bracket i or

$$p_{ij} = \frac{Pr(\zeta_{i-1} \leq Y^p < \zeta_i \text{ and } \xi_{j-1} \leq Y^c < \xi_j)}{Pr(\zeta_{i-1} \leq Y^p < \zeta_i)}$$

where $\sum_{j=1}^{m} p_{ij} = 1$. Furthermore, it is common to let π_i denote the probability that a child's parents were in income bracket i [i.e., $Pr(\zeta_{i-1} \leq Y^p < \zeta_i)$].

Table 1 provides a quartile transition matrix derived for white males from the NLSY79 data (a full discussion

[2]There is a distinction between 'size' transition matrices and 'quantile' transition matrices. The former defines boundaries of the matrix exogenously - for example, every \$10,000 - while the latter defines them endogenously - for example, every 25th percentile. We focus solely on quantile transition matrices. Interested readers can refer to Formby et al. (2004) and references within for an indepth discussion of size and quantile matrices.

of the data used is given in Section 3). Entries in Table 1 report the probability a son will be in each quartile of the income distribution conditional on parents having been in a specific quartile of the income distribution. The top left entry (0.403) tells us that for parents in the bottom quartile of the income distribution, their son has a 40% chance of ending up in the bottom quartile of his generation's income distribution. Similarly, the top right entry (0.119) tells us that the same son has only about a 12% chance of ending up in the top quartile of the income distribution. Comparatively, a son whose parents were in the top quartile of the income distribution has a 14% and 44% chance of ending up in the bottom and top quartile, respectively, of their generation's income distribution.

[Table 1 about here]

The intergenerational mobility portrayed in a transition matrix can be summarized through mobility indices, $M(P)$, which map the transition matrix P into a scalar value (Formby et al., 2004). These mappings can be useful in ranking transition matrices according to various welfare criterion (e.g. Kanbur and Stiglitz, 2015). A preferential, although not required, characteristic of a mobility index is that it be bounded between 0 and 1 and satisfy the condition that $0 \equiv M(I_m) < M(P) < M(PM) \leq 1$ where I_m is the identity matrix (i.e., zero mobility) and PM is the "perfect mobility" matrix (Jäntti et al., 2006). However, there is not a universally accepted specification for the "perfect mobility" matrix. Instead, a commonly used benchmark matrix is one where all outcomes are independent of origin and destination – i.e., $p_{ij} = 1/m$ for all i, j (Jäntti et al., 2006); we will refer to this as the 'independent' mobility matrix.[3]

Because there is not a consensus on how mobility should be measured, a number of mobility indices have been proposed; we consider four of these indices.[4] The first index, proposed by Prais (1955) and Shorrocks (1978), is based on the trace of the mobility matrix:

$$M_1 = \frac{(m - \sum_{i=1}^{m} p_{ii})}{m-1}. \tag{1}$$

M_1 is the normalized distance from the identity matrix and is equal to 0 for I_m and 1 for the independent matrix[5] (Formby et al, 2004). A criticism of M_1 is that it ignores the off-diagonal elements of P (Maasoumi,

[3] We use the independent mobility matrix as a standard of comparison only; it is not meant to be interpreted as a measure of "perfect mobility" or a policy goal.

[4] See Maasoumi (1998) or Checchi et al. (1999) for relatively comprehensive overviews of summary mobility measures and Formby et al. (2004) for a discussion of their asymptotic properties.

[5] M_1 can be interpreted as the (scaled) average probability that a child will be in a different bracket than his/her parents. By dividing by $(m-1)$ rather than m, the derivation for M_1 in equation (1) takes the value of 1 for the independent matrix and can take values above 1. For M_1 to represent the probability that the child is in a different bracket than his/her parents, the denominator would have to be m.

1998).

The second index is based on the second largest eigenvalue (λ_2) of the mobility matrix (Sommers and Conlisk, 1979):

$$M_2 = 1 - |\lambda_2|. \tag{2}$$

The second largest eigenvalue (λ_2) is the correlation coefficient between parental brackets and child brackets (Formby et al., 2004). Therefore, M_2 provides a measure of the speed at which a child escapes their parents' income bracket (Maasoumi, 1998; Theil, 1972). M_2 is bounded in the unit interval and takes a value of 1 for the independent mobility matrix (Jäntti et al., 2006).

The third index, proposed by Bartholomew (1982), is based on the expected number of income brackets crossed:

$$M_3 = \frac{1}{m-1} \sum_{i=1}^{m} \sum_{j=1}^{m} \pi_i p_{ij} |i - j|. \tag{3}$$

The fourth and final index is defined as the scaled Frobenius distance between a transition matrix and the independent matrix:

$$M_4 = 1 - \frac{\sqrt{\sum_{j=1}^{m} \sum_{i=1}^{m} (p_{ij} - \frac{1}{m})^2}}{m(1 - \frac{1}{m})^2}. \tag{4}$$

M_4 takes the value of 1 for the independent mobility matrix and is bounded below by 0. Table 2 reports the four index values for the independent mobility matrix and the NLSY79 mobility matrix for white men reported in Table 1.

[**Table 2 about here**]

Most of the existing literature on intergenerational mobility use transition matrixes or mobility indices as the final output – that is, a simple way to summarize empirical estimates of economic mobility (e.g., Checchi et al., 1999; Corak and Heisz, 1999; Dearden et al., 1997; Jäntti et. al., 2006; Peters, 1992). What we wish to do is take this a step further and understand the potential driving forces behind the transition matrix and summary indices. To evaluate such forces, we define the independent matrix as our baseline matrix from which we compare our actual (empirically estimated) matrix. For example, consider the trace index (M_1) of the NLSY79 mobility matrix for white men (0.84) and its difference from the M_1 value for the independent matrix (i.e., 1 - 0.86 = 0.14). Alternatively, we can compare specific matrix entries. For example, consider the probability that a son ends up in the top quartile given his parents were also in the top quartile (bottom right entry - p_{44}). The difference between the independent transition matrix and our empirical matrix is -0.19 (0.25 - 0.44); that is, given parents from the top quartile, men in our sample are 19% more likely to end up in the top quartile than if mobility were reflected by the independent mobility matrix.

The questions we seek to answer are how much of these differences are due to differences in observed characteristics between children from different households (referred to as the "composition" effect) and how much is due to differences in returns to observable and unobservable characteristics for these children (referred to as the "structure" effect). While understanding the roles of these components is important for understanding what drives mobility, we note again that our benchmark matrix is not a policy goal but simply a standard for comparison; even in a perfect meritocracy there will likely persist a 'mobility gap' due to differences in ability and other characteristics driven by a genetic component. Rather, in this paper, our objective is to provide a general framework through which one can decompose differences between any baseline matrix of interest and an empirically estimated transition matrix.

The indices presented above and the underlying transition matrix are functions of conditional distributions of incomes (i.e., distributions of children's incomes conditional on parental income). In order to answer our questions of interest, we must conduct counterfactual experiments on the conditional distributions and then inquire what the counterfactual conditions imply for the matrix and indices. In this vein, we connect the literature on economic mobility with the literature on decomposition methods.

2.2 Aggregate Decomposition

The seminal papers by Oaxaca (1973) and Blinder (1973) first introduced techniques for decompositions, and since then, the Oaxaca-Blinder decomposition has become a staple in labor economics. The original Oaxaca-Blinder decomposition was designed to decompose mean differences in a variable of interest between two groups; for example, what portion of the gender wage gap is due to different levels of education and experience (composition effect) and what portion is due to different market returns to these traits (structure effect – interpreted as a measure of market discrimination). However, in many cases, researchers are interested in differences beyond the mean. Recent developments have led to tools that allow such parallel decompositions (see Fortin et al. 2011 for a review).

We begin with several non-overlapping groups $g \in \{1, 2, ..., N\}$, such as children from households in the four quartiles of the income distribution. For each child in group $g \in \{1, 2, ..., N\}$, we observe the outcome Y^g (child's adult log income) and a d-dimensional vector of observables X^g with distributions F_Y^g and F_X^g, respectively, and conditional CDF $F_{Y|X}^g$ which is implicitly defined by a wage structure $Y^g = w_g(X, \epsilon)$.

Let $\nu(F_Y^1, F_Y^2, ..., F_Y^N)$ be a function of a transition matrix based on the distributions $F_Y^1, F_Y^2, ...,$ and F_Y^N; for example, a summary index (M_i) or a specific entry in the transition matrix (p_{ij}). In order to implement the decomposition, we have to specify a distribution which will underlie our independent matrix - a 'baseline' group; for ease of exposition, we let the last group (N) be our baseline group and therefore define the "independent mobility" value as $\nu(F_Y^N, F_Y^N, ..., F_Y^N)$.[6] For example, consider our case where groups are based on the four quartiles of parental income distribution. Then N represents children from households in the top 25% of the income distribution and the independent mobility value is derived from the case where all groups have the same income distribution as children from households in the top 25% of the income distribution. Given the independent mobility value, we define the overall 'mobility gap' as follows:

$$\Delta_O^\nu = \nu(F_Y^N, F_Y^N, ..., F_Y^N) - \nu(F_Y^1, F_Y^2, ..., F_Y^N).$$

To identify how much of this mobility gap is due to structural or compositional effects, we rely on the use of counterfactual distributions. Specifically, for groups $g \neq g'$, define the following counterfactual distribution:

$$F_Y^{g|g'}(y) = \int F_{Y|X}^g(y, x) dF_X^{g'}(x). \tag{5}$$

This is a counterfactual distribution of incomes based on individuals with characteristics as those from group g' and a wage structure (ie. returns to those characteristics) as those from group g. The structure effect, or the portion explained by differences in returns to characteristics between the baseline group and all others, is defined as

$$\Delta_S^\nu = \nu(F_Y^N, F_Y^N, ..., F_Y^N) - \nu(F_Y^{1|N}, F_Y^{2|N}, ..., F_Y^{N|N}). \tag{6}$$

The composition effect, or the portion explained by differences in observed characteristics between the

[6]The independent mobility value is derived from the case where all groups have the same distribution as those in baseline group N. While the general approach is unchanged, interpretation will vary based on the selected baseline group. For example, an alternative choice of baseline group could be the full population (i.e., comparison to population distribution).

baseline group and all others, is defined as[7]:

$$\Delta_X^{\nu} = \nu(F_Y^{1|N}, F_Y^{2|N}, ..., F_Y^{N|N}) - \nu(F_Y^1, F_Y^2, ..., F_Y^N). \tag{7}$$

Two key assumptions are needed to separately identify the structure and composition effect: (1) common support and (2) ignorability. With these assumptions, the 'aggregate' decomposition, $\Delta_O^v = \Delta_S^v + \Delta_X^v$, is well identified. Here, we briefly summarize these assumptions and how they apply to our analysis.[8]

Identifying Assumption 1 - Common support: *Let the support of all wage setting factors* $[X, \epsilon]$ *be* $\mathcal{X} \times \varepsilon$*. Define* D_g *as a discrete variable denoting membership in group* g*. For all* $[x, e]$ *in* $\mathcal{X} \times \varepsilon, 0 < Pr[D = g|X = x, \epsilon = e] < 1$.

The common support assumption requires that there be no set of observables/unobservables that uniquely define membership (or non-membership) in a parental income quartile. In other words, parental income cannot be a direct component of the child's structural wage setting function. However, parental income can be correlated with characteristics that directly affect earnings, e.g. ability (observable as AFQT) or motivation (unobserved), as long as no single value of these variables uniquely defines group membership.

Identifying Assumption 2 - Ignorability: *For* $g = 1, 2, ..., N$*, let* (D_g, X, ϵ) *have a joint distribution. For all* x *in* $\mathcal{X}$*:* ϵ *is independent of* D_g *given* $X = x$*.*

The ignorability assumption assures that, conditional on observables, the distribution of unobservables are not dependent on membership in a quartile of parental income distribution. So while the group from wealthier families may have more 'motivated' children (unobserved), the distribution of motivation is identical across groups when conditioned on observables. This assumption does not require the groups to be compensated equally for unobservables – compensation differences are captured in the structure effect.

Decomposing the overall differences into separate structure and composition effects will provide interesting insight into the sources of mobility differences. However, we are also interested in identifying which characteristics (or covariates) drive the composition effect; that is, what portion of the composition effect can be

[7] The sequence of the decomposition, which defines what counterfactuals are being examined, is parallel to the choice of reference or base group in the standard decomposition literature. We choose this sequence because it prices characteristics in the composition effect at prices individuals actually face. And since we will only perform a detailed decomposition on the composition effect, we feel this is more appropriate.

[8] Assumptions presented here are from Fortin et al. (2011); see Fortin et al. (2011) for a more comprehensive discussion of the technical assumptions (e.g. no general equilibrium effects, invariance of the structural form, etc.).

assigned to specific variables of interest? To investigate this question, we employ a 'detailed' decomposition of the composition effect in equation 7.[9] Several decompositions have been proposed (see DiNardo et al. 1996; Machado and Mata 2005; Chernozhukov et al. 2013) but these are in general 'path dependent' (i.e., they depend on the order of the covariate inclusion in the decomposition). We apply a non-path dependent decomposition based on copula theory introduced by Rothe (2015). In what follows, we briefly outline our specific use of his procedure and intuition; however, interested readers should refer to Rothe (2015) for a complete description of the procedure in a more traditional decomposition setting.

2.3 Detailed Decomposition

Rothe (2015) notes that the detailed decomposition cannot be additively decomposed into marginal components solely due to separate variables but must contain 'interaction' terms. To see how our decomposition proceeds, we must refer to the concept of the copula, which was first introduced by Sklar (1959).[10] Sklar's theorem states that any d-dimensional distribution F_X^g can be decomposed into two parts - the d marginal distributions $F_{X_i}^g$ for the random variables $(X_1, X_2, ..., X_d)$ and the copula function C^g which captures the dependence structure of the distribution:

$$F_X^g(x) \equiv C^g(F_{X_1}^g(x_1), ..., F_{X_d}^g(x_d)) \text{ for } g \in \{1, 2, ..., N\}.$$

In other words, the dependence structure of the covariates in F_X^g – captured by the copula – can be separated from the individual marginal distributions - the $F_{X_i}^g$'s. Also note that the copula takes as its arguments $F_{X_i}^g(x_i)$'s, which are uniform random variables, rather than the x's themselves.

Now consider a d-dimensional product set $\{1, 2, ..., N\}^d$ where an element of the product set (denoted with bold font) represents a set of d covariate marginal distributions. For example, if $\mathbf{k} = (N, N, 1, 1, ..., 1)$, this would denote the set of covariate marginal distributions where the distributions for covariates one and two are equal to that of group N and the remaining covariate distributions are equal to that of group 1. In this vein, define $\mathbf{1} = (1, 1, ..., 1)$, $\mathbf{2} = (2, 2, ..., 2)$, ..., and $\mathbf{N} = (N, N, ..., N)$. Using this notation, we extend the counterfactual setup from equation (5). Let the counterfactual distribution of Y, where the wage structure is of group g, the copula is of group g', and the marginal distribution of the l^{th} covariate is equal to the

[9]Performing a detailed decomposition on the structure effect has conceptual problem regarding the choice of omitted/baseline group for covariates of interest even for simple linear models (see Fortin et al. 2011 for a discussion) and therefore we do not pursue such a decomposition.

[10]See Rothe (2015) for a specific discussion or Nelsen (2006) for a general handling of copula theory.

group denoted in $\mathbf{k}_l$, be:

$$F_Y^{g|g',\mathbf{k}}(y) = \int F_{Y|X}^{g}(y,x)dF_X^{g',\mathbf{k}}(x) \tag{8}$$

with

$$F_X^{g',\mathbf{k}}(x) \equiv C^{g'}(F_{X_1}^{\mathbf{k}_1}(x_1), ..., F_{X_d}^{\mathbf{k}_d}(x_d)).$$

Now the composition effect can be decomposed into a *dependence effect* (Δ_D^ν), which is due to differences between groups in their copulas, and a *total marginal effect* (Δ_M^ν), which is due to differences in the marginal distributions between groups:

$$\Delta_X^\nu = \Delta_D^\nu + \Delta_M^\nu$$

where

$$\begin{aligned}
\Delta_D^\nu &= \nu(F_Y^{1|N,\mathbf{N}}, F_Y^{2|N,\mathbf{N}}, ..., F_Y^{N|N,\mathbf{N}}) - \nu(F_Y^{1|1,\mathbf{N}}, F_Y^{2|2,\mathbf{N}}, ..., F_Y^{N|N,\mathbf{N}}) \\
\Delta_M^\nu &= \nu(F_Y^{1|1,\mathbf{N}}, F_Y^{2|2,\mathbf{N}}, ..., F_Y^{N|N,\mathbf{N}}) - \nu(F_Y^1, F_Y^2, ..., F_Y^N).
\end{aligned}$$

Next, we further decompose the total marginal effect (Δ_M^ν) into portions attributable to specific covariates (and their interactions). With d covariates and N groups, there are several potential counterfactual distributions for Y that can be derived from equation (8), not all of which are of direct interest for our decomposition. Recall that the aggregate decomposition, and thus the composition effect, was derived with group N as the baseline group. Therefore, to decompose the total marginal effect, we need to consider all potential counterfactuals for groups $1, 2, ..., N-1$ relative to group N. That is, for all possible covariate combinations, we need to evaluate the counterfactual where all $N-1$ groups take group N's marginal distributions for the covariates in that combination but their own marginal distributions for the remaining covariates. This requires additional notation. Let $\tilde{\mathbf{k}}$ (or any other bold face letter with a tilde) be an element of the d-dimensional product set $\{0,1\}^d$ and let it represent the set of $N-1$ elements $\mathbf{k}^i$ with $i = \{1, 2, ...N-1\}$ where $\mathbf{k}_l^i = N$ if $\tilde{\mathbf{k}}_l = 1$ and $\mathbf{k}_l^i = i$ if $\tilde{\mathbf{k}}_l = 0$. In addition, let $\tilde{\mathbf{e}}^l$ be the l^{th} unit vector such that all entries equal 0 except the l^{th} which equals 1. Then, given a feature ν of interest, we can denote:

$$\beta^\nu(\tilde{\mathbf{k}}) = \nu(F_Y^{1|1,\mathbf{k}^1}, F_Y^{2|2,\mathbf{k}^2}, ..., F_Y^{N-1|N-1,\mathbf{k}^{N-1}}, F_Y^N) - \nu(F_Y^1, F_Y^2, ..., F_Y^N).$$

This last equation should be interpreted as the effect of a counterfactual alteration where, for elements with $\tilde{\mathbf{k}}_l = 1$, the marginal distributions of all groups $1, 2, ..., N-1$ are altered to the marginal distributions of group N while holding all else constant (including the copula). For example, if $\tilde{\mathbf{k}} = (1, 1, 0, 0.., 0)$, then $\beta^v(\tilde{\mathbf{k}})$

represents the effect of changing the marginal distributions of the first two variables to those of group N while holding all else constant.

Now, letting $|\tilde{\mathbf{k}}| = \sum_{l=1}^{d} \tilde{\mathbf{k}}_l$, define

$$\Delta_M^v(\tilde{\mathbf{k}}) = \beta^v(\tilde{\mathbf{k}}) + \sum_{1 \le |\tilde{\mathbf{m}}| \le |\tilde{\mathbf{k}}|-1} (-1)^{|\tilde{\mathbf{k}}|-|\tilde{\mathbf{m}}|} \beta^v(\tilde{\mathbf{k}}).$$

Thus, $\Delta_M^v(\tilde{\mathbf{e}}^l) = \beta^v(\tilde{\mathbf{e}}^l)$. And note if $|\tilde{\mathbf{k}}| = 2$ then $\Delta_M^v(\tilde{\mathbf{k}})$ is equal to the counterfactual $\beta^v(\tilde{\mathbf{k}})$ with the individual 'parts' of the effect from $\beta^v(\tilde{\mathbf{m}})$ where $|\tilde{\mathbf{m}}| = 1$ removed. Thus, the 'full' counterfactual of changing, for instance, the distribution of the first two covariates to that of group N for all groups $1, 2, ..., N-1$ is split up into two 'direct effects' [$\beta^v(\tilde{\mathbf{e}}^l) = \Delta_M^v(\tilde{\mathbf{e}}^l)$ where $|\tilde{\mathbf{e}}^l| = 1$ and $l = 1, 2$] and one 'interaction effect' [$\Delta_M^v(\tilde{\mathbf{k}})$ where $|\tilde{\mathbf{k}}| = 2$, $\tilde{\mathbf{k}} = (1, 1, 0, ..., 0)$].

Therefore, the detailed decomposition (of the composition effect) contains three parts: (1) "direct contributions" from each covariate due to differences between groups in the marginal distribution of these variables [$\Delta_M^v(\tilde{\mathbf{e}}^l)$], (2) several "interaction effects" which are due to interactions between the marginals ($\Delta_M^v(\tilde{\mathbf{k}})$ where $|\tilde{\mathbf{k}}| > 1$), and (3) a "dependence" effect due to different dependence structures among the covariates between the groups (Δ_D^v). Thus, we have:

$$\Delta_X^v = \sum_{l=1}^{d} \Delta_M^v(\tilde{\mathbf{e}}^l) + \sum_{|\tilde{\mathbf{k}}|>1} \Delta_M^v(\tilde{\mathbf{k}}) + \Delta_D^v \tag{9}$$

where the interaction effects are summed over all possible interactions. For example, the direct contribution term may be due to groups having a different distribution of ability levels or non-cognitive traits, the interaction effects may be the additional effect due to groups having different distributions for *both* ability level *and* non-cognitive traits, and the dependence term is due to the groups having different dependence structures between ability and non-cognitive traits.

In order to identify this detailed decomposition of the composition effect we must assume independence of the error terms.

Identifying Assumption 3 - Independence: *Given the wage structural model $Y^g = w_g(X, \epsilon)$, for every $g \in \{1, 2, ..., N\}$, ϵ is independent of X given D_g.*

The independence assumption requires that the error term is independent of the covariates of interest. This is a stronger assumption than ignorability (Identifying Assumption 2), but is needed to identify the detailed decomposition (the detailed decomposition requires us to be able to attribute causal effects to the parameters of the conditional $CDF_{Y|X}$). Ignorability simply assumes the error term is independent of group assignment conditional on the covariates X, independence goes further and assumes that the error term is independent of the covariates. We will discuss this assumption further within our estimation approach; in particular, for estimation of the detailed decomposition in the presence of endogenous school choice.

2.4 Estimation

The decomposition is based on estimating the relevant counterfactual outcomes discussed above. Thus, estimation rests on estimating the marginal distributions of the covariates for each group, the conditional CDF of $Y|X$ for each group, and the copula function for each group. Once these components are estimated, any counterfactual outcome can be numerically approximated using estimated components for population components (equation 8).

Marginal distributions are simply estimated with the empirical CDF. The conditional CDFs are estimated using the distributional regression approach suggested by Foresi and Peracchi (1995); the foundation for this approach is a Probit model.[11] The conditional CDF is estimated using multiple standard binary choice models where we vary the cut-off across a grid along the outcome space. In other words, we repeatedly estimate $Pr(y_i \leq \tilde{y}|x) = \Phi(x\beta_{\tilde{y}})$ for $\tilde{y} \in \mathcal{Y}$ where X is a vector of covariates (e.g., education, experience, cognitive, and non-cognitive measures) and Y is log income.

Lastly, the copula is estimated with the Gaussian copula model:

$$C_\Sigma(u) = \Phi^d_\Sigma(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)) \tag{10}$$

This model does not impose the joint distribution of X to be Gaussian; rather, it assumes the dependence structure of the transformed variables $u_i = F_i(x_i)$ follows the Gaussian model. For estimation we rely on the maximum pseudo-likelihood approach (Genest et al. 1995).[12]

[11] Koenker and Bassett (1978) propose an alternative approach based on quantile regression. We refer readers interested in the relationship between the alternative approaches to Koenker et al. (2013).

[12] Specifically, we use the R package 'copula' and the methods within (Hofert and Maechler 2011; Hofert et al. 2015; Kojadinovic and Yan 2010a; Yan 2007). Through simulation, Kojadinovic and Yan (2010b) show that the pseudo-likelihood approach performs favorability over several other approaches.

Therefore, our general approach is as follows: (1) estimate the copulas, conditional CDFs and marginal distributions, (2) simulate counterfactual outcomes (equation 8), (3) given these counterfactual outcomes, estimate the counterfactual transition matrix, and (4) use the counterfactual transition matrix to account for the overall empirical-benchmark difference in the transition matrix and related indices.

2.4.1 Instrumenting Education

Education is one of the variables of interest in our empirical analysis, and the endogenous nature of education in earnings equations is well known. To estimate causal relationships in the presence of endogenous education choices, we reestimate the model using an instrumental variables approach. The reestimation applies an intermediate step consisting of an instrumental variable (IV) probit approach (control function) within the estimation procedure of the conditional CDF (Rivers and Vuong, 1988). Specifically, we estimate a 'first stage' (predicting education) and then construct the normalized residual term (the 'control variable'). The control variable is then included in the estimation of the conditional CDF estimation step. Assuming the standard instrument assumptions hold, the causal structure will be consistently estimated.

Furthermore, the 'control' variable is included in the rest of the estimation procedure as a covariate. The control variable is a composite term that captures all things correlated with schooling decisions and adult earnings. However, for ease of discussion, we will follow the terminology of Matinez-Sanchis et al. (2012) and refer to the control variable as a (constructed) measure of unobserved ability. Therefore, the IV approach not only allows us to recover causal parameters but also allows us to investigate the role that unobserved ability plays in the decomposition.

3 Data

The primary source of data for our analysis is the 1979 National Longitudinal Survey of Youth (NLSY79) including the restricted-use geocode data.[13] The NLSY79 is a panel survey of youths aged 14-22 in 1979. It includes a cross-sectional representative survey (n = 6,111), an over sample of minorities and poor whites (n = 5,295), and a sample of military respondents (n = 1,280).[14] We use only the cross-sectional representative

[13] The use of the NLSY79 geocode data is subject to a special agreement with the Bureau of Labor Statistics.

[14] The over sample of military and poor whites were discontinued in 1984 and 1990, respectively

survey.

We limit the sample to white males who reported living with a parent for at least two of the first three years of the survey and with reported parental income for those years.[15] A key variable of interest is parental status based on the parents' (average) income. The outcome of interest is the individual's economic status based on their average reported wage and salary income between 1988 and 1990. All incomes are deflated to 1982-1984 dollars using the CPI. Further, since regional differences in cost-of-living may lead to imperfect and potentially biased measures of real income (Fuchs 2004), we adjust parental and child incomes according to the state-level cost of living indices provided by Berry et al. (2000). The sample is further limited to individuals not enrolled in school over the period of interest and with available Arms Force Qualifying Test (AFQT) scores. With these restrictions, the final sample includes 1,407 individuals born between 1957 and 1964 with a mean age of 27.5 during our outcome years of interest (1988-1990). Table 3 provides summary statistics for the entire sample and by parental income quartile.

[Table 3 about here]

The variables we include in our decomposition are based on an extended Mincer equation. The traditional Mincer equation includes education, experience, and experience squared (Mincer, 1974). We extend this basic model to include other variables that have been related to income determination. In particular, we include a measure of cognitive ability ($AFQT$) and two measures of non-cognitive ability (*Esteem*, *Rotter*).

The NLSY79 does not provide a direct measure of experience. Therefore, we construct a measure of 'full time equivalent' (FTE) years of experience using the weekly array of hours worked for all years the individual was not enrolled in school.[16] One FTE year of experience is assumed to equal 52 weeks times 40 hours (hours worked are top coded to 40). A few older individuals in our sample completed their education prior to the beginning of the survey and therefore were already working during the first round of interviews in 1979. Without information on previous work experience for these individuals, we construct the following 'pre-survey' estimate of FTE years ($FTE_{<79}$) based on age, years of schooling, and FTE years of experience earned in the initial survey year: $FTE_{<79} = (\text{Age}_{79} - \text{Years of Schooling}_{79} - 6) \cdot FTE_{79}$. We then add the pre-survey FTE years to the (observed) survey FTE years.

[15]Parental income is identified through a comparison of total household income and respondent's income. We exclude individuals who lived with a spouse or child during these years.

[16]Our measure of experience is very similar to, but slightly different from, the measure used by Regan and Oaxaca (2009).

The measure of ability used in our analysis is Arms Force Qualifying Test scores (AFQT). Since different individuals took the test at different ages, the measure used is from an equi-percentile mapping used across age groups to create age-consistent scores (Altonji et al., 2012). The use of AFQT scores as a measure of ability warrants a brief discussion. Some argue AFQT scores are proxies for IQ scores while others draw serious doubts to this interpretation (Ashenfelter and Rouse, 2000). While it may be appealing to interpret AFQT scores as a measure of IQ, it is also not entirely clear what IQ scores measure. For example, there have been large gains in IQ scores over time in nearly every country on record (Flynn, 2004). Flynn (2004) argues that these differences are too large to uncautiously equate IQ with 'intelligence.' Therefore, we interpret AFQT scores as some combination of innate ability and accumulated human capital as a youth that is valued in the labor market. However, for ease of expression, we will refer to AFQT scores as our measure of 'cognitive ability.'

We use two measures for non-congitive ability. First, we use information from the Rosenberg Self-Esteem Scale (1965). The Rosenberg Self-Esteem Scale contains 10 statements on self-approval and disapproval; we use a summary measure of the individual's responses to these 10 statements (*Esteem*). Second, we use a summary measure from the Rotter-Locus of Control Scale (*Rotter*) which measures the "extent to which individuals believe they have control over their lives through self-motivation or self-determination (internal control) as opposed to the extent that the environment (that is, chance, fate, luck) controls their lives (external control)" (BLS, 2015).

Finally, educational attainment is measured as years of schooling. The endogenous nature of education in explaining income outcomes is well documented and this measure of educational attainment likely incorporates unobserved characteristics. Therefore, we first estimate the model using our direct measure of educational attainment (years of schooling) and then reestimate the model using an instrumental variables approach. The instruments used in our analysis are based on Kling (2001) and Carneiro et al. (2011): presence of a four year college in the county of residence at age 14, minimum tuition in public 4 year college in the state of residence at age 14, local wage in the county of residence at age 17, and local unemployment in state of residence at age 17.[17]

[17] Carneiro et al. (2011) summarize the main papers that use similar instruments.

4 Results

The general decomposition method we propose can be applied to any function of transition matrices. For brevity, we present decomposition results for the following: (1) the four summary indices introduced in equations (1) - (4) (table 4), (2) transition matrix entries for those coming from households in the top quartile of the earnings distribution (p_{4j} for $j = 1, 2, 3, 4$) (table 5), and (3) transition matrix entries for those coming from households in the bottom quartile of the earnings distribution (p_{1j} for $j = 1, 2, 3, 4$) (table 6). The total decomposition effects reported in tables 4 - 6 will differ slightly from the differences from tables 1 and 2 because Tables 4 - 6 are based on simulated data rather than the empirical data. However, the estimated values are very similar, suggesting that our parametric specification of the conditional CDF and copula fit the data reasonably well.

Our presentation of results is structured around our theoretical discussion of the decomposition method. First, we report the overall difference and its aggregate decomposition into the structure and composition effects. Second, we report the initial step in the detailed decomposition of the composition effect - the dependence and total marginal effects. Then, we report the covariate direct contributions to the marginal effect followed by two-way interaction contributions to the marginal effect. Higher order interaction terms are not reported for brevity.

4.1 Baseline Results

We begin with the index decompositions (table 4). For each index, we decompose the difference between the index value for the independent matrix and the index value for our matrix. The structure-composition split varies among the indices; decomposition differences reflect the fact that each index measures mobility differently. The composition effect accounts for 62% for M_1, about one half for M_2 and M_3, and 40% for M_4. So if one were only concerned about mobility as measured by the trace (M_1), then a majority of the mobility gap is explained by differences in characteristics between children from different households. However, if one is interested in a more comprehensive distance measure (M_4), then only about 40% of the gap is explained by these differences; this latter value implies that even if all children had characteristics similar to those coming from the upper quartile households, 60% of the mobility gap would still be present. These decomposition differences highlight the role of the choice of index in mobility analysis.

[Table 4 about here]

Turning to the detailed decomposition of the composition effect, the dependence effect is minimal for all four indices and the majority of the composition effect comes from the full marginal effect. Further decomposition of the marginal effect indicates that AFQT accounts for 16-20% of the marginal effect in indices $M_2 - M_4$ but has an (imprecise) zero effect in M_1. Non-cognitive ability measures (*Esteem*, *Rotter*) have no significant contribution to the marginal effects of any index. However, education and experience are significant drivers of the marginal effects in all indices. The contribution from education is near or even larger than the total marginal effect; this implies that negative contributions from other variables offset some of education's contribution. One of those variables is experience. A negative contribution from experience indicates that if all children had experience as those from the top quartile, the mobility gap would widen. This result is not surprising since individuals from the top quartile tend to have more education and less experience (see Table 3) and experience is expected to have a positive effect on earnings. The only interaction term of statistical significance is the schooling-experience term for the M_1 index. The interaction effect together with the direct effects indicate that the mobility gap would be reduced by 6.26 points if all children had education *and* experience levels as those from the top quartile of households: 11.08 (direct effect from education) - 2.04 (direct effect from experience) - 2.78 (interaction effect).

Next, we focus on children from households in the top quartile of the parental earnings distribution (p_{4j} for $j = 1, 2, 3, 4$) (table 5). In this case, the total difference represents the difference between the entries in the bottom row of the independent matrix (0.25 for all cells) and the bottom row of Table 1. The total, structure and composition effects are all statistically significant for the first, second and fourth quartiles but not the third quartile. The composition effect explains around 37% of the mobility gap for the bottom and second quartile and 46% for the top quartile. These results indicate that differences in the distribution of characteristics explain less than half of the large presence of children from the top households ending up in the top quartile of their earnings distribution, and it only explains about a third of their relative omission from the bottom two quartiles.

[Table 5 about here]

In the detailed decomposition, the dependence structure plays a statistically significant role for the first and second quartile and accounts for 22% and 33%, respectfully; the total marginal effect is statistically significant for the first, second and fourth quartiles. Looking at the direct contributions to the total marginal effect, AFQT has a statistically significant effect at the first and third quartile, accounting for about 55% and 100% of the total marginal effect, respectfully. As in the index decompositions, non-cognitive ability measures do not provide any statistically significant contribution to the marginal effects. Schooling and experience exhibit effects of similar magnitude but opposing directions on the first quartile at 3.1 and -3.6 points; the interaction of schooling and experience contributes an additional (positive) 2.4 points. So if everyone had education levels as those in the top quartile the first quartile gap would shrink, if everyone had the same level of experience as the top quartile the gap would widen, and if everyone had the same level of both education and experience as the top quartile the gap would shrink by 1.93 points (3.10 - 3.55 + 2.37). Education also plays a large role in the second quartile contributing 5 points, while experience exhibits a statistically insignificant positive direct effect; however, the interaction term of education and experience exhibits a strong negative effect at -3.78 points, therefore offsetting some of education's direct contribution. For the fourth quartile, education exhibits a negative effect at -11.14 points, while experience contributes a small positive effect at 1.63 points. We also find a negative AFQT-education interaction term in the fourth quartile at -1.99, which would further narrow the mobility gap beyond just the direct effects of AFQT or schooling and imply some positive interaction in the top quartile.

Finally, we consider children from households in the bottom quartile of the parental earnings distribution (p_{1j} for $j = 1, 2, 3, 4$) (table 6). The composition effect is estimated at 62% for the bottom quartile and 45% for the top quartile (aggregate decomposition components are statistically insignificant for the second and third quartile). Therefore, differences in characteristics explain 62% of the overabundance of children from bottom quartile homes that end up in the bottom quartile of their generation's earnings distribution, and it explains 45% of their relative absence in the upper quartile.

[Table 6 about here]

Interestingly, while the composition effect is not statistically significant for the second quartile, it is the only quartile where the dependence effect is statistically significant; the dependence effect accounts for a positive 1 point of the negative 2 point total difference. So if everyone had the same dependence structure of covariates as those from the top quartile of households, there would be 1% more children from the lowest quartile in the second quartile of adult earnings. Also, while the composition effect is not statistically significant for the third quartile, it is the only quartile where the AFQT direct marginal effect is statistically significant; AFQT accounts for about 4 points out of the 5.2 point difference. So if everyone had the same distribution of

AFQT scores as those from the upper quartile households, the gap at the third quartile of adult earnings for children from the lowest quartile households would reduce to only 1%. In the first and fourth quartiles, which have significant composition effects, education plays the largest role accounting for -8.3 points and 5 points, respectfully (approximately 88% of the marginal effect in both cases). Experience provides a statistically significant (positive) contribution to the marginal effect only in the first quartile; if all children had (the lower) levels of experience of those from the top most households, the overabundance of children from bottom quartile households that end up in the bottom quartile would increase by almost 3%. None of the interaction terms appear to play a significant role in this decomposition.

4.2 IV Results

Since concerns regarding endogeneity of education may invalidate the causal interpretation of our estimated conditional CDF and therefore the interpretation of our detailed decomposition, we reestimate our decompositions using an IV approach. IV results are provided in Tables 7 - 9. Recall, however, that the baseline (non-IV) total decomposition results (i.e., structure and composition effects) remain valid even in the presence of endogeneity (see Section 2). Minor discrepancies in the IV and non-IV total decomposition results are due to finite sample effects in the construction of our control variable.

In the detailed decomposition for the summary indices (Table 7), we see several changes from the baseline results. The direct effect of AFQT is much larger for all indices (almost twice as large), but only significant for M_4. With the exception of M_1, the direct effect of education is also smaller. The most notable decline in the education effect is in M_4 where the direct effect is less than half of the baseline effect and no longer statistically significant. Experience has similar effects in the baseline and IV results, although the direct effect of experience on M_1 is about 50% larger in the IV model. Unobserved ability, captured in the control variable, has sizable direct effects for M_2 and M_3, but the effects are not statistically significant.

[Table 7 about here]

Table 8 provides IV results for children from households in the top quartile of the parental earnings distribution (p_{4j} for $j = 1, 2, 3, 4$). Again we find, in general, a much larger direct effect for AFQT. For the fourth quartile, the effect increases nearly eight fold from -0.53 to -3.91 and is now statistically significant. We also see a large drop in the direct effect of education at the first, second and fourth quartiles; the effect of education is slightly increased at the third quartile but remains statistically insignificant. Interestingly, the direct effects of unobserved ability at the first and second quartiles are substantially larger than the direct effect of education, but again, these effects are not statistically significant. We do, however, find a significant interaction effect between unobserved ability and experience on the second quartile with a point effect of -3.5. In addition, a few interaction terms increase in magnitude and become statistically significant: the AFQT-experience effect on the second quartile and the education-experience effect on the fourth quartile.

[**Table 8 about here**]

Finally, Table 9 provides IV results for children from households in the bottom quartile of the parental earnings distribution (p_{1j} for $j = 1, 2, 3, 4$). The changes in the direct effect of AFQT are mixed with reduced effects in the first and third quartiles and increased effects in the second and fourth quartiles. Also, quite different than the index and top quartile results, the direct effect of education appears much larger in the IV model for the first and fourth quartile but only statistically significant for the first quartile. Again, point estimates of the direct effect of unobserved ability are quite large for all quartiles, at times larger than the direct effects of experience or education, but these effects are not precisely measured and not statistically different than zero.

[**Table 9 about here**]

5 Conclusion

Economists have long been interested in intergenerational mobility. Over the past few decades, great gains have been made in estimating mobility, both in mean effects (e.g., intergenerational elasticity of income) and through alternative measurements that provide a more complete picture of mobility (e.g., transition matrices). As a result, interest has started to turn from estimating more accurate point estimates to developing a better understanding of the forces that drive income persistence. While much work has been done to understand the driving mechanisms of mean effects, the lack of an appropriate framework has limited similar work on transition matrices and related indices. In this paper, we use recent advances in the decomposition literature to fill this gap.

We introduced a method to ‘decompose’ transition matrices and related indices. Given a benchmark matrix of interest, the method decomposes the difference between the benchmark matrix and the empirical matrix into portions attributable to differing characteristics between children from different economically advantaged households (a composition effect) and differing returns to these characteristics (a structure effect). Our approach also includes a detailed decomposition of the composition effect based on copula theory. While the method we presented is not a decomposition in the traditional sense, it draws directly on the decomposition literature; the procedure entails simulating several simultaneous counterfactuals which are then recast into counterfactual transition matrices and from which we can identify decomposition effects.

We illustrate our method using data on white men from the 1979 NLSY. We base our decomposition on an extended Mincer equation that includes education, experience, cognitive and non-cognitive measures. Moreover, to address endogeneity concerns regarding schooling decisions, we reestimate our decomposition with an IV method that allows us to recover causal effects as well as investigate the role of unobserved ability. We apply our method to specific transition matrix entries and four summary indices. We find that the relative importance of the structure versus composition effects varies substantially across the measures of interest; these findings highlight the importance of being able to understand the driving forces of mobility beyond mean effects. Similarly, the importance of specific characteristics (e.g., education, experience, cognitive, non-cognitive measures) exhibits substantial heterogeneity across the various measures. And while education plays an overwhelming role in our baseline decomposition, this role is substantially diminished in our IV approach with ability playing a more important role.

References

Altonji, Joseph, Prashant Bharadwaj and Fabian Lange (2012) "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," *Journal of Labor Economics*, 4, pp. 783-828.

Ashenfelter, Orley, and Cecilia Rouse (2000) "Schooling, Intelligence, and Income in America." Chapter 5 in Denneth Arrow, Samuel Bowles, and Steven Durlauf (eds.), Meritocracy and Economic Inequality. Princeton: Princeton University Press.

Bartholomew, D. (1982) *Stochastic Models for Social Processes*, Third Edition, Wiley: London.

Becker, Gary and Nigel Tomes (1979) "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87, pp. 1153-1189.

Becker, Gary and Nigel Tomes (1986) "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4,(3), pp. S1-S39.

Berry, William, Richard Fording and Russell Hanson (2000) "An Annual Cost of Living Index for the American States, 1960-1995," *The Journal of Politics*, 62(2), 550-567.

Bhattacharya, Debopam and Bhashkar Mazumder (2011) "A Nonparametric Analysis of Black-White Differences in Intergenerational Income Mobility in the United States," *Quantitative Economics*, 43(1), pp. 139-172.

Björklund, A., Lindahl, M., Plug, E., 2006 "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data", *Quarterly Journal of Economics*, 121(3) 999-1028.

Björklund, A., Roine, J., Waldenstrm, D., 2012 "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?," *Journal of Public Economics* 96, 474-484.

Black, Sandra and Paul Devereux (2011) "Recent Developments in Integenerational Mobility" Chapter 16 in *Handbook of Labor Economics*, Volume 4(B): 1487-1541.

Blanden, Jo, Paul Gregg, and Lindsey Macmillan (2007) "Accounting for Intergenerational Income Persistance: Noncognitive Skills, Ability and Education," *The Economic Journal*, 117, pp. C43-C60.

Blau, Peter and Otis Duncan (1967) *The American Occupational Structure*, John Wiley & Sons: New York.

Blinder, A. (1973) "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 8, 436-455.

Bowles, Samuel and Herbert Gintis (2002) "The Inheritance of Inequality," *Journal of Economic Perspectives*, 16(3), pp. 3-30.

Bureau of Labor Statistics (BLS). (2015) "NLSY79 Appendix 21: Attitudinal Scales," Bureau of Labor Statistics - National Longitudinal Survey of Youth 1979. Accessed June 29, 2015. https://www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/codebook-supplement/nlsy79-appendix-21-attitudinal-scales#references

Cardak, Buly, David Johnston, Vance Martin (2013) "Intergenerational Income Mobility: A New Decomposition of Investment and Endowment Effects," *Labour Economics*, 24, pp. 39-47.

Carneiro, P., J. Heckman, and E. Vytlacil. (2011). "Estimating Marginal Returns to Education," *American Economic Review*, 101(6), 2754-2781.

Checchi, Daniele, Andrea Ichino, and Aldo Rustichini (1999) "More Equal but Less Mobile? Education Financing and Intergenerational Mobility in Italy and in the US," *Journal of Public Economics*, 74, 351-393.

Chernozhukov, V., I. Fernandez-Val, and B. Melly (2013) “Inference on Counterfactual Distributions,” *Econometrica*, 81, 2205-2268.

Corak, Miles and Andrew Heisz (1999) “The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data,” *The Journal of Human Resources*, 34, 504-533.

Dearden, Lorraine, Stephen Machin and Howard Reed (1997) “Intergenerational Mobility in Britain,” *The Economic Journal*, 107, 47-66.

DiNardo, J., N. Fortin, and T. Lemieux (1996) “Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach,” *Econometrica*, 64, 1001-1044.

Flynn, James (2004) “IQ Trends over Time: Intelligence, Race, and Meritocracy.” Chapter 3 in Denneth Arrow, Samuel Bowles, and Steven Durlauf (eds.), Meritocracy and Economic Inequality. Princeton: Princeton University Press.

Foresi S., and Peracchi F. (1995) “The Conditional Distribution of Excess Returns: An Empirical Analysis,” *Journal of the American Statistical Association*, 90: 451-466.

Formby, John, James Smith, Buhong Zheng (2004) “Mobility Measurement, Transition Matrices and Statistical Inference,” *Journal of Econometrics*, 120, 181-205.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo (2011) “Decomposition Methods in Economics”, Chapter 1 in *Handbook of Labor Economics*, Volume 4, Part A: 1 - 102.

Fuchs, Victor. (2004). “Reflections on the Socio-economic Correlates of Health,” *Journal of Health Economics*, 4(23), 653-661.

Genest, C., Ghoudi, K., and Rivest, L.-P. (1995) “A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions”, *Biometrika* 82, 543-552.

Hofert, Marius, Ivan Kojadinovic, Martin Maechler and Jun Yan (2015). Copula: Multivariate Dependence with Copulas. R package version 0.999-13

Hofert, Marius and Martin Maechler (2011) “Nested Archimedean Copulas Meet R: The Nacopula Package” *Journal of Statistical Software*, 39(9), 1-20.

Kanbur, Ravi and Joseph Stiglitz (2015) “Dynastic Inequality, Mobility and Equality of Opportunity,” Dyson School Working Paper WP 2015-05.

Kling, J. (2001). “Interpreting Instrumental Variables Estimates of the Returns to Schooling”, *Journal of Business & Economics Statistics*, 19(3), 358-364.

Koenker R., and Bassett G. (1978) “Regression Quantiles,” *Econometrica*, 46: 33-50.

Koenker, Roger, Samantha Leorato and Franco Peracchi (2013) “Distributional vs. Quantile Regression,” *Center for Economic and International Studies Working Paper Series*, 11(15), No. 300.

Kojadinovic, Ivan and Jun Yan (2010a) “Modeling Multivariate Distributions with Continuous Margins Using the Copula R Package”, *Journal of Statistical Software*, 34(9), 1-20.

Kojadinovic, Ivan and Jun Yan (2010b) “Comparison of Three Semiparametric Methods for Estimating Dependence Parameters in Copula Models”, *Insurance: Mathematics and Economics*, 47, 52-63.

Jäntti, Markus, Bernt Bratsberg, Knut Røed, Oddbjørn Raaum, Robin Naylor, Eva Österbacka, Anders Björklund and Tor Eriksson (2006) “American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States,” IZA Discussion Paper No. 1938.

Lefgren, Lars, Matthew Lindquist, and David Sims (2012) "Rich Dad, Smart Dad: Decomposing the Intergenerational Transmission of Income," *Journal of Political Economy*, 120, pp. 268-303.

Liu, Haoming and Jinli Zeng (2009) "Genetic Ability and Intergenerational Income Mobibility," *Journal of Population Economics*, 22, pp. 75-95.

Maasoumi, Esfandiar (1998) "On Mobility," Chapter 5 in Ullah, Amand and David Giles (Eds.), Handbook of Applied Economic Statistics. Marcel Dekker: New York, pp. 119-175.

Machado, J. and J. Mata (2005) "Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression," *Journal of Applied Econometrics*, 20, 445-465.

Martinez-Sanchis, E, J. Mora and I. Kandemir (2012) "Counterfactual Distributions of Wages via Quantile Regression with Endogeneity," *Computational Statistics and Data Analysis*, 56, 3212-3229.

Mayer, Susan, and Leonard M. Lopoo (2008) "Government Spending and Intergenerational Mobility," *Journal of Public Economics*, 92, pp. 139-158.

Mazumder, Bhashkar (2005) "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Income Data," *The Review of Economics and Statistics*, 87(2), pp. 235-255.

Mincer, Jacob (1974) *Schooling, Experience and Earnings*, New York: National Bureau of Economic Research.

Oaxaca, R. (1973) "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14, 693-709.

Peters, Elizabeth (1992) "Patterns of Intergenerational Mobility in Income and Earnings," *The Review of Economics and Statistics*, 74, 456-466.

Prais, S. (1955) "Measuring Social Mobility," *Journal of the Roayal Statistical Society*, Series A, 118(1): 56-66.

Regan, Tracy and Ronald Oaxaca (2009) "Work Experience as a Source of Specification Error in Earnings Models: Implications for Gender Wage Decompositions," *Journal of Population Economics*, 22, 463-499.

Richey, Jeremiah and Alicia Rosburg (2015) "Changing Roles of Ability and Education in U.S. Intergenerational Mobility," Unpublished Manuscript.

Rivers, Douglas and Quang Vuong (1988) "Limited Information Estimatros and Exogeneity Tests for Simultaneous Probit Models," *Journal of Econometrics*, 39, 347-366.

Rosenberg, M. (1965). *Society and the Adolescent Self-image*, Princeton, NJ: Princeton University Press.

Rothe, Christoph (2015) "Decomposing the Composition Effect," *Journal of Business and Economic Statistics*, (forthcoming).

Sacerdote, Bruce (2002) "The Nature and Nurture of Economic Outcomes," *American Economic Review*, 92, pp. 344-348.

Shea, John (2000) "Does Parents' Money Matter?," *Journal of Public Economics*, 111(3), pp. 358-368.

Shorrocks, A. (1978) "The Measurement of Mobility," *Econometrica*, 46: 1013-1024.

Sklar, A. (1959) "Fonctions de repartition n dimensions et leurs marges", Universite Paris 8, pp. 229-231.

Solon, Gary (1999) "Intergenerational Mobility in the Labor Market." In Handbook of Labor Economics, Volume 3 edited by Orley Ashenfelter and David Card. Elsevier Science.

Sommers, P. and J. Conlist (1979) "Eigenvalue Immobility Measures for Markov Chains", *Journal of Mathematical Sociology* 6: 253-276.

Theil, H. (1972) *Statistical Decomposition Analysis*, North Holland: Amsterdam.

Table 1: Transition Matrix for White Males in NLSY79

Parental Quartile	Childs' Quartile 1st	2nd	3rd	4th
1st	0.403 (0.026)	0.278 (0.024)	0.199 (0.023)	0.119 (0.017)
2nd	0.287 (0.025)	0.284 (0.023)	0.233 (0.023)	0.196 (0.021)
3rd	0.174 (0.019)	0.285 (0.024)	0.293 (0.022)	0.248 (0.022)
4th	0.136 (0.018)	0.151 (0.018)	0.270 (0.0230	0.443 (0.028)

Notes: Standard errors in parenthesis. Incomes are adjusted by state-level cost of living indices.

Table 2: Mobility Indices

Index	M(Independent matrix)	M(Table 1)
M_1	1	0.86
M_2	1	0.67
M_3	0.42	0.32
M_4	1	0.85

Notes: These are summary indices values for an 'independent' matrix as discussed in section 2.1 and the same indeces values for the empirical matrix in Table 1 based on white men from the 1979 NLSY.

Table 3: Summary Statistics - NLSY79 White Men

Variable	Parental Income Quartile				All
	Q1	Q2	Q3	Q4	
Parental income	12,435	25,044	35,126	57,416	32,503
	(4,804)	(3,436)	(4,176)	(14,713)	(18,411)
Youth income	12,511	14,852	16,864	20,260	16,121
	(7,943)	(8,974)	(8,878)	(10,700)	(9,600)
Youth income rank	0.38	0.46	0.54	0.64	0.5
	(0.27)	(0.28)	(0.26)	(0.28)	(0.29)
Experience	224	222	212	184	210
	(119)	(123)	(117)	(118)	(120)
Education	12.3	12.9	13.5	14.6	13.3
(years of schooling)	(2.5)	(2.2)	(2.2)	(2.4)	(2.5)
AFQT	0.05	0.38	0.52	0.78	0.43
	(1.05)	(0.89)	(0.89)	(0.79)	(0.95)
Rotter	8.81	8.57	8.42	8.23	8.51
	(2.49)	(2.14)	(2.26)	(2.46)	(2.35)
Esteem	21.76	22.17	22.64	23.05	22.4
	(3.87)	(3.90)	(3.96)	(4.09)	(3.98)
Birth year	61.2	61.3	61.4	61.2	61.3
	(2.03)	(2.03)	(2.06)	(2.03)	(2.04)
County wages	11,191	11,831	12,163	12,604	11,947
(age 17)	(2,635)	(2,609)	(2,684)	(2,487)	(2,653)
State unemployment rate	6.94	7.12	7.06	7.06	7.04
(age 17)	(1.72)	(1.89)	(1.83)	(1.83)	(1.82)
Minimum college tuition	494	517	533	510	513
(4 year public; age 14)	(220)	(237)	(234)	(252)	(236)
College in county	0.42	0.47	0.56	0.68	0.53
	(0.49)	(0.50)	(0.50)	(0.47)	(0.50)

Notes: Standard deviations in parenthesis. Incomes are constant 1982-84 dollars and adjusted for state-level cost of living indices. AFQT score is a standardized measurement.

Table 4: Index Decompositions

	M1		M2		M3		M4	
Total Difference: Δ_o^v	14.319***	(1.829)	32.426***	(2.577)	10.005***	(0.843)	15.35***	(1.222)
Structure Effect: Δ_s^v	5.5***	(1.791)	16.963***	(3.397)	4.9***	(0.887)	9.277***	(1.211)
Composition Effect: Δ_x^v	8.819***	(1.964)	15.463***	(3.195)	5.105***	(0.881)	6.073***	(1.145)
Dependence Effect: Δ_d^v	-0.368	(0.459)	-0.419	(1.175)	-0.167	(0.157)	0.175	(0.365)
Marginal Effect: Δ_m^v	9.187***	(2.1)	15.882***	(3.277)	5.272***	(0.904)	5.898***	(1.116)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	-0.057	(1.216)	2.946**	(1.412)	0.837*	(0.503)	1.152*	(0.677)
Esteem	0.353	(0.445)	0.573	(0.587)	0.191	(0.197)	0.223	(0.268)
Rotter	0.003	(0.34)	0.031	(0.425)	-0.007	(0.145)	-0.038	(0.211)
Education	11.084***	(2.123)	14.207***	(3.415)	5.236***	(0.935)	4.918***	(1.395)
Experience	-2.041*	(1.104)	-3.924***	(1.238)	-1.35***	(0.399)	-1.979***	(0.588)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\lvert k \rvert = 2$								
AFQT:Esteem	0.343	(0.47)	0.088	(0.381)	0.095	(0.154)	0.112	(0.232)
AFQT:Rotter	-0.063	(0.275)	0.454	(0.334)	0.073	(0.092)	0.105	(0.143)
AFQT:Education	0.687	(1.391)	2.502	(2.048)	0.564	(0.432)	1.459	(1.002)
AFQT:Experience	0.108	(1.231)	0.23	(1.023)	0.036	(0.321)	0.179	(0.554)
Esteem:Rotter	-0.203	(0.168)	0.004	(0.119)	-0.03	(0.047)	0.008	(0.066)
Esteem:Education	-0.256	(0.549)	-0.185	(0.637)	-0.044	(0.165)	-0.11	(0.342)
Esteem:Experience	0.107	(0.497)	-0.176	(0.399)	0.002	(0.148)	-0.036	(0.19)
Rotter:Education	0.093	(0.397)	0.262	(0.648)	0.088	(0.124)	0.115	(0.245)
Rotter:Experience	-0.012	(0.312)	0.002	(0.24)	0.017	(0.082)	0.035	(0.137)
Education:Experience	-2.776*	(1.576)	0.391	(1.905)	-0.557	(0.429)	0.854	(0.88)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.

Table 5: Decomposition of Top (Parental) Quartile

	1st Q		2nd Q		3rd Q		4th Q	
Total Difference: Δ_o^v	11.546***	(1.739)	9.47***	(1.814)	-2.326	(2.033)	-18.69***	(2.23)
Structure Effect: Δ_s^v	7.382***	(1.772)	5.964***	(2.079)	-3.2	(2.697)	-10.146***	(2.64)
Composition Effect: Δ_x^v	4.164***	(1.169)	3.506*	(1.969)	0.874	(2.341)	-8.544***	(2.587)
Dependence Effect: Δ_d^v	0.936***	(0.319)	-1.18***	(0.346)	-0.858	(0.73)	1.102	(0.778)
Marginal Effect: Δ_m^v	3.228***	(1.09)	4.686**	(1.969)	1.732	(2.488)	-9.646***	(2.718)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	1.784*	(0.962)	0.516	(1.254)	-1.772*	(1.004)	-0.528	(0.835)
Esteem	0.508	(0.346)	0.058	(0.475)	-0.478	(0.391)	-0.088	(0.242)
Rotter	-0.176	(0.235)	0.292	(0.303)	0.036	(0.361)	-0.152	(0.341)
Education	3.106**	(1.254)	5.002**	(2.149)	3.03	(2.794)	-11.138***	(2.783)
Experience	-3.55***	(1.042)	1.492	(1.322)	0.426	(1.116)	1.632*	(0.839)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\lvert k \rvert = 2$								
AFQT:Esteem	-0.33	(0.294)	0.466	(0.561)	0.004	(0.508)	-0.14	(0.245)
AFQT:Rotter	0.152	(0.183)	-0.054	(0.283)	0.26	(0.361)	-0.358	(0.299)
AFQT:Education	0.796	(1.192)	-0.462	(1.745)	1.658	(1.34)	-1.992*	(1.136)
AFQT:Experience	0.97	(1.461)	-1.742	(1.841)	0.814	(1.091)	-0.042	(0.775))
Esteem:Rotter	-0.004	(0.1)	0.06	(0.147)	-0.06	(0.137)	0.004	(0.113)
Esteem:Education	0.05	(0.347)	-0.018	(0.633)	0.65	(0.612)	-0.682	(0.424)
Esteem:Experience	-0.264	(0.443)	0.344	(0.639)	-0.072	(0.413)	-0.008	(0.189)
Rotter:Education	0.174	(0.194)	-0.214	(0.306)	0.78	(0.582)	-0.74	(0.513)
Rotter:Experience	0.174	(0.206)	-0.116	(0.294)	-0.1	(0.339)	0.042	(0.283)
Education:Experience	2.374*	(1.241)	-3.778**	(1.798)	-1.176	(2.007)	2.58	(1.593)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.

Table 6: Decomposition of Bottom (Parental) Quartile

	1st Q		2nd Q		3rd Q		4th Q	
Total Difference: Δ_o^v	-16.132***	(2.332)	-1.97	(2.12)	5.204***	(1.989)	12.898***	(1.612)
Structure Effect: Δ_s^v	-6.126***	(2.334)	-2.632	(2.962)	1.63	(2.951)	7.128**	(2.814)
Composition Effect: Δ_x^v	-10.006***	(2.284)	0.662	(2.859)	3.574	(2.757)	5.77**	(2.239)
Dependence Effect: Δ_d^v	-0.66	(0.474)	0.932*	(0.534)	-0.346	(0.638)	0.074	(0.692)
Marginal Effect: Δ_m^v	-9.346***	(2.168)	-0.27	(2.708)	3.92	(2.734)	5.696**	(2.214)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	-2.288	(1.96)	-2.19	(2.076)	4.106**	(1.729)	0.372	(1.038)
Esteem	-0.09	(0.703)	0.098	(0.801)	-0.14	(0.733)	0.132	(0.475)
Rotter	0.366	(0.53)	-0.496	(0.57)	-0.168	(0.507)	0.298	(0.385)
Education	-8.286***	(2.494)	3.876	(3.175)	-0.582	(3.692)	4.992*	(2.689)
Experience	2.96**	(1.441)	-1.838	(1.558)	-1.252	(1.371)	0.13	(0.952)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\|k\| = 2$								
AFQT:Esteem	-0.084	(0.545)	-0.314	(0.712)	0.406	(0.598)	-0.008	(0.155)
AFQT:Rotter	0.002	(0.291)	-0.35	(0.43)	0.262	(0.349)	0.086	(0.188)
AFQT:Education	-0.498	(1.588)	-0.424	(2.175)	0.186	(1.89)	0.736	(0.981)
AFQT:Experience	-0.59	(1.273)	1.572	(1.669)	-0.842	(1.266)	-0.14	(0.81)
Esteem:Rotter	-0.018	(0.12)	-0.14	(0.199)	0.14	(0.175)	0.018	(0.078)
Esteem:Education	0.904	(0.609)	-0.666	(0.97)	-0.704	(0.881)	0.466	(0.429)
Esteem:Experience	0.148	(0.567)	-0.212	(0.754)	0.146	(0.51)	-0.082	(0.233)
Rotter:Education	-0.194	(0.365)	0.316	(0.548)	-0.424	(0.642)	0.302	(0.51)
Rotter:Experience	-0.148	(0.256)	0.184	(0.393)	0.32	(0.361)	-0.356	(0.256)
Education:Experience	0.08	(1.621)	-1.354	(2.813)	0.984	(2.732)	0.29	(1.464)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.

Table 7: Index Decompositions - IV results

	M1		M2		M3		M4	
Total Difference: Δ_o^v	14.24***	(1.708)	32.608***	(2.509)	10.021***	(0.793)	15.401***	(1.161)
Structure Effect: Δ_s^v	5.287**	(2.1)	16.073***	(3.637)	4.657***	(1.042)	9.106***	(1.275)
Composition Effect: Δ_x^v	8.953***	(1.948)	16.535***	(3.352)	5.365***	(0.935)	6.295***	(1.259)
Dependence Effect: Δ_d^v	-0.232	(0.53)	0.439	(1.227)	-0.017	(0.214)	0.281	(0.415)
Marginal Effect: Δ_m^v	9.185***	(1.967)	16.097***	(3.197)	5.382***	(0.898)	6.014***	(1.227)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	1.252	(2.436)	6.837	(4.301)	1.957	(1.311)	2.589*	(1.527)
Esteem	-0.014	(0.537)	0.841	(0.786)	0.177	(0.24)	0.275	(0.336)
Rotter	0.001	(0.373)	0.123	(0.567)	0.009	(0.164)	-0.014	(0.233)
Education	12.153***	(4.506)	11.331*	(6.351)	4.634**	(2.207)	2.264	(2.674)
Experience	-3.056**	(1.219)	-3.296*	(1.774)	-1.35***	(0.487)	-1.943***	(0.719)
Control	-0.32	(2.414)	2.722	(4.61)	0.945	(1.239)	0.523	(1.519)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\|k\| = 2$								
AFQT:Esteem	0.753	(0.675)	-0.142	(1.279)	0.236	(0.209)	0.108	(0.37)
AFQT:Rotter	0.093	(0.394)	0.382	(0.856)	0.102	(0.126)	0.159	(0.242)
AFQT:Grade	0.143	(2.472)	-6.049	(4.83)	-0.884	(0.919)	-1.187	(1.92)
AFQT:Experience	-1.559	(1.243)	-0.945	(1.854)	-0.461	(0.346)	-1.09	(0.751)
AFQT:Control	-0.287	(2.396)	-5.586	(4.616)	-0.048	(0.8)	-2.684	(1.699)
Esteem:Rotter	-0.307	(0.198)	-0.018	(0.198)	-0.054	(0.057)	-0.013	(0.084)
Esteem:Grade	0.048	(0.699)	-0.149	(1.084)	0.000	(0.225)	0.422	(0.438)
Esteem:Experience	0.755	(0.595)	0.095	(0.53)	0.186	(0.161)	0.145	(0.246)
Esteem:Control	0.417	(0.682)	-1.372	(1.104)	0.022	(0.206)	-0.447	(0.427)
Rotter:Education	0.063	(0.348)	0.344	(0.804)	0.039	(0.118)	0.082	(0.298)
Rotter:Experience	-0.122	(0.297)	0.031	(0.313)	0.005	(0.077)	-0.024	(0.132)
Rotter:Control	0.07	(0.319)	0.037	(0.671)	0.092	(0.101)	0.103	(0.217)
Education:Experience	-1.812	(1.572)	-1.882	(3.074)	-0.784	(0.497)	0.398	(1.392)
Education:Control	-0.714	(2.17)	-3.303	(5.094)	-1.131	(0.79)	0.952	(2.17)
Experience:Control	-1.254	(0.959)	-0.59	(2.139)	-0.352	(0.321)	-0.58	(0.726)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.

Table 8: Decomposition of Top (Parental) Quartile - IV results

	1st Q		2nd Q		3rd Q		4th Q	
Total Difference: Δ_o^v	11.844***	(1.753)	9.294***	(1.724)	-2.492	(2.093)	-18.646***	(2.079)
Structure Effect: Δ_s^v	7.224***	(1.708)	5.174***	(1.916)	-2.944	(2.536)	-9.454***	(2.664)
Composition Effect: Δ_x^v	4.62***	(1.544)	4.12**	(1.965)	0.452	(2.392)	-9.192***	(2.369)
Dependence Effect: Δ_d^v	0.57*	(0.308)	-0.682*	(0.392)	-0.85	(0.869)	0.962	(0.96)
Marginal Effect: Δ_m^v	4.05***	(1.53)	4.802**	(1.945)	1.302	(2.487)	-10.154***	(2.411)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	2.372*	(1.371)	3.226	(2.041)	-1.688	(2.099)	-3.91*	(2.341)
Esteem	0.68	(0.462)	0.116	(0.6)	-0.572	(0.602)	-0.224	(0.5)
Rotter	-0.144	(0.244)	0.318	(0.284)	0.016	(0.378)	-0.19	(0.387)
Education	0.976	(2.259)	1.26	(2.546)	5.77	(4.764)	-8.006	(5.909)
Experience	-2.678***	(0.999)	1.9*	(1.02)	-0.216	(1.066)	0.994	(0.985)
Control	1.572	(1.311)	3.17	(2.346)	-1.022	(2.205)	-3.72	(2.294)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\|k\| = 2$								
AFQT:Esteem	-0.464	(0.482)	0.808	(0.858)	0.348	(0.787)	-0.692	(0.46)
AFQT:Rotter	0.114	(0.202)	0.07	(0.332)	0.532	(0.473)	-0.716*	(0.413)
AFQT:Education	-0.636	(1.615)	-2.694	(2.664)	2.422	(2.968)	0.908	(2.338)
AFQT:Experience	1.77	(1.24)	-3.958***	(1.725)	0.266	(1.435)	1.922*	(0.99)
AFQT:Control	-0.082	(1.584)	0.838	(2.736)	-1.154	(2.766)	0.398	(1.771)
Esteem:Rotter	0.014	(0.115)	0.018	(0.232)	-0.044	(0.315)	0.012	(0.224)
Esteem:Education	-0.262	(0.535)	0.074	(0.815)	1.12	(0.82)	-0.932	(0.582)
Esteem:Experience	-0.072	(0.455)	-0.444	(0.716)	0.606	(0.652)	-0.09	(0.341)
Esteem:Control	-0.538	(0.466)	0.96	(0.911)	0.046	(0.868)	-0.468	(0.482)
Rotter:Education	0.072	(0.216)	-0.034	(0.329)	0.382	(0.54)	-0.42	(0.468)
Rotter:Experience	0.106	(0.211)	-0.156	(0.285)	0.112	(0.36)	-0.062	(0.297)
Rotter:Control	0.134	(0.192)	0.092	(0.329)	0.31	(0.455)	-0.536	(0.383)
Education:Experience	0.35	(1.217)	-2.238	(1.616)	-2.52	(2.198)	4.408**	(1.841)
Education:Control	-0.62	(1.391)	-2.284	(2.495)	1.238	(2.982)	1.666	(2.16)
Experience:Control	1.194	(1.194)	-3.492**	(1.625)	0.802	(1.486)	1.496	(0.985)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.

Table 9: Decomposition of Bottom (Parental) Quartile - IV results

	1st Q		2nd Q		3rd Q		4th Q	
Total Difference: Δ_o^v	-16.176***	(2.349)	-1.984	(1.992)	5.1**	(1.984)	13.06***	(1.433)
Structure Effect: Δ_s^v	-6.99***	(2.555)	-1.656	(3.432)	1.888	(3.247)	6.758	(2.867)
Composition Effect: Δ_x^v	-9.186***	(2.437)	-0.328	(3.374)	3.212	(2.987)	6.302***	(2.439)
Dependence Effect: Δ_d^v	-1.2**	(0.542)	1.364*	(0.746)	-0.492	(0.883)	0.328	(1.056)
Marginal Effect: Δ_m^v	-7.986***	(2.465)	-1.692	(3.239)	3.704	(3.001)	5.974***	(2.06)
"Direct" Contibution to Marginal Effect: $\Delta_m^v(e^j)$								
AFQT	-1.116	(3.702)	-4.116	(3.818)	3.1	(3.406)	2.132	(3.827)
Esteem	0.16	(0.759)	-0.258	(0.964)	-0.23	(0.867)	0.328	(0.615)
Rotter	0.392	(0.535)	-0.554	(0.593)	-0.236	(0.567)	0.398	(0.431)
Education	-13.146**	(6.611)	2.13	(6.301)	2.68	(5.774)	8.336	(7.151)
Experience	4.226**	(1.836)	-2.418	(1.632)	-2.104	(1.627)	0.296	(1.396)
Control	3.668	(3.544)	-3.52	(3.636)	-1.808	(3.39)	1.66	(3.621)
"Two-Way" Interactions: $\Delta_m^v(k)$ with $\|k\| = 2$								
AFQT:Esteem	0.444	(0.746)	-0.592	(1.106)	0.024	(0.992)	0.124	(0.701)
AFQT:Rotter	0.07	(0.318)	-0.096	(0.529)	-0.38	(0.612)	0.406	(0.435)
AFQT:Grade	-2.476	(2.678)	7.922	(4.964)	-0.996	(5.42)	-4.45	(3.445)
AFQT:Experience	0.12	(1.23)	-0.096	(1.727)	0.698	(1.547)	-0.722	(1.057)
AFQT:Control	2.726	(2.786)	-1.412	(4.566)	-2.896	(4.437)	1.582	(2.418)
Esteem:Rotter	-0.02	(0.145)	-0.15	(0.262)	0.146	(0.329)	0.024	(0.234)
Esteem:Education	-0.024	(0.682)	0.666	(1.243)	-0.616	(1.167)	-0.026	(0.679)
Esteem:Experience	-0.464	(0.576)	0.238	(0.835)	0.29	(0.685)	-0.064	(0.393)
Esteem:Control	0.964	(0.737)	-0.226	(1.031)	-0.826	(0.853)	0.088	(0.595)
Rotter:Education	-0.088	(0.308)	0.04	(0.607)	0.484	(0.771)	-0.436	(0.596)
Rotter:Experience	-0.068	(0.213)	-0.012	(0.351)	0.3	(0.419)	-0.22	(0.303)
Rotter:Control	-0.116	(0.246)	0.138	(0.433)	-0.468	(0.6)	0.446	(0.414)
Education:Experience	-0.716	(1.528)	0.98	(2.612)	1.572	(2.664)	-1.836	(1.586)
Education:Control	-1.524	(2.314)	7.056	(4.423)	-1.344	(4.938)	-4.188	(3.214)
Experience:Control	-0.954	(1.348)	0.956	(1.88)	0.806	(1.56)	-0.808	(0.97)

Notes: Standard errors, based on 200 bootstraps, are in parenthesis. Statistical significance is denoted by *** for the 1% level, ** for the 5% level, and * for the 10% level. All results are multiplied by 100 for readability.