Decomposing Joint Distributions via Reweighting Functions: An Application to Intergenerational Economic Mobility

Richey, Jeremiah and Rosburg, Alicia

Office of the Comptroller of the Currency, University of Northern Iowa

24 October 2016

Decomposing Joint Distributions via Reweighting Functions: An Application to Intergenerational Economic Mobility

Jeremiah Richey
Credit Risk Analysis Division
Office of the Comptroller of the Currency
Washington, D.C.
Phone: 202-649-5537
Jeremiah.richey@occ.treas.gov

Alicia Rosburg
209 Curris Business Building
University of Northern Iowa
Cedar Falls, Iowa 50614
Phone: (319)273-3263
alicia.rosburg@uni.edu

October 25, 2016

Abstract

We propose a decomposition method that extends the traditional Oaxaca-Blinder decomposition to a continuous group membership setting that can be applied to any distributional measure of interest. This is achieved by reframing the problem as a decomposition of joint distributions: we decompose the difference between an empirical and a (hypothetical) independent joint distribution of membership index and an outcome of interest. Differences are divided into a composition effect and a structure effect. The method is based on the estimation of a counterfactual joint distribution via reweighting functions that can be caste into various distributional measures to investigate the drivers of the empirical relationship. We apply the method to U.S. intergenerational economic mobility and investigate multiple versions of the intergenerational elasticity of income (IGE): the traditional linear IGE, quantile regression counterparts, and a nonparametric IGE. Quantile results reveal a U-shaped effect which is primarily compositional in nature; nonparametric results indicate the composition effect is the main driver of the mean parental-offspring link at low levels of parental income while the structural effect is the main driver at high levels of parental income. Both of these effects are masked by the traditional IGE which implies an even 50-50 split between the composition and structure effect.

JEL Classification: C14, C20, J31, J62

Key Words: intergenerational mobility, decomposition methods

Acknowledgements: The views expressed in this paper are those of the authors alone and do not necessarily reflect those of the Office of the Comptroller of the Currency or the US Department of the Treasury. All errors are our own.

1 Introduction

Kindled by the seminal papers of Becker and Tomes (1979, 1986), a large body of economics literature has sought to understand the persistence of incomes across generations. The literature that seeks to understand the drivers of this intergenerational link has focused primarily on simple mean effects, such as the intergenerational elasticity of income (IGE) (Björklund *et al.*, 2006; 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, 2016a; Shea, 2000). Two exceptions are Bhattacharya and Mazumder (2011), who estimate conditional directional mobility measures and transition matrices, and Richey and Rosburg (2016b), who decompose transition matrices and related indices. While both studies are revealing and provide important additions to the literature, they have limitations that restrict their generalizability.[1] This somewhat narrow focus, due in part to a lack of methods, has limited our understanding of intergenerational mobility. Thus, we propose a simple estimation method that allows us to investigate the drivers behind the intergenerational link that can be applied to *any* measure of the parental-offspring relationship. The method, based on reweighting functions, extends the traditional aggregate Oaxaca-Binder decomposition to a continuous group membership setting applied to any distributional measure. The method 'decomposes' the joint distribution of parental and offspring incomes into the portion due to a link between parental income and offspring characteristics (a composition effect) and the portion due to a link between parental income and the returns to characteristics (a structure effect). The decomposition is achieved by estimating a counterfactual joint distribution based on the removal of the compositional component. This counterfactual joint distribution can then be used to analyze any measure of (im)mobility and identify what portion is compositional or structural in nature.

More generally, the proposed method builds on a large body of literature which seeks to understand differences in some outcome between groups. A prime example in labor economics is wage differences between men and women. The method often used to investigate such differences is some variation of the Oaxaca-Blinder (OB) decomposition (Oaxaca 1973, Blinder 1973). While the traditional OB decomposition focuses on understanding simple mean differences, recent literature has extended the traditional OB decomposition beyond simple mean comparisons to evaluate differences between two groups across the full distribution and functions thereof (Chernozhukov *et al.*, 2013; DiNardo *et al.*, 1996; Firpo *et al.*, 2007; Machado and Mata, 2005; Rothe, 2015). In some instances, however, the relevant question pertains to an outcome that varies along a continuous group membership rather than binary group membership. In particular, the literature on economic mobility seeks to understand to what degree and why offspring incomes vary with parental incomes. In this case, the outcome (offspring income) varies along a continuous group membership (parental income). While there is a large body of decomposition literature on binary groups, there is relatively limited work on 'continuous group' decompositions. Two notable exceptions are Ulrick (2012) and Nopo (2008) who evaluate

[1]Notably, the non-parametric procedure proposed by Bhattacharya and Mazumder (2011) is limited by the curse of dimensionality, and as a result, their application is limited to conditioning on a single covariate at a time. The decomposition approach by Richey and Rosburg (2016b) relies on arbitrary segmentations of the population (e.g., parental income quartiles).

mean differences but allow 'group membership' to be continuous and compare means between 'levels' of this group assignment; the procedure suggested here, which provides an aggregate decomposition of the full distribution of the outcome of interest while allowing group membership to vary continuously, is a natural extension of their work. The proposed method differs from traditional decompositions in that we do not explicitly ask what explains differences in observed outcomes between groups; rather, we ask what explains differences in the observed joint distribution and a hypothetical joint distribution in which offspring incomes are independent of parental income. The estimation process uses a reweighting method that can be seen as an extension to the method of DiNardo *et al.* (1996) beyond a binary group setting.

To better understand the driving forces behind (im)mobility in the US, we apply the decomposition method to intergenerational economic mobility of white males surveyed in the 1979 National Longitudinal Survey of Youth (NLSY). We base our wage structure on an extended Mincer equation that includes education, experience, and cognitive and non-cognitive measures. We decompose multiple mobility measures including standard IGEs, quantile regression counterparts to the IGE, and a nonparametric version of the IGE. To preview our results, we find an IGE of 0.43 with the composition and structure effect each accounting for about half of this relationship. This simple mean relationship, however, does not provide a complete picture. Quantile regressions reveal a U-shaped effect across quantiles (most pronounced in the lower quantiles) suggesting parental income has a strong safety-net effect that is largely driven by a composition effect. Nonparametric regressions of conditional mean offsprings income on parental income levels also reveal the composition effect is large at low levels of parental income but the structure effect dominates at high levels of parental income. Therefore, the simple IGE hides large heterogeneity in effects across offspring's conditional quantiles as well as large nonlinearities across parental income levels.

2 Method

2.1 Set-Up and Estimation

Our goal is to understand the relationship between two variables - parental income and offspring income. If y_c is the offspring's (adult) income and y_p is parental income, we are interested in why $f(y_c, y_p) \neq g(y_c, y_p) \equiv f(y_c)f(y_p)$. That is, why do we not see an independent joint density and what are the mechanisms through which parental income contribute to offspring income? Specifically, we wish to identify what part of the relationship is due to children from different households having different characteristics (composition effect) and what part is due to similar children from different households receiving different returns for those characteristics (structure effect).

Mobility measures of interest are typically some function of the joint distribution, $\nu(f(y_c, y_p))$, such as the IGE or a nonparametric version of the IGE. Therefore, we want to understand $\nu(f(y_c, y_p)) - \nu(g(y_c, y_p))$, which we define as the overall mobility gap:

$$\Delta_O^\nu = \nu(f(y_c, y_p)) - \nu(g(y_c, y_p)) = \nu_f - \nu_g. \tag{1}$$

Let us denote the actual joint density of (y_c, y_p) as $f_{a|a}$, which can be expressed as:

$$f_{a|a} = f(y_c, y_p) = \int f(y_c|x : y_p) f(y_p, x) dx \tag{2}$$

where $f(y_c|x : y_p)$ is implicitly defined by a continuous set of wage structures $y_c = w_{y_p}(x, \epsilon)$ with ϵ representing the set of unobserved characteristics. The $x : y_p$ notation in equation (2) indicates that parental income affects the conditional distribution of offspring incomes through varying returns on x.[2]

In a similar manner, denote the independent joint density $g(y_c, y_p)$ as $f_{i|i}$, which can be expressed as:

$$f_{i|i} = g(y_c, y_p) \equiv f(y_c) f(y_p) = \int g(y_c|x) g(y_p, x) dx \tag{3}$$

where the joint density $g(y_p, x) \equiv f(y_p) f(x)$ represents the case where the distribution of offspring characteristics are independent of parental income. In addition, the conditional CDF $g(y_c|x)$ no longer allows for returns to depend on parental income.

Now consider the following counterfactual joint density denoted as $f_{a|i}$:

$$f_{a|i} = \int f(y_c|x : y_p) g(y_p, x) dx \tag{4}$$

This is the counterfactual that would prevail if all children, regardless of parental incomes, had the same distribution of characteristics but retained their returns to characteristics (i.e., actual returns, independent characteristics). Letting $\nu(f_{a|i}) = \nu_{a|i}$, the composition and structural effects can be expressed as:

$$\Delta_X^\nu = \nu_f - \nu_{a|i} \tag{5}$$

$$\Delta_S^\nu = \nu_{a|i} - \nu_g, \tag{6}$$

which represents an aggregate decomposition of the mobility gap:

$$\Delta_O^\nu = \Delta_X^\nu + \Delta_S^\nu. \tag{7}$$

[2] This aspect, in the traditional decomposition literature, is made explicit by defining separate conditional distributions (wage structures) for each group, e.g., $f_g(y|x) = m_g(x, \epsilon)$ for $g = \{1, 2, 3...\}$. Given the continuous nature of parental income, we denote this by including parental incomes in the conditional notation.

Now note:

$$\begin{aligned} f_{a|i} &= \int f(y_c|x:y_p)g(y_p,x)dx && (8)\\ &= \int f(y_c|x:y_p)\frac{g(y_p,x)}{f(y_p,x)}f(y_p,x)dx && (9)\\ &= \int f(y_c|x:y_p)\Psi(y_p,x)f(y_p,x)dx && (10)\end{aligned}$$

where $\Psi(y_p,x) = \frac{g(y_p,x)}{f(y_p,x)}$ is the 'reweighting' function. Therefore, the counterfactual is simply the actual data reweighted.[3] Assuming the counterfactual distribution of x (i.e., $f(x)$) is identical to that in the observed population (a point we return to in the next section), we have:

$$\Psi(y_p,x) = \frac{g(y_p,x)}{f(y_p,x)} = \frac{f(y_p)f(x)}{f(y_p|x)f(x)} = \frac{f(y_p)}{f(y_p|x)}. \tag{11}$$

The distribution $f(y_p)$ can be estimated via kernel density estimation.[4] The estimation of $f(y_p|x)$ is done in two steps. First, we estimate the conditional CDF ($F(y_p|x)$) via the distributional regression approach suggested by Foresi and Peracchi (1995); this is based on estimating multiple standard probit models where we vary the cut-off across a grid along the outcome space. Specifically, we repeatedly estimate $Pr(y_{pi} \leq \tilde{y}_p|x) = \Phi(x\beta_{\tilde{y}_p})$ for $\tilde{y}_p \in \mathcal{Y}_p$ where X is the vector of productivity characteristics (e.g., education, experience, cognitive, and non-cognitive measures) and y_p is log parental income. Second, we use kernel density estimation on a numerical approximation to this distribution to obtain $f(y_p|x)$. Once weights for each observation are estimated, they are simply used in the calculation of any statistic of choice to obtain the counterfactual value.

The decomposition results will be well identified as long as *ignorability* holds (Fortin et al., 2011).

Identifying Assumption - Ignorability: *Let (y_p, x, ϵ) have a joint distribution. For all x in $\mathcal{X}$: ϵ is independent of y_p given $X = x$.*

The ignorability assumption assures that, conditional on observables, the distribution of unobservables are not dependent on parental income.[5] For example, if wealthier families had more 'motivated' or 'unmotivated' children (unobserved), the decomposition will still be identified as along as the distribution of motivation is identical across

[3] In a previous version of this paper, under a different title, we proposed an alternative, more 'brute force' approach to the counterfactual (Richey and Rosburg, 2016c). That approach paralleled Machado and Mata (2005) and took an 'estimate-and-simulate' approach that was very computationally intensive and required estimation of complex interactive conditional CDFs (i.e., $F(y_c|x:y_p)$). As a result, that process was more susceptible to misspecification. The approach provided here builds off the foundation of that paper, but provides an improved estimation procedure that avoids estimation of complex interactive conditional CDFs.

[4] Specifically, we use the KernSmooth package in R which has been shown by Deng and Wickham (2011) to be both fast and accurate.

[5] Ignorability is a less restrictive assumption than *independence*, which would require the unobservables to be independent of the covariates. Only ignorability is needed for identification of the structure-composition decomposition.

parental income when conditioned on observables.

In the above, we made implicit assumptions regarding the structure of $f(x)$ and the counterfactual to consider. We briefly discuss these choices in the following subsection.

2.2 Choice of Counterfactual

To estimate our counterfactual, we must choose an appropriate $g(y_p, x)$, or more specifically, $f(x)$. In the traditional decomposition literature, this is usually the distribution of characteristics for one of the groups, however that is not an option in this setting. An appealing choice for $f(x)$, and the one implicitly assumed above, is to assume the observed unconditional distribution of characteristics would still prevail if the connection between y_p and x were removed. This distribution of characteristics aligns with Nopo's (2008) definition of the composition effect - the part due to differences in characteristics between individuals and the 'average' individual.[6]

In all decomposition settings we must choose which counterfactual to consider, or the order to approach the decomposition. Above we considered the counterfactual outcome if the 'compositional' portion were removed from the joint distribution. Alternatively, we could consider the counterfactual if the 'structural' portion were removed. In general, they will lead to different results because characteristics are 'priced' differently in the two counterfactuals and returns to characteristics are 'sized' differently. For example, consider the counterfactual considered above ($f_{a|i}$). We first remove differences in characteristics between children from different parental income levels; this identifies and closes the composition effect. The remaining gap - the structure effect - represents the portion explained by children from different homes receiving different returns for characteristics after removing any differences in characteristics. As a result, differences in characteristics are priced at the actual (divergent) returns, while differences in returns to them are sized at the equalized distribution of characteristics. Which counterfactual is more appropriate will depend on the research question of interest. If one is interested in how a policy aimed at equalizing education and test scores across all households might shrink the mobility gap, then the counterfactual we consider would likely be more appropriate. Not all questions or contexts will yield a clear counterfactual choice.[7]

[6] If one wished to consider a different $f(x)$ for the counterfactual there would be an additional term in the constructed weights - $\frac{f'(x)}{f(x)}$ - to control for the fact that they are no longer equal and do not cancel in equation (11).

[7] If one did wish to consider ($f_{i|a}$), you would first need to obtain the 'independent data' to reweight. This could be done by randomizing parental income data over the rest of the variables and then continuing as above.

3 Measuring Mobility

A key benefit of our decomposition method is that it does not limit how one can investigate the joint distribution - it is completely flexible. Specific to our application regarding intergenerational mobility, alternative measures of mobility found in the literature represent different ways to summarize or capture some aspect of the joint parent-offspring distribution of incomes. Thus, once the weights have been estimated, we can investigate the role of the structure and composition effects for any mobility measure of interest. Here, we focus on the traditional intergenerational elasticity of income (IGE), quantile regression counterparts to the IGE, and a nonparametric version of the IGE. This method could easily be extended to other measures of mobility such as transition matrices or upward mobility measures.

Much of the economic mobility literature estimates intergenerational mobility by modeling expected log earnings of a child $[ln(y_c)]$ as a linear function of log parental earnings $[ln(y_p)]$:

$$E[ln(y_c)] = \alpha + \beta ln(y_p). \tag{12}$$

The value β is the IGE and $(1-\beta)$ is a measure of intergenerational economic mobility. This simple relationship is the workhorse of much of the existing literature on economic mobility. By comparing the observed (actual) IGE with the counterfactual IGE, we can ascertain what portion of the observed IGE is structural or compositional in nature.

While the standard IGE approach (equation 12) tells us how the conditional mean of offspring income varies with parental income, it is often informative to look beyond the simple mean. Therefore, a natural extension to the basic IGE is through quantile regressions. The goal of a quantile regression, first introduced by Koenker and Bassatt (1978), is to identify the effect of an explanatory variable on different points of the conditional distribution of the independent variable. For example, previous results for the U.S. have shown larger IGE effects at the lower end of the income distribution, which may reflect parental income acting as a safety net (Eide and Showalter, 1999). Our method allows us to investigate not only how the relationship varies across the distribution but also what mechanism (compositional or structural) drives the relationship across the distribution. Specifically, we model the conditional quantile of offspring log income as:

$$Q_{ln(y_c)|ln(y_p)}(\tau) = \alpha_\tau + \beta_\tau ln(y_p) \tag{13}$$

where τ is the quantile of interest such as the median, 25th percentile, or 75th percentile. The coefficient vector β_τ will, in general, differ for each quantile. Our decomposition method will identify the portion of each β_τ that is structural or compositional in nature.

Both of the above approaches are based on linear functional forms and are quite common in the literature on economic mobility. However, a key theme in the original work by Becker and Tomes (1986) is that, in the face of credit constraints, nonlinearities may exist in the relationship between generations (see Corak and Heisz (1999) and Grawe (2004) for work along these lines). Moreover, if nonlinearities exist, simple linear IGE estimates can cloud mobility comparisons.[8] For both of these reasons, nonparametric versions of mobility measures have become more common in the literature. A nonparametric framework amounts to adding flexibility to Equation 12. Specifically, let:

$$E[ln(y_c)] = h(ln(y_p)). \tag{14}$$

where $h(\cdot)$ is some unknown function we wish to estimate. There are multiple ways to estimate $h(\cdot)$ such as kernel or local polynomial regressions. Once we have our estimated weights, we can investigate counterfactual versions of $h(\cdot)$ and ask what part of the empirical $h(\cdot)$ is structural or compositional in nature.

4 Data

The data for our analysis is the 1979 National Longitudinal Survey of Youth (NLSY79). 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).[9] We use only the cross-sectional representative 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 since parents' (average) income is a key variable of interest, to those with reported parental income for those years.[10] Further, we drop observations where the parental incomes were deemed outliers based on a rule of thumb of 1.5 times the interquartile range below the first quartile.[11] The outcome of interest is the individual's economic status based on their average reported wage and salary income from 1994, 1996 and 1998. All incomes are deflated to 1982-1984 dollars using the CPI.

The sample is further limited to individuals not enrolled in school over the period of interest and with available Armed Forces Qualifying Test (AFQT) scores. The final sample includes 1,368 individuals with a mean age of 33.6 during our outcome years of interest (1994-1998).[12] Table 1 provides summary statistics.

[8] Bratsberg et al. (2007) illustrate how international comparisons of simple IGE estimates can be misleading in the presence of nonlinearities.

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

[10] We exclude individuals who lived with a spouse or child during these years, and identify parental income through a comparison of total household income and respondent's income. Measurement error in parental income is a common concern in the mobility literature. Though recent research indicates that some mobility measures are less susceptible to such errors relative to others (Nybom and Stuhler, 2015).

[11] We found the estimator to be somewhat sensitive to outliers. Weights can explode in magnitude as a result of outliers (linked to very small estimated denominators in the weights), causing a few individuals to have disproportionate influence. The trimming here controls for this. This is related to our relatively small sample size which is sparse in the lower tail and simulations suggest larger sample sizes will be **less susceptible** to this concern.

[12] The literature on intergenerational mobility has identified the possibility of life-cycle biases is estimates depending on age

[Table 1 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 three measures of non-cognitive ability (*Esteem*, *Rotter*, and *Perlin*).

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.[13] 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.

Educational attainment is measured as years of schooling. The endogenous nature of education in explaining incomes is well documented; however, as long as the appropriate conditions hold (ignorability), this is not a concern for the *aggregate* decomposition. The measure of ability used in our analysis is Armed Forces 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).

We use three measures for non-cognitive 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). *Rotter* and *Esteem* were measured in the first two rounds of the survey (1979 and 1980, respectively). Therefore, the third measure we include is the Perlin Mastery Scale measured in 1992 when respondents were in their late 20s or early 30s. The Perlin Mastery Scale measures the extent to which individuals "perceive themselves in control of forces that significantly impact their lives" (BLS, 2015).

at which children are surveyed (Böhlmark and Lindquist, 2006; Haider and Solon, 2006; Nybom and Stuhler, 2015). However, this literature seems to indicate such biases are minimized or eliminated when youths reach their mid-30s.

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

5 Results

First, we consider the standard IGE results reported in column 1 of Table 2 (note: results are multiplied by 100 for readability). The total mobility gap is 42.85. This implies as parental income increases by 1%, offsprings' expected income increases by 0.43%. The composition and structure effect each account for about 50% of the measured mobility gap, meaning differences in characteristics of offspring across households and differences in returns to characteristics both contribute evenly to the measured mobility gap. While the structure effect is commonly interpreted as a measure of discrimination in many traditional decomposition settings, Richey and Rosburg (2016b) argue that a more fitting interpretation in this context is some form of (loosely-termed) 'privilege' such as parental connections, parental knowledge/awareness of job market and education opportunities or perhaps greater financial flexibility to facilitate job search. All of these would result in higher returns to similar productive characteristics.

[**Table 2 about here**]

The quantile IGE results in Table 2 reveal a U-shaped effect with larger effects in the two tails, especially in the lower tail which is nearly twice the effect at the median. This suggests that as parental income increases, offspring's conditional wage distribution exhibits compression at the lower end (decreasing 50-10 gap) and an extending upper tail (increasing 90-50 gap). This increased effect at the lower end is consistent with the pattern reported by Eide and Showalter (1999) and suggests that parental income has a strong safety-net effect. Further, decomposition results suggest that this safety-net effect is primarily compositional in nature; the composition effect increases dramatically from the median to the 10th quantile (nearly tripling) while the structure effect increases less than 50%. As we move upwards from the median to the 90th quantile, the composition effect again increases dramatically while the structure effect is nearly cut in half; therefore, the extending upper tail effect also appears to be compositionally driven. This means if all characteristics were equalized across households, we would still see some income compression at the lower end of the distribution for offspring from wealthier households, though not nearly as much as we see empirically. Further, rather than seeing an extending upper tail, we would see compression there as well.

Figure 1 presents the empirical and counterfactual nonparametric relationships between offspring's (conditional) mean income and parental income. The figure illustrates the nonlinearities in the empirical relationship. At low levels of parental income, the empirical relationship is very steep. This relationship flattens out in the middle income range before becoming steep again. The counterfactual relationship reveals a very interesting result. Removal of the composition effect appears to wipe out most of the relationship at low levels of parental income, but has little impact at higher levels of parental income (i.e., steepness remains). And while a linear approximation to the empirical relationship may not appear too bad, it is clearly not a very good one for the counterfactual. We can more directly see the absolute magnitudes of the composition and structure effect by comparing the plotted curves to the unconditional mean (the horizontal line at 9.5).[14] The space between the empirical relationship and the counterfactual relationship is the composition effect, while the space between the counterfactual and the unconditional mean represents the structure effect. These are plotted separately in Figures 2 and 3 with confidence intervals based on 2,000 bootstraps.[15] The structure effect is small and insignificant at low levels of parental income and increases greatly for children from wealthier homes; as a result, the structure effect appears to dominate at the top end of the parental distribution while the composition effect has limited contribution. Conversely, at low levels of parental income the composition effect is large and significant and appears to be the driving force to the empirical income gap, though is insignificant at higher parental income levels. This implies equalizing characteristics across all households would have a large effect on the intergenerational link at low parental income levels but have minimal effect on the intergenerational link at high parental income levels.

[**Figures 1-3 about here**]

6 Conclusion

A wealth of research over the past few decades has improved our understanding and empirical estimation of intergenerational economic mobility. Much of the recent empirical literature that aims at identifying the potential driving forces behind this intergenerational link has focused primarily on mean effects. While understanding mean effects is important, it is also somewhat limiting. A few studies have evaluated driving forces beyond mean effects but are hindered by a lack of generalizability. In this paper, we proposed a simple decomposition method that circumvents these previous limitations. The method we propose directly decomposes the joint parent-offspring distribution of incomes. In particular, we decompose the difference between the empirical joint distribution and a hypothetical independent joint distribution. Our decomposition is based on reweighting functions that remove the link between parental income and offspring's characteristics (a composition effect). This counterfactual can be caste into any measure of mobility and used to identify the portion of the measured mobility gap that is structural or compositional in nature.

To better understand intergenerational mobility in the U.S., we apply this method to a cohort of white males surveyed in the 1979 NLSY and to multiple versions of intergenerational elasticity of income (IGE). We find an IGE of 0.43 with the composition and structure effects each accounting for about half of this relationship. However, this simple mean relationship hides large heterogeneity in effects across offspring's conditional quantiles as well as large nonlinearities across parental income levels. Quantile regressions reveal a U-shaped effect across quantiles (most pronounced in the

[14]This implicitly assumes that the conditional distribution of offsprings' incomes on covariates ($f(y_c|x)$) would remain in the hypothetical 'independent' world. This would lead to the same unconditional mean. Without this implicit assumption we can only discuss the shape of the nonparametric relationship and not its distance from the independent case.

[15]The large confidence intervals at the tail ends is not unexpected as nonparametric estimators already tend to have poor performance on the edges, and we are seeing that behavior mixed with the uncertainty of our weights.

lower quantiles) driven primarily by the composition effect. Nonparametric regressions of conditional mean offsprings income on parental income levels reveal the composition effect is large at low levels of parental income but the structure effect dominates at high levels of parental income.

While we focus our application on economic mobility, the method proposed here can be applied to any joint distribution where the relationship is believed to be due to differing levels of intermediate variables and differing returns to them. Furthermore, while we focus on the mean regression and quantile and nonparametric versions of it, the method can be applied to any function of joint distributions such as transition matrices, directional migration measures or other correlation measures.

References

Altonji J, Bharadwaj P, Lange F. 2012. Changes in the characteristics of american youth: implications for adult outcomes. *Journal of Labor Economics* **4**: 783-828. DOI: 10.1086/666536

Becker G, Tomes N. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* **87**: 1153-1189.

Becker G, Tomes N. 1986. Human capital and the rise and fall of families. *Journal of Labor Economics* **4**(3):S1-S39. DOI: 10.1086/298118.

Bhattacharya D, Mazumder B. 2011. A nonparametric analysis of black-white differences in intergenerational income mobility in the United States. *Quantitative Economics* **43**(1): 139-172. DOI: 10.3982/QE69

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. DOI: 10.1162/qjec.121.3.999

Blanden J, Gregg P, Macmillan L. 2007. Accounting for intergenerational income persistance: noncognitive skills, ability and education. *The Economic Journal* **117**:C43-C60. DOI: 10.1111/j.1468-0297.2007.02034.x

Blinder A. 1973. Wage discrimination: reduced form and structural estimates. *Journal of Human Resources* **8**:436-455. DOI: 10.2307/144855

Böhlmark A, Lindquist M. 2006. Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. *Journal of Labor Economics* **24**(4): 879-896. DOI: 10.1086/506489

Bowles S, Gintis H. 2002. The inheritance of inequality. *Journal of Economic Perspectives* **16**(3): 3-30. DOI: 10.1257/089533002760278686

Bratsberg B, Røed K, Raaum O, Naylor R, Jäntti M, Eriksson T, Österbacka E. 2007. Nonlinearities in intergenerational earnings mobility: consequences for cross-country comparisons. *The Economic Journal* **117** C72-C92. DOI: 10.1111/j.1468-0297.2007.02036.x.

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 B, Johnston D, Martin V. 2013. Intergenerational income mobility: a new decomposition of investment and endowment effects. *Labour Economics* **24**: 39-47. DOI:10.1016/j.labeco.2013.05.007

Chernozhukov V, Fernandez-Val I, Melly B. 2013. Inference on counterfactual distributions. *Econometrica* **81**: 2205-2268. DOI: 10.3982/ECTA10582.

Corak M, Heisz A. 1999. The intergenerational earnings and income mobility of Canadian men: evidence from longitudinal income tax data. *The Journal of Human Resources* **34**: 504-533.

Deng H, Wickham, H. 2011. Density estimation in R. Unpublished Manuscript

DiNardo J, Fortin N, Lemieux T. 1996. Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach. *Econometrica* **64**: 1001-1044. DOI: 10.2307/2171954

Eide, E, Showalter, M. 1999. Factors affecting the transmission of earnings across generations: a quantile regression approach. *The Journal of Human Resources* **34**(2): 253-267. DOI: 10.2307/146345

Firpo, S, Fortin, N, Lemieux, T. 2007. Decomposing wage distributions using recentered influence function regressions. Mimeo, University of British Columbia. http://www.economics.uci.edu/files/docs/micro/f07/lemieux.pdf

Foresi S, Peracchi F. 1995. The conditional distribution of excess returns: an empirical analysis. *Journal of the American Statistical Association* **90**: 451-466. DOI: 10.2307/2291056

Fortin N, Lemieux T, Firpo S. 2011. Decomposition methods in economics. Chapter 1 in O. Ashenfelter, R. Layard, and D. Card (eds.), *Handbook of Labor Economics*. Volume 4(A): 1 - 102. New Holland.

Grawe, Nathan. 2004. Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints. *Journal of Human Resources* **39** 813-827.

Haider S, Solon G. 2006. Life-cycle variation in the association betwen current and lifetime earnings. *American Economic Review* **96**(4): 1308-1320. DOI: 10.1257/aer.96.4.1308

Koenker R, Bassett G. 1978. Regression quantiles. *Econometrica* **46**: 33-50. DOI: 10.2307/1913643

Koenker R, Leorato S, Peracchi F. 2013. Distributional vs. quantile regression. *Center for Economic and International Studies Working Paper Series* **11**(15), No. 300. https://ideas.repec.org/p/eie/wpaper/1329.html

Lefgren L, Lindquist M, Sims D. 2012. Rich dad, smart dad: decomposing the intergenerational transmission of income. *Journal of Political Economy* **120**: 268-303. DOI: 10.1086/666590

Liu H, Zeng J. 2009. Genetic ability and intergenerational income mobility. *Journal of Population Economics* **22**: 75-95. DOI: 10.1007/s00148-007-0171-6

Machado J, Mata J. 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics* **20**: 445-465. DOI: 10.1002/jae.788

Mayer S, Lopoo L. 2008. Government spending and intergenerational mobility. *Journal of Public Economics* **92**: 139-158. DOI: 10.1016/j.jpubeco.2007.04.003.

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

Nopo, H. 2008. An extension of the Blinder-Oaxaca decomposition to a continuum of comparison groups. *Economics Letters* **100**: 292-296. DOI: 10.1016/j.econlet.2008.02.011

Nybom M, Stuhler J. 2015. Biases in standard measures of intergenerational income dependence. Working Paper. Accessed October 9, 2015: https://janstuhler.wordpress.com/research/.

Oaxaca R. 1973. Male-female wage differentials in urban labor markets. *International Economic Review* **14**: 693-709. DOI: 10.2307/2525981.

Regan T, Oaxaca R. 2009. Work experience as a source of specification error in earnings models: implications for gender wage decompositions. *Journal of Population Economics* **22**: 463-499. DOI: 10.1007/s00148-007-0180-5

Richey J, Rosburg A. 2016a. Changing roles of ability and education in U.S. intergenerational mobility. *Economic Inquiry,* forthcoming. DOI: 10.1111/ecin.12362.

Richey J, Rosburg A. 2016b. Decomposing economic mobility transition matrices. Working paper, https://mpra.ub.uni-muenchen.de/66485/

Richey J, Rosburg A. 2016c. Understanding intergenerational economic mobility by decomposing joint distributions. Working paper, https://ideas.repec.org/p/pra/mprapa/72665.html

Rothe C. 2015. Decomposing the composition effect: the role of covariates in determining between-group differences in economic outcomes. *Journal of Business and Economic Statistics* **33**(3): 323-337. DOI:10.1080/07350015.2014.948959

Shea J. 2000. Does parents' money matter? *Journal of Public Economics* **111**(3): 358-368. DOI: 10.1016/S0047-2727(99)00087-0

Ulrick S. 2012. The Oaxaca decomposition generalized to a continuous group variable. *Economics Letters* **115**(1): 35-37. DOI: 10.1016/j.econlet.2011.11.037

Table 1: Summary Statistics - NLSY79 White Males

Variable	Mean	St. Dev
Parental income	33,678	18,079
Offspring income	16,317	11,215
Experience	12.8	3.5
Education	13.7	2.6
AFQT	0.51	0.93
Age	33.6	2.2
Rotter	8.4	2.3
Esteem	22.5	4.0
Perlin	22.6	3.0

Notes: Incomes are constant 1982-84 dollars. AFQT score is a standardized measurement.

Table 2: Decomposition of Traditional IGE and Quantile Regression IGEs for White Males in the NLSY79

Effect	IGE	10th	25th	50th	75th	90th
Total	42.85	61.65	43.09	33.08	37.91	43.59
	(3.92)	(11.47)	(5.34)	(3.69)	(4.1)	(5.48)
Structure	21.94	28.16	22.31	20.74	20.22	12.28
	(5.29)	(10.68)	(5.23)	(3.83)	(8.14)	(13.68)
Composition	20.91	33.49	20.78	12.34	17.69	31.32
	(5.57)	(12.59)	(5.92)	(4.4)	(7.96)	(13.28)

Notes: Standard errors, based on 2,000 bootstraps, are in parentheses. 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.

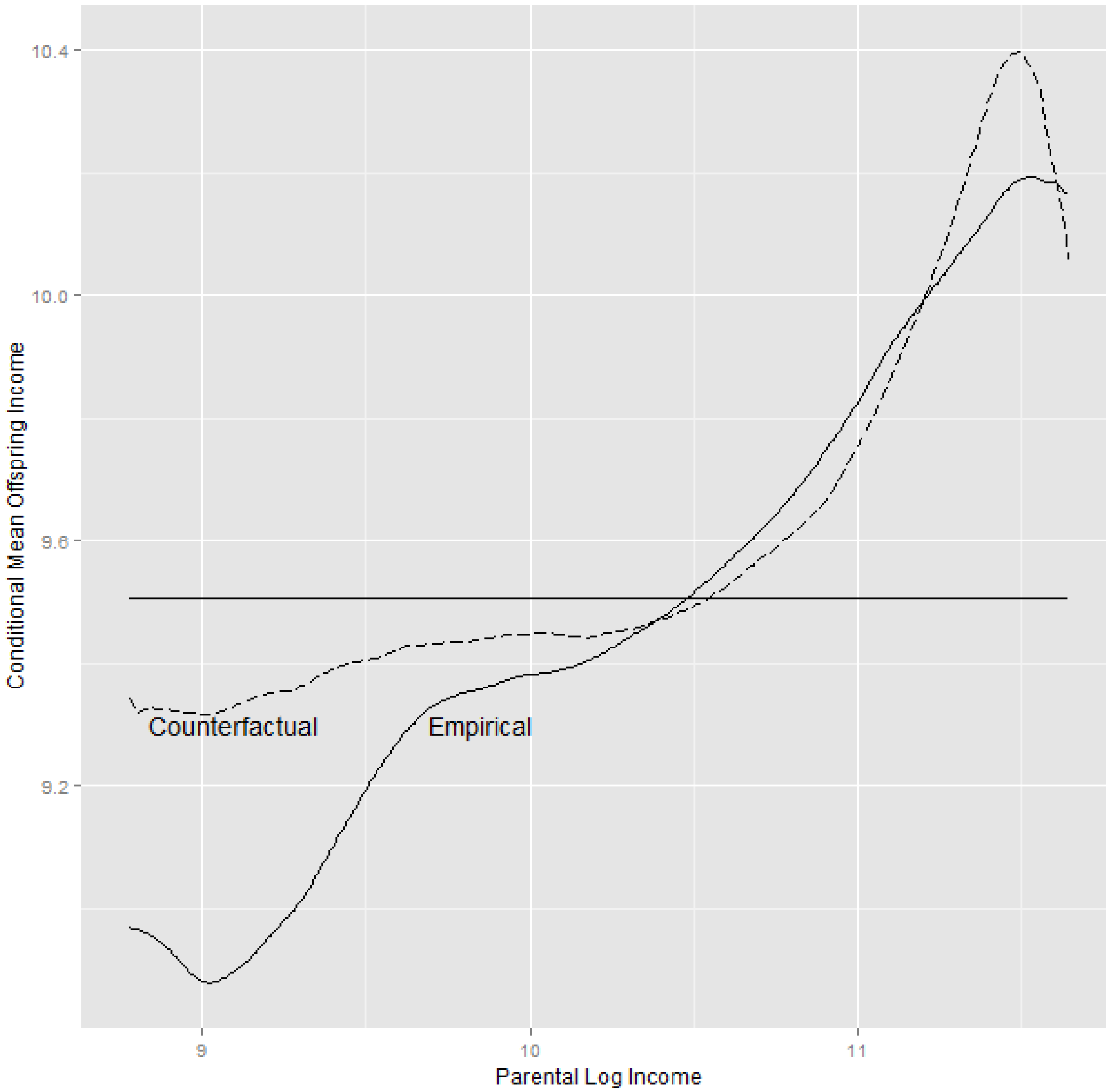

Figure 1: Nonparametric IGEs.

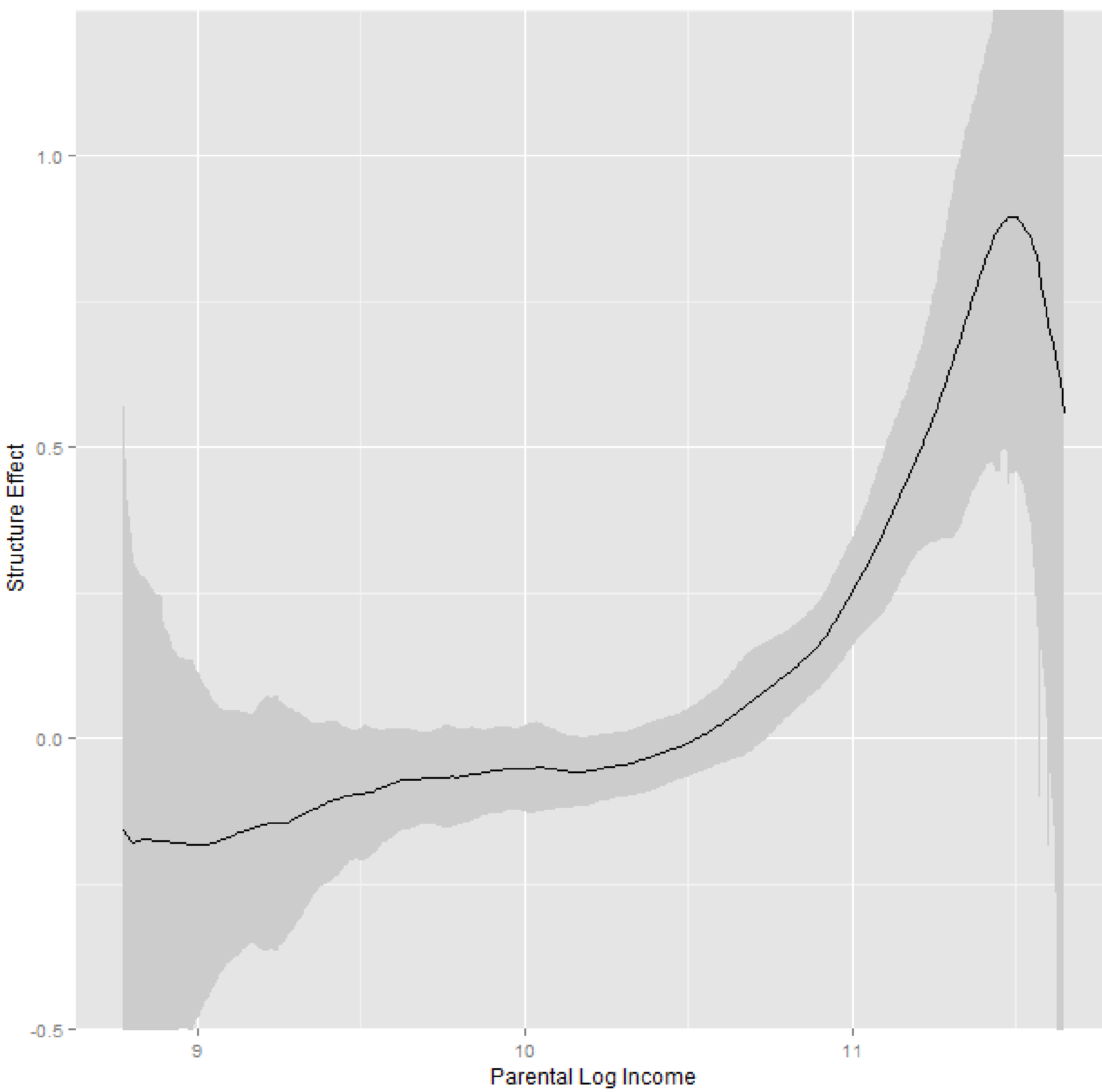

Figure 2: Structure Effect.

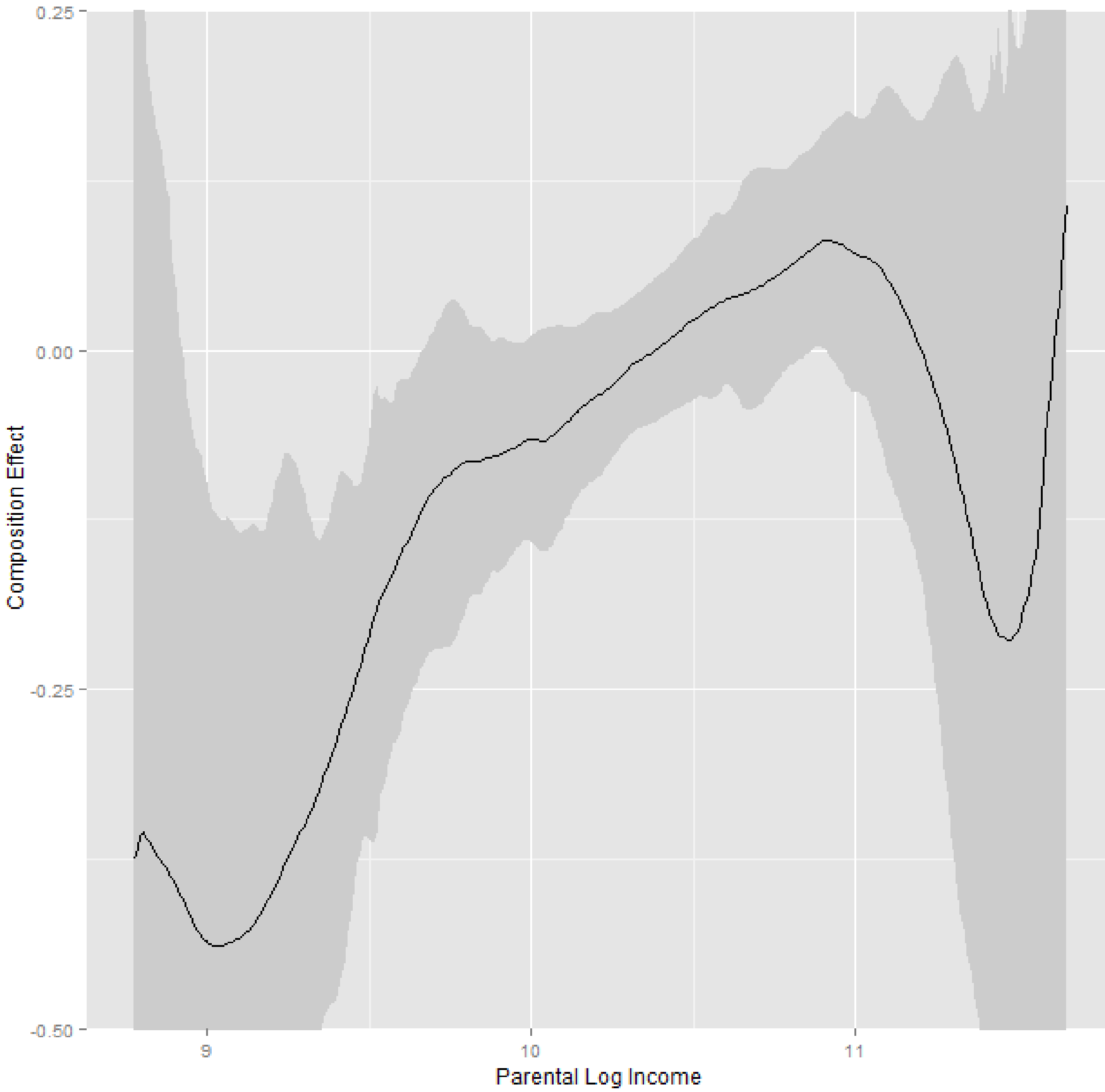

Figure 3: Composition Effect.