CAEPR Working Paper
#2015-016

Your American Dream is Not Mine!
A New Approach to Estimating Intergenerational Mobility Elasticities

Yonghong An
Texas A&M University

Le Wang
The University of Alabama

Ruli Xiao
Indiana University

September 28, 2015

Your American Dream is Not Mine!

A New Approach to Estimating Intergenerational Mobility Elasticities

Yonghong An* Le Wang† Ruli Xiao‡

August 2015

Abstract

This paper provides a new framework to estimate intergenerational mobility elasticities (IGE) of children's income with respect to parental income. Our approach is nonparametric allowing for heterogeneity in IGEs and nonlinearity by leaving the relationship unspecified, while acknowledging the *latent* nature of both child and parental *permanent* incomes. Our approach also addresses life-cycle bias directly in estimation without requiring knowledge of *permanent* income as the previous literature does. We confirm some of the previous findings and also have some novel results. First, we uncover that the association between observed annual income and permanent income varies over the life cycle; and that the observed patterns differ over generations, although the latter result is statistically insignificant due to large variances. Second, we find strong evidence that there exists substantial heterogeneity in IGEs across population and that the mobility function is nonlinear. The implied IGE exhibits a U-shape pattern with "twin-peaks". Specifically, there is a considerable degree of mobility among the broadly defined middle class, but both the children of high and low income fathers are more likely to be high and low income adults, respectively. This result suggests that the U.S. is indeed the "land of opportunity" to live out the "American Dream", just not for everyone! These results survive a battery of robustness checks. Finally, we also find that there exist a great deal of *within-group* heterogeneity. Our approach is also immediately applicable in many other similar contexts both within and outside labor economics.

Keywords: Intergenerational Mobility, Inequality, Measurement Error Models, Nonparametric Estimation

*Department of Economics, TAMU, College Station, TX 77843. E-mail: y.an@tamu.edu; Phone: (979) 845-8685; Fax: (979) 847-8757.

†Department of Economics, Finance, and Legal Studies, The University of Alabama, Tuscaloosa, AL, 35473. Email: Le.Wang.Econ@gmail.com; Phone: (205) 348-8967; Fax: (205) 348-0590.

‡Department of Economics, Indiana University, 100 S. Woodlawn Ave. Bloomington, IN 47405. Email: rulixiao@indiana.edu; Phone: (812) 855-3213 ; Fax: Fax: (480) 855-3736.

1 Introduction

Equality of opportunity, that individual success is determined by one's effort and motivation, but not one's family background, is a value central to the so-called "American Dream". In a society with fair equality of opportunity, any observed measure of one's success should be independent of the success of her parents. To the extent that income is a good measure of success, the size of the correlation between child income and parental income is often used to gauge the degree of the equality of opportunity, and to provide important insights into the evolution of income distributions over generations.[1] A larger correlation, often called intergenerational mobility (IGM) measure, implies a lower degree of equality of opportunity, and of mobility in moving up or down the social ladder across generations. Economists and social scientists have long been interested in documenting this measure (Black and Devereux, 2011; Durlauf and Shaorshadze, forthcoming), and estimation of it presents much more than academic curiosity. In this paper, we advance the literature by providing a new framework to estimate IGM.

Our paper is motivated by the often contradictory views on the size of the IGM in the U.S. between the general public and economists, and even among economists ourselves. On the one hand, popular writings on both the "underclass" and "the very wealthy" have emphasized the stickiness in the intergenerational transmission of income, and hence the lack of mobility: the poor have little opportunity to escape the poverty of their parents, while the very rich continue to enjoy the same success as their parents. On the other hand, early estimates of the correlation between the log of child and parental incomes (using linear regressions) in the economics literature are often very low (0.2 or less), indicating a high degree of IGM in the U.S. (e.g., Behrman and Taubman, 1985; Table 1 of Becker and Tomes, 1986 for more examples). It was these estimates that led Becker to conclude, in his influential presidential address to the American Economic Association, that "low earnings as well as high earnings are not strongly transmitted from fathers to sons" (Becker, 1988); and in Becker and Tomes (1986) that "Almost all the earnings advantages or disadvantages of ancestors are wiped out in three generations." The divergent views between popular writings and academic works have long existed dating back to at least early 1980s (see Solon (1992) for a nice summary).

We argue in this paper that the different views are not necessarily conflicting. Specifically, the divergent views highlight the fact that there exists potential heterogeneity/nonlinearity in the intergenerational transmission of income across the population: the views among the general public focus on the correlation at both extremes of the income distribution, while the regression estimates are better thought of as the correlation for "average" individuals (or average sizes of the correlations in the population). It is well documented that the family plays a vital role in shaping child outcomes

[1] Occupation is another popular measure of success, especially in the sociology literature.

through genetics, investment in human capital, and through choice of child environments (Cunha and Heckman, 2007). Economic theories have predicted that families with different socio-economic statuses differ in their resources and incentives to invest in their children, which in turn influence the future success of their children. Various mechanisms such as education have also been shown to give rise to heterogeneity in IGMs across socio-economic groups. As such, the correlations can be drastically different across the population, and therefore it would not be surprising if the views based on different sub-populations could differ from one another. Heterogeneity matters, but in *a priori* unknown fashion, thereby calling for a necessarily nonparametric approach.

Traditional nonparametric approach is, however, not applicable in this context. What makes estimation of IGM uniquely challenging is that neither the dependent nor the independent variables are observed. In his seminal work (attempting to reconcile the aforementioned differences), Solon (1992) points out this more fundamental problem with estimation: the relevant income measure is *permanent income* instead of annual income (which can be thought of as a combination of both permanent and transitory incomes). Solon's significant contribution to this literature is to acknowledge the latent nature of permanent income, and to recast the estimation problem as a standard, linear error-in-variables model, in which any measure of income in a given year is a less than perfect representation of permanent income, measured with error. Treating this issue as a *classical* measurement error problem, Solon finds much larger estimates and hence lower mobility, which reconciles the divergent views to some extent. His research has since spawned much of the subsequent literature to pay more attention to this issue. We need to emphasize that while data of better quality (such as federal income tax records used in Chetty et al. (2014a)) can help reduce self-reported errors in *observed* annual income, it cannot necessarily eliminate the error due to the fact that such observed annual income is an imperfect measure of permanent income. Due to the difficulties in addressing this problem, the literature within economics has found vastly different degrees of mobility in the U.S., even for the same cohort of children (Durlauf and Shaorshadze, forthcoming).[2]

This paper is among the first attempts to provide a unified framework confronting *both* measurement error *and* heterogeneity/nonlinearity issues surrounding estimation of IGM. Most of the previous literature has only dealt with each issue separately, using different approaches that are not necessarily comparable. On the one hand, the literature on estimation of IGE typically employs simple regression methods and follows Solon's method that averages parental annual income across *all* time periods to reduce the measure error issue. The average method assumes that the association between annual and permanent income is the same over the life cycle, an assumption challenged by Haider and Solon (2006) and others (e.g., Grawe, 2006) using the non-classical mea-

[2]For example, for the cohorts of children born in the 1950s and 1960s, the estimates vary from 0.2, 0.4, to 0.6 (Durlauf and Shaorshadze, forthcoming).

surement error framework; violation of the assumption leads to the life-cycle bias in estimates of IGE. On the other hand, a typical solution to get at heterogeneity is to employ the transition matrix approach, whose results heavily depend on the discretization of the income distribution and are not directly comparable to the regression estimates of IGE. We review each approach in detail below to highlight potential shortcomings with the existing approaches. Building on recent advances in the literature on nonparametric identification of errors-in-variables models, especially those by Hu and Sasaki (2014), we overcome many of the shortcomings associated with the existing approaches. Our approach is nonparametric allowing for nonlinearity by leaving the relationship unspecified, while acknowledging the *unobservable* nature of both dependent and key independent variables (i.e., child and parental *permanent* incomes). This approach can accommodate non-classical measurement error, and hence, addresses the life-cycle bias directly in estimation. Additionally, this nonparametric approach has a closed-form estimator, which is practically more convenient and is preferred to many alternative estimators relying on numerical optimizations.

We further extend the analytical form of the approach to derive an explicit expression of the life-cycle bias for our nonparametric framework, which nests many interesting cases such as linear models. This expression allows us to relax the normalization assumption, typically assumed in the broad literature on measurement error or factor models, to bound the true function. This extension itself is a novel contribution to the econometrics literature. Our approach is immediately applicable in many other similar contexts both within and outside labor economics. Monte Carlo simulation exercises indeed indicate superior finite-sample performances of this approach.

Our approach has some additional advantages. First, a by-product of our approach is a *model-free* way to estimate life-cycle variation in the association between annual and permanent incomes, subject to a required normalization. The association is also allowed to vary across generations. This result has important implications for many influential economic studies where the relevant economic variable is *permanent* income, but only short-term or annual incomes are available for estimations. Second, although this literature is primarily descriptive, our approach can be easily adapted to include covariates to control for potential mechanisms through which parental income is correlated with child income, which has not been systematically done in the previous literature, especially when dealing with nonlinearity.

To facilitate the comparison to the existing literature, we also use data from Panel Studies of Income Dynamics (PSID), a commonly used data in the literature. As noted in Corak and Heisz (1999), "[t]he data for all of the U.S. surveys is based on only two different surveys (either the Panel Study of Income Dynamics or the National Longitudinal Survey)." We reach several important conclusions, some of which are new findings in the literature. First, we uncover that the association between single-year and permanent incomes varies over the life cycle. Specifically, the single-year income in younger years generally underestimate the permanent income, while the opposite is true for the single-year income in one's later stages. In addition, the life-cycle patterns has changed between generations (fathers and children). Second, we find that there exists substantial heterogeneity in IGMs across population. The mobility function is nonlinear. The degree of mobility differs greatly at the two extremes of the earnings distribution. Specifically, there is a considerable degree of mobility among the broadly defined middle earnings group, but both the children of high and low earning fathers are more likely to grow up to be, respectively, high and low earning adults. These results survive a battery of robustness checks. Finally, we also find evidence of substantial within-group (as opposed to between-group) heterogeneity in IGMs, which is previously undocumented.

Our methodology is related to the literature on nonlinear models with measurement error. For nonlinear models with classical measurement error, Li (2002) and Schennach (2004a,b) use double measurements of explanatory variables for identification and estimation, Schennach and Hu (2013) show nonparametric identification and semiparametric estimation without side information of explanatory variables. On the other hand, if explanatory variables are subject to non-classical measurement error, Hu (2008) and Hu and Schennach (2008) provide general identification results for nonlinear models with misclassification and measurement error, respectively, using instrumental variables. The methodology of our paper is closely related to and based on the results in Hu and Sasaki (2014), who propose closed-form estimators for nonparametric regressions using two measurements with non-classical errors. Our approach differs from Hu and Sasaki (2014) in several aspects. First, we allow for the case that both dependent and explanatory variables may be subject to measurement errors; whereas only explanatory variables are mis-measured in their model. The impacts of mis-measured dependent variable on estimation of nonlinear models are no longer trivial when the measurement error is nonclassical. Second, we relax the normalization assumption, which is typically assumed in the broad literature on measurement error or factor models to bound the latent function. This is a novel result in the literature.

The next section has a lengthy discussion of the literature, where we highlight both important and some of the more subtle issues that motivate our method. Section 3 lays out the theoretical results about our method and Section 4 conducts some Monte Carlo simulation exercises to assess its finite-sample performances. Section 5 describes the data and Section 6 presents the results. Section 7 discuss the further implications of our results. Finally, Section 8 concludes this paper. Proofs, tables and figures are included in the Appendix.

2 Related Empirical Literature

Due to its importance, the literature on estimation of IGM measures is abundant. Solon (1999), Björklund and Jäntti (2009), and Black and Devereux (2011) provide excellent reviews of a substantial body of work on this topic. Here we will focus only on the empirical issues that motivate our approach, as well as the limitations common to this literature including this work.

2.1 Issues with Different Approaches

The literature has typically focused on the following benchmark regression:

$$Y^* = \alpha + \beta X^* + \epsilon$$

where Y^* and X^* are the log of permanent income of children and parents, respectively. In this log-log specification, β is the parameter of interest, often called intergenerational elasticity of income (IGE). It is *the* most commonly used summary measure of IGM due to its easy interpretation much like a correlation (Lefgren et al., 2012; Durlauf and Shaorshadze, forthcoming). Solon (1992) acknowledges the fact that permanent income is unobservable; and that simply replacing it with observed annual income can thus lead to a measurement error problem. In the case of the classical measurement error (i.e., $X_t = X^* + U_t$ where t is age and U_t can be considered as exogeneous transitory shocks), a well-known result is that OLS estimates of β are biased downward to zero, thereby leading one to find a smaller IGE and hence overstate the level of intergenerational mobility. For the linear specification it is relatively easy to address the measurement error problem. One of the solutions proposed by Solon (1992) is to average father's observed annual incomes across *all* time periods.[3] Solon (1992) finds that addressing the measurement error problem indeed leads to a larger estimate of IGE, which can reconcile the divergent views to some extent. Solon's averaging method has been traditionally followed in the subsequent literature (Black and Devereux, 2011).

An implicit assumption in the averaging method is that the association between annual and permanent incomes is the same over the life cycle. This assumption has been challenged in Solon's later work (Haider and Solon, 2006) and others (Grawe, 2006; Böhlmark and Lindquist, 2006; and Nilsen et al., 2012). By considering a non-classical measurement error framework, Haider and Solon (2006) show that annual incomes in one's early career tend to understate one's permanent income, but this downward bias (also known as life-cycle bias) becomes less severe as one matures. This issue has since often been acknowledged, but we are not aware of any papers that systematically deal with this problem in estimation of IGE. In their work, Haider and Solon (2006) derive bias formulae based on the linear specification to correct for life-cycle bias, but their method, as noted in Grawe (2006), actually assumes that we have data on permanent income, which is clearly unavailable![4]

[3] Another solution proposed in Solon (1992) is the instrumental variable (IV) approach. See II. Section E in Mazumder (2005) for a brief summary of the existing results using this approach. However, this approach has not been applied widely, presumably because it is not a trivial task to find a valid IV.

[4] Additional distributional assumptions are also required for estimation of permanent income in Haider and Solon (2006). Grawe (2006) discuss inferential issues inherent in their multi-step approach.

Moreover, Haider and Solon (2006) also implicitly assumes that the life-cycle bias remains constant over generations, which can fail to hold in practice.

Further, measurement error in child's permanent income has often been ignored in this literature since, in the linear context, classical measurement error in child's income can affect only the precision, but not the consistency of the estimates. This log-log specification has also been shown to be sensitive to inclusion of very small income values (Chetty et al., 2014a).

It is obvious that the benchmark regression, assuming a constant IGE (β) across population, cannot accommodate heterogeneity. A typical solution to allow for heterogeneity/nonlinearity is to employ the transition-matrix approach (e.g., Zimmerman (1992) and Corak and Heisz (1999)). This approach calculates transition probabilities to describe the rates of movement across specific parts (e.g., quartiles) of the distribution over a generation. While useful to capture heterogeneity to some degree, it also has two important shortcomings: one is related to measurement errors and the other to its "overtly disaggregate nature".

First, Bhattacharya and Mazumder (2011) note that in the presence of measurement errors, the results using the transition matrix approach will hold only if the ranks of individuals are preserved despite measurement errors. This assumption is often restrictive and has been shown to fail to hold in practice. O'Neill et al. (2007) show via limited simulation exercises that measurement error can bias the estimates as much as 20% in some cases; and that the bias is most severe in the tails of the distribution (which are often of main interest to policymakers and alike). However, there has not been a systematic way to deal with measurement error in this framework.[5] Notable exceptions include Corak and Heisz (1999) and Bhattacharya and Mazumder (2011), but again using the averaging method.

Another substantive shortcoming of the transition matrix approach is "its overtly disaggregate nature" (Bhattacharya and Mazumder, 2011). One needs to discretize the support of the income distribution, and only a finite number of income categories is used in practice (sometimes as few as four groups e.g., Zimmerman, 1992). An immediate consequence is that mobility depends on the chosen threshold for an income group, and that any mobility within a particular group is ignored. This could also produce an artificial nonlinearity where people in the very top and very bottom of the father's income distribution tend to have higher probabilities to stay in the same income category, simply because those at the very top are restricted from further upward movement and those at the very bottom further downward movement; the so-called floor-ceiling effect associated with the transition-matrix approach (Corak and Heisz, 1999). More importantly, the estimated transition probabilities are not directly comparable to IGEs. For example, if the estimated probability from the second to the top quartiles is 25%, this figure cannot tell us by how much the child's income will increase should the father's income increase by a certain amount, as IGE does. Furthermore, Hertz

[5] O'Neill et al. (2007) discuss several papers on other topics that attempt to account for measurement error.

(2005) points out that with transition matrices, "there is no best way to summarize their content", prompting development of an easier-to-interpret summary measure of mobility. Our approach based on continuous measures of income allows a simple metric of persistence.

Other approaches used to address heterogeneity/nonlinearity such as spline regressions (e.g., Björklund et al., 2012), standard nonparametric estimations (e.g., Corak and Heisz, 1999), and rank-rank regressions (e.g., Chetty et al., 2014a) share some similar issues with the transition-matrix approach.

Considering the foregoing issues, we propose a nonparametric approach to address heterogeneity, while acknowledging the latent nature of both child and parental permanent incomes. Unlike the averaging method, our approach does not require that *all* observed annual incomes be accurate representations of permanent income. Neither do we require knowledge of permanent income to address the life-cycle bias in estimation. To anchor our approach to the past studies, we choose a nonparametric log-log specification since IGE is the most commonly used measure of IGM and is easy to interpret.

2.2 Common Limitations to this Literature

Prior to continuing, it is also important to point out some of the limitations common to most of the literature including ours. First, the literature is largely descriptive. Most of the literature is still focusing on obtaining precise measures of IGEs, although increasing attention has been paid to causal mechanisms underlying the observed patterns (Black and Devereux, 2011). While some attempts in this direction have been made, they are still far from satisfactory. Notable examples among others include Björklund et al. (2012), Chetty et al. (2014a), and Lefgren et al. (2012), which, for example, consider the impacts of parental income on several potential intermediate outcomes. Fagereng et al. (2015) isolate the impact of parental income from genetic differences in abilities by using a sample of adoptees. The goal of this paper is to provide credible estimates of the patterns of IGEs, which is a necessary step before any research can be done finding the potential mechanisms.

Second, despite the number of studies on this topic, most U.S. studies are basing on only two different surveys, either the Panel Study of Income Dynamics (PSID) or the National Longitudinal Survey of Youth (NLSY), both of which have some shortcomings. For example, the intergenerational samples based on these datasets can be small, especially for minorities such as blacks (e.g., Hertz, 2005; Mazumder, 2005; Corak and Heisz, 1999). Few studies have used alternative data sources such as SIPP matched to Social Security Adminstration data (Mazumder, 2005) and federal income tax records (Chetty et al., 2014a,b). These alternative datasets are generally not readily available to the public, and hence we utilize the PSID in this paper. Our use of this commonly used dataset facilitates direct comparisons to the existing results to highlight the flexibility of our approach.

Although all the data problems pertaining to the use of PSID/NLSY still carry through to our paper as well, we try our best to follow the standard procedure to deal with the shortcomings (discussed in more detail below). Moreover, there is also some simulation evidence that some of these issues (such as small sample sizes) may not necessarily impact our estimates severely.

3 Empirical Strategy

3.1 The Model

Consider a model that describes an intergenerational income transition:

$$Y^* = g(X^*, Z) + \varepsilon \tag{1}$$

where Y^* and X^* are respectively the permanent (log) income of children and parents. Z is an exogenous vector of covariates of childrens, e.g, race, gender and region; ε are other factors that might affect the intergenerational income transition that we cannot control for. $g(\cdot)$ is an unspecified mobility function relating parental income X^* to child income Y^*. This nonparametric specification nests the linear specification commonly used in the literature as a special case. The IGE at a realized $Z = z$ is captured by the derivative of the mobility function, i.e., $\partial g(X^*, Z = z)/\partial X^* |_{X^*=x^*}$, which is allowed to vary with the parental income X^*.

Note further that the econometrician cannot observe permanent income directly in the data. Instead, we can only have access to actual annual incomes of each individual at different stages of their lives, which are imperfect measures of the permanent income. Following Haider and Solon (2006), we can model the relationship between the observed and the permanent incomes (by incorporating age-dependent property due to heterogeneous age-earnings profiles):

$$\begin{aligned} Y_t &= \delta_t Y^* + U_t, \\ X_t &= \alpha_t X^* + V_t, t = 1, 2. \end{aligned} \tag{2}$$

where t represents the stage of each individual in her life cycle such as ages, Y_t and X_t are the observed incomes at period t for the two generations respectively, U_t and V_t are transitory shocks, and $\delta_t > 0$ and $\alpha_t > 0$ capture the weight of the income in age t in the permanent income (or how well annual income in a particular period approximates the permanent income); δ_t and α_t are often called life-cycle coefficients.[6]

This specification implies that the averaging method cannot reduce measurement error in the presence of non-classical measurement error. Consider a simple case where one is a permanent

[6]As will be shown, two periods of data are enough for identification and estimation.

income with a classical measurement error, and the other is not (i.e., $X_1 = X^* + V_1$, $X_2 = \alpha X^* + V_2$). The averaging method gives $\overline{X} = \frac{1}{2}(X_1 + X_2) = \widetilde{\alpha} X^* + \widetilde{V}$ with $E[\overline{X} - X^* | X^*] \neq 0$ where $\widetilde{\alpha} = (\frac{1}{2} + \frac{1}{2}\alpha)$ and $\widetilde{V} = \frac{1}{2}(V_1 + V_2)$. This implies that the averaging method, although it may reduce the impact of transitory income shocks, actually leads to a non-classical measurement error (or life-cycle bias), whose direction of bias is *priori* unknown.

3.2 Identification and Estimation of Mobility Function and IGE

The identification and estimation of the IGE is obtained by first identifying and estimating the mobility function $g(\cdot)$. Below we discuss how we can use repeated measures of the latent variables to nonparametrically identify $g(\cdot)$: in our case, the observed incomes X_t and Y_t from multiple years are repeated measurements of the permanent income X^* and Y^*, respectively. Intuitively, the joint distribution of the observed incomes from multiple periods reveals information of the distribution of the permanent income, whereas the joint distribution of parental and children's incomes further helps us recover the unspecified mobility function $g(\cdot)$. We consider both point and partial identification of the mobility function. For the sake of simplicity, we suppress z below whenever there is no ambiguity.

3.2.1 Point Identification

We first make the following observation:

Claim 1. *Not all the δ_t and α_t, $t = 1, 2$ are uniquely determined from the observed Y_t and X_t in model (1).*[7]

This claim implies that additional restrictions or normalizations are necessary for any attempts for point identification. Normalization is commonly required even in the linear case (see, e.g., Cunha et al., 2010; Black and Smith, 2006). For example, Black and Smith (2006) consider a linear GMM method to estimate the returns to the latent school quality. They normalize the variance of the underlying latent school quality to one. Such normalization is obviously not plausible in our context when the underlying variable is permanent income. As is commonly used in the previous literature, our normalization is instead on the coefficient on *only* one of the single year income. This normalization is more natural and economically plausible: it simply requires that there exists a stage

[7] This claim can be easily verified. We exemplify it using α_t. The original specification can be transformed to an alternative specification as follows: $X_t = \alpha_t X^* + V_t = \alpha_1 \alpha_t (X^*/\alpha_1) + V_t$. The two specifications would be observably equivalent with the permanent income being X^* and X^*/α_1 and the coefficients being α_t and $\alpha_1 \alpha_t$, respectively. Note that in Haider and Solon (2006) the coefficient α_t is identified as $Cov(V, X_t)/Var(X_t)$ under a simple dynamics for the observed income $X_t = a + bt$, where V is the present discounted value of lifetime earnings and $V \approx a + r - logr + \frac{b}{r}$ (r is a constant real interest rate).

of one's life where the annual income in this year is the permanent income plus some unobservable transitory shocks. Formally,

Assumption 1. *There exist a period* $t_0, t_1 \in \{1, 2\}$ *such that (i)* $X_{t_1} = X^* + V_{t_1}$ *and (ii)* $Y_{t_0} = Y^* + U_{t_0}$*, where* t_0 *and* t_1 *are not necessarily to be the same. We call* t_0, t_1 *baseline years.*

In the absence of any transitory shocks (measurement errors), one's income will eventually reach their permanent income at one point (i.e., classical measurement errors). There is strong evidence supporting this normalization.[8] Unlike the average method, we allow for the scale of the annual incomes in any other years to differ from permanent income with a non-classical measurement error ($\alpha_2, \delta_2 \neq 1$). For ease of notation, we denote the normalization period as 1 and the other periods as 2 for both children and parents.

To recover the mobility function, we first need to identify the life-cycle coefficient for the other period at least for parents by the results in the measurement error literature. We first provide the assumptions needed for identifying the other coefficients, i.e., δ_2 and α_2.

Assumption 2. *(i)* $Cov(U_t, X^*) = Cov(V_1, U_t) = 0$*, (ii)* $Cov(V_t, Y^*) = Cov(U_1, V_t) = 0$*, t=1, 2.*

Parts (i) and (ii) impose restrictions on parental and children's income, respectively. The first part of (i) is standard in the literature; it states that the parental permanent income is uncorrelated with the transitory shocks to parental income in both periods. The second part of (i) imposes restrictions on the transitory shocks in the baseline year (in which the observed income is equal to the permanent income plus transitory shocks). This assumption, while not necessarily less restrictive than the assumption of no serial correlation in transitory shocks that is typically assumed in the literature, is more plausible. To see this, suppose baseline years for both parents and children are age 36, and a second-year observation for children age 26. $Cov(U_2, V_1) = 0$ requires that the transitory shock to a child's income at age 26 is uncorrelated with any transitory shock to father's income at age 36. It is hard to imagine these two shocks will be correlated because they are at least 15 years apart (assuming father's age at his birth is 25). A similar argument applies to $Cov(U_1, V_1) = 0$. The discussion above can be readily extended to part (ii) for children.

Under Assumptions 1 and 2, δ_2 and α_2 can be identified as follows.

Lemma 1. *The coefficients* δ_2 *and* α_2 *are identified under Assumptions 1 (i), 2 (i), and 1 (ii) and 2 (ii), respectively.*

$$\delta_2 = \frac{Cov(X_1, Y_2)}{Cov(X_1, Y_1)}, \ \ \alpha_2 = \frac{Cov(X_2, Y_1)}{Cov(X_1, Y_1)}. \tag{3}$$

[8] For example, Haider and Solon (2006) estimate α_t and δ_t for the 1951-1991 Social Security Administration earnings histories of the members of the Health and Retirement Study sample. Assuming the life-cycle coefficients of children and parents are the same, i.e., $\alpha_t = \delta_t$, they show that at the age of 32 the coefficient crosses one and keeps insignificantly different from one till mid or late forties. Böhlmark and Lindquist (2006) provide a similar analysis based on the Sweden income tax data, and they find that the coefficient is closest to one at the age of 34.

Once α_2 and δ_2 are identified, we obtain repeated measurements for the permanent income X^* and Y^* which allow us to identify their distributions nonparametrically. Such an approach has been widely used in the literature of measurement errors. The basic idea is to investigate the identifying power of characteristic function (ch.f. hereafter) for a density function. Let R denote a random variable and $f_R(r)$ be its density evaluated at $R = r$.

To identify the mobility function, additional assumptions are needed and presented as follows.

Assumption 3. *(i) X^* is independent of V_2. (ii) $E[V_1|X_2] = 0$.*

Part (i) is a textbook assumption on transitory shocks, requiring the permanent income X^* to be independent of the transitory shock V_2 at period 2. Note that in the specification (2), we have $E[X_2 - X^*|X^*] = E[(\alpha_2 - 1)X^* + V_2|X^*] = E[(\alpha_2 - 1)X^*] \neq 0$ whenever $\alpha_2 \neq 1$. Thus the independence assumption in part (i) does not require the measurement error to be classical (i.e., $E[X_2 - X^*|X^*] = 0$). Part (ii) imposes mean independence of the transitory shocks at the normalized period from the remaining observed income. In practice, we have the flexibility to choose the period $t = 2$ so that this part of assumption can be easily satisfied. For example, given $t = 1$ being 36, we can choose $t = 2$ to be, say, 40 so that $E[V_1|X_2] = 0$ is plausibly met because it is likely that one's annual income at age 40 has no predictive power of the transitory shock to her income at age 36. Note that our Assumption 3 is less restrictive than the assumption of mutual independence of V_1, V_2, and X^* for parents, which is commonly imposed in the literature.

Lemma 2. *The density f_{X^*} is nonparametrically identified under Assumptions 1 (i), 2 (i) and 3.*

Proof. See Appendix. ■

The identification of the density f_{X^*} from two measurements of X^* follows the insights in Li (2002) where the information of the latent variable X^* is explored in the characteristic functions of its measurements. Similarly, we may identify the density of the children's permanent income f_{Y^*} under a similar restriction to Assumption 3. Nevertheless, as will be shown below, the density f_{Y^*} is not necessary for identifying the mobility function $g(\cdot)$.

Assumption 4. *(i). $E[\varepsilon|X_2] = 0$. (ii). $E[U_2|X_2] = 0$.*

Part (i) of this assumption is assumed in nearly all the studies of IGE, stating that those unobserved and random factors ε that affect intergenerational income mobility are mean independent of parental income X_2. Violation of this assumption simply means that we cannot necessarily interpret the result as a causal relationship. Part (ii) requires that there exists an age (not necessarily the same for parents and sons) where the transitory shock to children' income is mean independent of parental income.

The ch.f. of the density f_R is defined as $\phi_R(s) \equiv E[e^{isr}] = \int_{-\infty}^{\infty} e^{isr} f(r)dr$. Similarly, a joint ch.f. of two random variables R, R' is $\phi_{R,R'}(s_1, s_2) \equiv E[e^{is_1r+is_2r'}] = \int_{-\infty}^{\infty} e^{is_1r+is_2r'} f(r, r')drdr'$.

Assumption 5. *(i) ϕ_{X_2} does not vanish on $\mathbb{R}$. (ii) f_{X^*} and $\phi_{X^*}(\cdot)$ are continuous and absolutely integrable. (iii) The convolution of f_{X^*} and $g(\cdot)$, $f_{X^*} * g(\cdot)$ is continuous and absolutely integrable, so is $\int_{-\infty}^{\infty}(f_{X^*} * g)dx$.*

Part (i) is a widely imposed restriction in the literature on measurement errors, where characteristic functions oftentimes appear as denominators. Many commonly used distribution families satisfy this requirement, e.g., exponential, gamma, chi-squared, and normal distributions. Parts (ii) and (iii) are commonly used regularity conditions that enable us to apply the Fourier transform and inversion to those functions.

Proposition 1. [*Hu and Sasaki (2014)*] *Suppose Assumptions 1 (i), 2 (i), 3-5 hold, then the mobility function $g(X^*)$ is nonparametrically identified from two periods' income of parents and one period of children's income. Moreover, $g(X^*)$ has a closed-form.*

Proof. See Appendix. ■

Proposition 1 provides a closed-form solution to the unknown function $g(X^*)$ conditional on the covariate Z. The closed-form estimator does not rely on any optimization algorithms and it has several advantages over a MLE or a GMM: (i) A closed-form estimator is global by construction. By contrast, an optimization algorithm, e.g., MLE can only guarantee a local maximum or minimum even a global solution exists. (ii) It allows us to analyze how parameters affect the estimate constructively while this can only be done numerically for an estimator using optimization algorithms. (iii) A closed-form estimator is computationally more convenient since most of the optimization algorithms involve iterations. Also, note that our method may be preferred especially because of the paucity of long panel data. We need simply the baseline year and another year of the data for parents, whereas only one year data for children are necessary.

3.2.2 Partial Identification and Bounds

Following the literature, we have thus far considered a necessary normalization to *point* identify the mobility function. A useful advantage of our closed-form solution is that it allows us to provide informative bounds on the mobility function when such normalization is mis-specified. Below, we consider a scenario in which we have only some imperfect knowledge of the range of the life-cycle coefficient, as opposed to having precise knowledge of which period the coefficient equals to 1. Specifically, we allow the coefficient to vary around 1 within a plausible neighborhood. The following identification result on bounds without normalization is new and important in the literature.

Assumption 6. *Suppose the observed income Y_1 and X_1 in baseline years satisfy*

$$\begin{aligned} Y_1 &= \delta_1 Y^* + U_1, \delta_1 \in [\underline{\delta}_1, \bar{\delta}_1] \\ X_1 &= \alpha_1 X^* + V_1, \alpha_1 \in [\underline{\alpha}_1, \bar{\alpha}_1], \end{aligned} \tag{4}$$

where δ_1 and α_1 are unknown, but $\underline{\delta}_1, \bar{\delta}_1$, and $\underline{\alpha}_1, \bar{\alpha}_1$ are known and positive.

This assumption relaxes the normalization requirement in Assumption 1. The well-established empirical evidence in the literature provides with us reasonable bounds for α_1 and δ_1. For instance, the point-wise confidence interval of α_1, which is assumed to equal δ_1, in Haider and Solon (2006) can be used as a reasonable bound.

Below we show that δ_1 and α_1 affect the closed-form of the function $g(\cdot)$ in a particular way such that a bound of $g(\cdot)$ can be obtained given that $\underline{\delta}_1, \bar{\delta}_1, \underline{\alpha}_1$ and $\bar{\alpha}_1$ are known. We first present an important result concerning the functional form of $g(\cdot)$ under Assumption 6.

Theorem 1. *Suppose Assumptions 2(i) and 3-6 hold. Let $\widetilde{g}(\cdot)$ denote the mobility function identified based on Proposition 1 by falsely assuming Assumption 1 holds. Then:*

$$\widetilde{g}(x^*) = \delta_1 g(x^*/\alpha_1) \text{ or equivalently } g(x^*) = \frac{1}{\delta_1}\widetilde{g}(\alpha_1 x^*). \tag{5}$$

Proof. See Appendix. ■

This theorem expresses explicitly how an inaccurate relationship between permanent income and an observed one affects the mobility function. Interestingly, the coefficient for the children (δ_1) affects only the scale of the mobility function, and the effect is independent of the underlying mobility function $g(\cdot)$, whereas the effects of the coefficient for parents α_1 depend on $g(\cdot)$.

The results in Theorem 1 have some important implications. Specifically, if the *true* life-cycle coefficients for the baseline years are the same across generations (i.e., $\delta_1 = \alpha_1$) as in Haider and Solon (2006), mis-specified normalization does not bias the estimation of the true mobility function when the true function is either homogeneous with degree one or linear. First, when the true mobility function is homogeneous with degree one $g(x^*/\alpha_1) = g(x^*)/\alpha_1$ (e.g., a linear function without an intercept), the estimated mobility function is $\widetilde{g}(x^*) = \frac{\delta_1}{\alpha_1} g(x^*)$; hence $\widetilde{g}(x^*) = g(x^*)$. Second, if the mobility function is linear $g(x^*) = \beta_0 + \beta_1 x^*$ as widely used in the literature, the following corollary holds.

Corollary 1. *Suppose the latent mobility function is linear $g(x^*) = \beta_0 + \beta_1 x^*$, and the two generations follow the same relationship between permanent and observed income $\delta_1 = \alpha_1$, then the intergenerational income elasticity (β_1) is identified as $\widetilde{g}'(\cdot)$ under the conditions for Theorem 1.*

Proof. See Appendix. ∎

The conditions of linearity for the mobility function and the same income model for two generations are standard and widely imposed in the literature of intergenerational mobility of income. Thus, Corollary 1 implies a distinct advantage of our method in dealing with a latent linear mobility function, which is the approximation of permanent income does not affect our estimate of the function if the true mobility function is linear.

For more general mobility function, the results in Theorem 1 also enable us to obtain bounds for both the mobility function $g(\cdot)$ and the income mobility elasticities when the normalization assumption fails to hold.

Theorem 2. *Under Assumptions 2(i) and 3-6, the bounds of the mobility function and its derivative are identified to be*

$$\frac{1}{\bar{\delta}_1} \min_{\alpha_1 \in [\underline{\alpha}_1, \bar{\alpha}_1]} \left(\widetilde{g}(\alpha_1 x^*)\right) \leq g(x^*) \leq \frac{1}{\underline{\delta}_1} \max_{\alpha_1 \in [\underline{\alpha}_1, \bar{\alpha}_1]} \left(\widetilde{g}(\alpha_1 x^*)\right) \tag{6}$$

and

$$\frac{\underline{\alpha}_1}{\bar{\delta}_1} \min_{\alpha_1 \in [\underline{\alpha}_1, \bar{\alpha}_1]} \left(\widetilde{g}'(\alpha_1 x^*)\right) \leq g'(x^*) \leq \frac{\bar{\alpha}_1}{\underline{\delta}_1} \max_{\alpha_1 \in [\underline{\alpha}_1, \bar{\alpha}_1]} \left(\widetilde{g}'(\alpha_1 x^*)\right), \tag{7}$$

respectively.

Proof. See Appendix. ∎

Theorem 2 states that our method provides bounds for the nonparametric mobility function when Assumption 1 fails but some prior information about the range of the coefficient may be available from the literature.

3.2.3 Estimation

The identification results on the mobility function and its bounds are constructive. This allows us to follow the identification argument in estimation. Given an observed sample of the two generations' incomes for two years, the mobility function is estimated in three steps. We sketch the steps of estimation and leave the technical details in the Appendix.

1. First, the coefficients α_2 and δ_2 are estimated using the sample analog of equation (3).

2. Second, we recover the density of the permanent income f_{X^*} basing on the estimated α_2, together with the closed-form expression provided in Lemma 2.

3. Finally, we estimate the mobility function by plugging the estimates in the first two steps into the closed-form mobility function in Proposition 1. The estimate of the mobility function also allows us to further estimate the derivative as well as the bounds (basing on Theorem 2).

As discussed before, the estimation of mobility function as well as its derivatives only requires two years of income data of parents and one year of income data of children. In the case where many years of data are available for both generations, one may obtain multiple mobility functions and use the average of those estimates as the estimator. Such an approach may improve the precision of the estimate.

4 Monte Carlo Experiments

In this section, we design several Monte Carlo studies to illustrate the finite-sample performance of our estimators. We assume that both α_t and δ_t are random draws from a uniform distribution $U[0.2, 1.2]$. They are not necessarily equal for a given t. The transitory shocks U_t, V_t, and the structure error ε are i.i.d. draws from the standard normal $\mathcal{N}(0, 1)$. Since our estimation procedure is conditional on Z, we need not worry about covariates in our Monte Carlo studies.

We consider two specifications of the function $g(X^*)$:

$$
\begin{aligned}
g(X^*) &= 0.6X^*, \\
g(X^*) &= X^* - 0.2X^{*2}, \qquad (8)
\end{aligned}
$$

where X^* is distributed according to a normal distribution $\mathcal{N}(0, 4)$. The number of periods $T = 2$ and the sample size $N = 100, 300, 500$ (we only present the results of $N = 300$ for brevity). We replicate each experiment for 1000 times.

In the first set of experiments, we maintain the normalization assumption $\alpha_1 = \delta_1 = 1$ as in proposition 1. We present the estimates of α_2, δ_2 in Table 1. As an intermediate step, we also estimate the densities for X^* and Y^* and report the results in Figure 1, where the top panels are for the linear specification, while the bottom ones for the quadratic one. In each subfigure, we present the true density, the estimated mean and the [5%, 95%] point-wise confidence intervals. The figure illustrates that our estimates perform very well even for a modest sample size of 300 observations, which is smaller than the sample size used in most of the existing studies in this context. The estimates of the mobility function $g(\cdot)$ are presented in the top panel of Figure 2 where the linear specification is shown on the left and the quadratic on the right. The results illustrate that the estimated function captures the true function closely for both the linear and quadratic cases. It is worth noting that the estimates at both tails are noisy, because the number of observations at tails is small for normal distributions. Nevertheless, the estimates, especially the point-wise confidence intervals, improve significantly as the sample size increases. Based on the estimated function $g(\cdot)$, we employ equation (B.3) in the Appendix to estimate the derivative $g'(\cdot)$. The results are presented in the middle panel of Figure 2.

In the second set of experiments, we consider estimation of bounds for the true function when α_1 and δ_1 are mis-specified. Specifically, we set $\alpha_1 = 0.8$ and $\delta_1 = 1.2$ but pretend that these values are unknown yet with knowledge that $\alpha_1 \in [0.6, 1]$ and $\delta_1 \in [0.9, 1.5]$. Following Proposition 2, we estimate the bounds of $g(\cdot)$. The estimates are displayed in the bottom panel of Figure 2, where the left one is for the linear specification. We observe that the estimated bounds (blue lines) successfully contain the true mobility function. We also use these bounds to obtain an interval of the marginal effect of x^* on $g(\cdot)$. We now turn to the empirical application of our proposing method.

5 Data

Our data are drawn from the Panel Study of Income Dynamics (PSID), a longitudinal survey consisting of a nationally representative sample of over 18,000 individuals living in 5,000 families in the U.S.. The original sample has been re-interviewed annually starting from 1968 through 1997 and biennially afterward. The PSID has also continued to follow children from the original sample even after they left their parents to form their own households. This allows us to match children to parents and obtain their incomes at different ages. More importantly, the PSID collects data on annual income for the preceding calendar year. Because of these advantages, the PSID has been widely used for empirical studies of IGM, despite its relatively small sample size. We therefore provide only limited detail here.

Following the standard practice in the literature, we focus on (the logarithm of) family income for both parents and children, as it is "the more inclusive measure of economic status and preferable for most applications" (Bowles and Gintis, 2002); and "reflects the objective of measuring mobility after accounting for secular income growth in a society" (Durlauf and Shaorshadze, forthcoming). See, also, Chadwick and Solon (2002) and (Lee and Solon, 2009, p.767) for more discussions of this point. In our analysis, we exclude negative income and adjust for top-coded earnings by a factor of 1.4, as in Lemieux (2006).

Because of the use of family income for both parents and children, we cannot separate the income of children from that of their parents' if children live with their parents. Following the literature (e.g., Mayer and Lopoo (2004, 2005); Lee and Solon (2009)), we measure children's income at different ages by the family earnings in the household in which they become the head or head's spouse. Since family income is used, we need only map the income of children to that of their fathers because mothers' incomes are the same as that of fathers.

As in Solon (1992), we use only the Survey Research Center component of the PSID, but exclude the Survey of Economic Opportunity (SEO) component to prevent over-representing the poverty sample. Previous literature has also pointed out some serious irregularities in the sampling of SEO respondents, which can "preclude easy generalization to any well-defined population" (Bloome, 2015).

We consider information on income for both children and parents at their ages from 26 to 50. There are two things about the number of observations available for our analysis that we need to keep in mind; these figures will become useful when making our choice of baseline and alternative measurement of incomes in our estimations (Table 2). We note that the number of observations available for children on income decreases with age monotonically, while the number of observations for parents exhibit an inverted-U shape. The number of observations would decrease further in estimations because we require information on income be available for both child and parent. Thus, the combination of child's and parent's ages that we choose will determine the sample size. We discuss our choice further below.

6 Results

6.1 Life-cycle Bias

We begin by showing estimates of association between annual and permanent incomes over the life cycle for both fathers and children. As shown in the prior literature, if the association varies with age, the use of annual income at different ages in traditional estimation can thus result in what is called life-cycle bias. An insight from Haider and Solon (2006) is that with complete information on earnings over a *full* career, it is possible to estimate permanent income and hence the corresponding associations over life-cycle (by simply regressing annual income at certain age on the permanent income). Such discussion is, however, only hypothetical as noted in Grawe (2006): if the data on the permanent income were available, we can simply regress child's permanent income on parental permanent income, and the estimates are free of life-cycle bias.

Our estimation is basing on observed annual incomes, and hence does not require data on permanent incomes. As mentioned previously, we normalize the association between annual income observed in one of the years and permanent income to be one. In other words, there must be one year when one's income reaches her permanent income (with some random transitory shocks). This normalization is economically plausible. Moreover, such normalization is innocuous in terms of interpretations. The estimated associations can be interpreted as *relative* association between permanent and annual incomes (or the (relative) importance of the income in each stage of the life cycle).

The previous literature has typically found that the associations become close to 1 between individuals' mid thirties and forties (e.g., Haider and Solon, 2006; Böhlmark and Lindquist, 2006; Grawe, 2006). As such, we choose $t_0 = t_1 = 36$, the first year during this period to illustrate our approach. Figure 3 present the estimates for the coefficients of α_2 and δ_2 when the income of different years ($t = 2$) are used for estimation. Examining the patterns for fathers, we reach quite similar conclusions to the previous literature. There exhibits an inverted-U shape in the relationship between association and age. We first notice that estimates are generally smaller than one before mid-thirties. Such differences are statistically significant since the bootstrapped 90% confidence intervals generally exclude one. This result indicates that relative to the baseline year, the income in one's early years are less precise in capturing her permanent income. During this period, the relationship between the association and age is nearly monotonic: as one gets older, her annual income gradually becomes closer to her permanent income. Afterward, the estimates remain relatively stable around 1 until mid-forties. Starting from one's late forties and on, the estimates are generally again smaller than one. This pattern certainly rejects the assumption of equal weights often imposed in the literature when using the average method. Consistent with the literature, these estimates suggest that life-cycle bias in IGEs for linear regressions could be more likely when fathers' incomes aged younger than 35 are used in estimations.

Turning to the estimates for children, we observe a slightly different pattern. While we continue to observe that the estimates are generally smaller than one before mid-thirties, these estimates are not necessarily statistically different from one. Moreover, there exists no inverted-U observed as for fathers. The estimates generally fluctuate around one. Comparing the results between children and fathers, we uncover the *relative* representation of the permanent income has changed between fathers and children. The weight of observed income relative to the permanent income (α_2, δ_2) changes not only over life cycle but also across generations. Also, we may conclude that at least in the case of linear framework, life-cycle bias is more likely a result of use of fathers' income at early ages.

6.2 Baseline Results

We now turn to our baseline results of IGEs that use the pooled sample. To illustrate our approach, the baseline year is set to be age 36 for both children and parents, as discussed above. The second period is set to be age 26 for children and 38 for fathers. The choice of different periods for children and fathers is motivated by two reasons. We first want to maximize the number of child-father pairs for the cohort of children born between 1952 and 1974, a cohort typically used in the previous literature (to facilitate comparison). Recall that we need information on incomes for *both* parent *and* child available not only for the baseline year, but also for the second year. This data requirement reduces the sample size. To visualize the patterns, we present the number of observations associated with each combination of the second years for children and parents in Table 3 and display these results in Figure 4 in the appendix. There are two distinct patterns of how the sample sizes vary with age for both parents and children when the baseline year for parents is 36. Given the choice of age for parents, the sample size decreases with child's age nearly monotonically. Given the choice for children, the sample size exhibits an inverted-U shape peaking at around 36. Thus, the combination of age 26 for children and 37 for parents appears to be the best choice. Recall, however, that our assumptions require that the baseline and second years should not be too close to each other. We thus choose 38 for parents instead, which, along with 26 for children, gives us the second largest sample for our empirical analysis.

Turning to our IGE results, we uncover substantial variation in the intergenerational mobility elasticities across population. To visualize this finding, we plot the histogram of the derivatives of the estimated mobility function in the left panel of Figure 5 and report quartiles (along with their corresponding bootstrapped 90% confidence intervals) in Table 4. We first note that similar to much of the previous literature, the median elasticity is 0.263 with 90% bootstrapped confidence interval $[0.124, 0.514]$. However, the elasticity from our nonparametric model is clearly not a constant. Specifically, the interquartile range, a measure of dispersion of the distribution, is as large as 0.353. The majority of the estimated IGEs lie in the neighborhood [0.130 (the first quartile), 0.483 (the third quartile)], covering the estimates typically found in the literature examining the same cohort of children.[9]

To further examine how the IGEs differ by parental income, we display the IGEs (i.e., the derivatives of the estimated mobility function $g(\cdot)$) evaluated at each possible log income level of parents in Figure 5 (the right panel). The solid line is our nonparametric estimates of IGEs, and the blue ones are 90% bootstrapped confidence intervals.[10] For comparison purposes, we also estimate the intergenerational income elasticity using the traditional linear approach (which is the horizontal line). The conventional OLS approach yields an estimate of the intergenerational elasticity of 0.231. This number falls into the range of the estimates typically found in the previous literature using PSID. However, this estimate drastically conceals important features of the nonlinear relationship in incomes across the generations. Our nonparametric results in Figure 5 reveal a nearly U-shape pattern of the IGEs. Specifically, we find that the income associations are high for the very poor, but the degree of associations quickly decreases as parental income increases and becomes relatively weak and closer to the average IGE for the broadly defined middle-income class (whose log income is between 8.5 and 11.5). Afterward, the income associations across generations quickly increase

[9] Note that as in the influential work by Corak and Heisz (1999), we also find that our nonparametric estimates cover some negative values, but none of them are statistically significant.

[10] The estimates for extremely low (log income less than 8) and high (log income greater than 12) parental income are excluded to facilitate presentation because they have been shown to be noisy and less reliable as illustrated in the Monte Carlo simulations.

and become high again in the top of the distribution of parental income. Note that the estimates are not significantly different from zero for log income from 8.6 to 10, which might be due to the small sample size.[11]

Our results corroborate previous findings of significant nonlinearity in the relationship in incomes across generations. Similar to Corak and Heisz (1999) and Björklund et al. (2012), we find that the degree of lack of mobility is concentrated in the extremes of the earnings distribution, while there is a considerable degree of mobility among the middle class. The finding of "twin peaks" is also prevalent in the studies using the transition matrix approach (e.g., Hertz 2005). Comparison of OLS and our nonparametric estimates highlights that the average IGE is drastically different from the IGE across the population. This finding indeed reconciles the divergent views discussed in the introduction that motivates our paper.

While largely similar to the previous studies examining nonlinearity, two distinct features of our results emerge. First, we find that there exists pronounced persistence in *both* the top *and* the bottom of the income distribution in the U.S.. This result is in contrast to Björklund et al. (2012) using Sweden data, where the authors find the marked persistence only in the very top of the income distribution, but much mobility in the lower tail (see their Figure 1) even more than other parts of the income distribution. Second, our results indicate mobility is potentially a more severe issue in the U.S. than in Canada. Although Corak and Heisz (1999), using Canadian data, find a similar nonlinear pattern as ours, their estimates are generally in smaller magnitudes. Specifically, out of three nonparametric specifications in Corak and Heisz (1999), the maximum value of the IGE (Figure 4 in their paper) is less than .8 among the top of the income distribution and close to 0.4 at the bottom of the distribution. By contrast, we find that the elasticity is around one at both extremes of the income distribution, indicating a close to complete transmission of income from parents to children (similar to Björklund et al. (2012)). Further, for very small proportions of the income values in the very top (whose log income is close to the maximum of 12), we find that the estimated IGE is even slightly greater than 1, although the difference is not statistically significant. Taken literally, our estimates paint a pessimistic picture: the very poor cannot escape the poverty trap, while the very rich continues to enjoy at least as much success as their parents.

6.3 Robustness Analysis

In the baseline results, we include all children regardless of their gender or race. For these results, we also normalize the association between annual income in age 36 and permanent income to one, but choose annual incomes in certain ages as the second measurement of permanent income. It

[11] Please see the left panel in Figure 6 for the density of parental permanent income, where the density from 8.6 to 10 is very small. The densities are estimated using the result in Lemma 2. The data used are parental income at 36 and 38, children income at 26 and 36.

is natural to ask whether our results are sensitive to these choices. In this section we assess the robustness of our results to the use of alternative second measurements, alternative normalization, relaxation of normalization, and different sub-samples. We find that the basic pattern uncovered above is generally robust to these changes. Recall that IGEs are the derivatives of the estimated mobility function. Any small changes in the mobility functions can result in significant changes in the derivatives. The fact that they look similar indicates the robustness of our results and increases our confidence in our estimates.

Prior to continuing, it is worth emphasizing that although the simulation exercises above have indicated superior finite-sample performances of our approach, it is still a nonparametric approach, which is data demanding. Many of the decisions, when illustrating our approach, are made so as to obtain a reasonably large sample of the particular cohort of interest and hence more reliable estimates.

6.3.1 Alternative Second Measurements of Permanent Income

Given the number of annual incomes observed across ages available for both child and parent, there are many possible combinations of the second measurements of permanent incomes for children and parents. Here we experiment with two types of alternative combinations, while maintaining that the base year is 36. In the first experiment, we set the second period to be age 26 for children and 39 for fathers. The idea behind this experiment is to see whether or not our results are robust to small departures from the original setting. The results are displayed in panel (a) of Figure 7. As we can see, we again find a nearly U-shape in the relationship between child and parent incomes. This result confirms our earlier finding of a lack of mobility at the both upper and lower tails of the income distribution, but a much higher degree of mobility in the middle.

In the second experiment, we set the second period to be age 30 for children and 40 for fathers. The idea behind this experiment is to use the years that are father away from the original setting, while still keeping a reasonably large amount of observations; this can make sure that the difference in estimates is not a result of large difference in the data. Only such comparison is meaningful. The results are displayed in panel (b) of Figure 7. We continue to find that there exists a nearly U-shape in the relationship between child and parent incomes. As is in the first experiment, we observe a large persistence of incomes across generations for *both* the very top and the very bottom of the income distribution. Comparing both experiments, we can see that the basic pattern is rather similar to what we find in the baseline results. Using the second measurements that are farther away from the original combination, we can see that the estimates of IGE at the parental (log) income very close to 10 become much noisier because of reduction in the number of observations, which is a pattern expected of our experiments.

We further conduct an experiment by averaging our baseline results and the robustness results in this subsection. In theory, because all these results are consistent, averaging them can potentially lead to some efficiency gain. This may not hold in practice, however, because the use of different years can reduce the sample sizes. To evaluate this claim, we present the average of three results with the second measurement being: (1) 26 (children) and 38 (parents), (2) 26 (children) and 39 (parents), and (3) 30 (children) and 40 (parents) in panel (c) of Figure 7. As we can see, while there does not appear to be significant efficiency gain, the main result continues to hold.

6.3.2 Alternative Normalization: Average Income

Earlier we borrow the insights from the literature on the life-cycle bias and consider annual income at age 36, the first year of the period that the literature has generally found to be good representation of permanent income. Here we instead consider the *average* of annual income in several adjacent years as permanent income. Specifically, we normalize the association between the three-year average of annual incomes in age $35, 36$ and 37 and permanent income to one. This exercise is motivated by the averaging method that the literature typically uses to address the life-cycle bias. Remember that even if the normalization is true, the three-year average income is only a measurement of permanent income with error. Supplementing it with second measurements as above (26 for children and 38 for fathers), our approach can be seen as an improvement of the averaging method to take into account the life-cycle bias. The results are displayed in panel (d) of Figure 7. We again find that estimated IGEs are much higher for both the top and the bottom tails of the income distribution, while estimates are much moderate in the middle; and that there are small proportions of the individuals who continue to move either downward and upward, depending on their parents' original situation.

6.3.3 Bounding Analysis: Completely Relaxing Normalization

Normalization is required for any methods in this context to *point* identify the model. However, we are able to bound the true mobility function when such normalization is relaxed. Using the results in Theorem 2, we estimate the bounds of the mobility function given various degrees of departures from the normalization. In other words, we require only that the annual income in the baseline year is *plausibly* (as opposed to *exactly*) good representation of permanent income.

For this exercise, we focus on fathers because the results above have shown the estimates of the life-cycle association for children are relatively stable over the life cycle. We experiment with three different choices of $\underline{\alpha}$ and $\bar{\alpha}$ (we assume $\underline{\alpha} = \underline{\delta}, \bar{\alpha} = \bar{\delta}$): $[\underline{\alpha}, \bar{\alpha}]$ being $[0.98, 1.02]$, $[0.96, 1.04]$ and $[0.95, 1.05]$, respectively. These neighborhoods correspond to small, moderate, and large departures from the "plausibly" good representation of permanent income. Their corresponding results are displayed in Panels (1), (2), and (3) of Figure 8, respectively. In all these figures, we contrast the bounds (along with our estimated mobility function) with a linear mobility function (the red line).

As we can see, we can clearly reject a linear mobility function in favor of a nonlinear one when the amount of violation is small. As the degree of the violation of normalization increases, we are less likely to reject the linearity as expected. However, this finding is mainly concentrated in the upper tail of the parental income distribution. We continue to reject linearity easily in the lower tail of the parental income distribution even when the degree of violation is relatively large. The bounding analysis here does not allow us to point identify the mobility function and corresponding derivatives (i.e., IGEs), but it does provide significant evidence supporting the nonlinearity and heterogeneity uncovered above.

6.3.4 Sub-sample Analysis

Excluding non-Whites As is in most of the previous literature, our baseline analysis includes children of all races. This does not take into account the fact that the transmission of income over generations may differ by races, and that there may exist permanent differences in opportunities between families of different ethnic groups (Durlauf and Shaorshadze, forthcoming). Consistent with this possibility, some recent research has shown that the degree of intergenerational mobility does differ by races. For example, Chetty et al. (2014a) find that the mobility is significantly lower in areas with a larger share of African-American population. Bhattacharya and Mazumder (2011) find more upward mobility for whites than blacks. As such, we exclude non-whites and repeat our analysis. The results are displayed in the left panel of Figure 9. The baseline result continues to hold when we focus only on whites. This is probably not surprising because PSID has a relatively small sample of non-whites, and thus exclusion of these observations should not drastically impact our estimates.[12] Furthermore, we also find that there exist substantial heterogeneity even within whites. Specifically, the interquartile range is about $.51 \approx 0.618$(the third quartile) $-$ 0.110(the first quartile), much greater than the interquartile range for the whole sample.

Excluding Females One may be concerned that income is not necessarily a good indicator of daughters' economic success because of relatively limited labor force participation of women (Mazumder, 2005). How selection into labor force affects estimation of IGE depends on the direction of selection, which can be correlated with fathers' income (Mazumder, 2005). For example, if it is positive selection, i.e., higher-income women who are more likely to come from higher-income families would work, we are more likely to observe higher persistence only in the upper tail of the income distribution, but not in the lower tail. There are two implications of positive selection for estimation. First, it is, on average, more likely to observe IGE of smaller magnitudes for women than for men. This is indeed what the literature has found; the average IGEs for daughters are

[12] Bhattacharya and Mazumder (2011) note that the intergenerational samples of blacks in PSID are indeed so small that it may likely inhibit research for blacks.

often in the 0.1 to 0.2 range using NLSY and in the vicinity of 0.4 using PSID (Mazumder, 2005).[13] Second, inclusion of daughters in our nonparametric analysis will more likely bias the estimates of IGE in the lower tail of the income distribution. To verify this, we exclude daughters in our analysis and the results are displayed in the right panel of Figure 9. And we indeed find that there is much larger persistence in the lower tail for sons, indicating lack of mobility. However, we continue to find the basic pattern in the IGE as above. Furthermore, we again find that there exist substantial heterogeneity even within sons. Specifically, the interquartile range is about $.24 \approx 0.392$(the third quartile) $- 0.153$(the first quartile), although smaller than the interquartile range for the whole sample.

Summary Prior to continuing, it is useful to mention that race and gender are by no means the only ways to split the sample, but they are the categories that the previous literature has typically considered. There could, of course, be many other between-group heterogeneity to consider as well. For example, the most recent paper by Chetty et al. (2014a) also finds substantial heterogeneity in IGEs across geographic areas. However, further sample splits can drastically reduce the number of observations in each group and hence result in imprecise estimates.[14] We leave more exercises of this kind to future research when larger datasets are readily available.

Two main messages are nevertheless clear. First, the nonlinear, U-shape pattern continues to hold, indicating substantial heterogeneity in IGEs. Second, such heterogeneity exists even within a more narrowly defined group. This *within-group*, as opposed to *between-group*, heterogeneity in IGEs is also a new finding, which is usually masked by the traditional approach.

6.4 Interpretations of the Main Result

A major question that emerges from our analysis is, what can explain the U-shape or twin-peak pattern in IGE, where children are stuck in poverty and affluence? There are several different mechanisms through which parental income can affect child economic success. For example, income can be spent on resources that are passed on to children in the form of expenditures on human capital (or various skills), financial assets, and wealth bequest as well as inter vivos wealth transfer. Families with different incomes may invest differently in their children; the existence of nonlinearity, particularly the "twin peaks", suggests that distinct transmission mechanisms may be at work at

[13]Chadwick and Solon (2002) find that the differences between daughters and sons are not always statistically significant. This may be due to a much smaller estimation sample of daughters, and the inference is not as reliable. This is indeed what we find as well: the estimates are much smaller for daughters than for sons, but the estimates for daughters are very imprecise. The results are available from the authors upon request.

[14]We indeed carry out such exercise by region, but find that most estimates are imprecisely estimated. The estimates, however, do continue indicate a similar U-shape pattern for the Western region, while the nonlinearity is relatively modest for the rest of the country.

different parts of the distribution of parental income (Bowles and Gintis, 2002).

The larger IGE for the families with very little income is consistent with the presence of credit constraints. In the standard intergenerational mobility models (e.g., Becker and Tomes 1986; and Mulligan 1997), families cannot fully borrow against the future incomes of either their own or children to finance child human capital investment. To the extent that credit constraints are most severe for families at the very low tail of the income distribution, these families may not be able to attain the optimal amount of human capital investment for their children.[15] As a result, income is positively related to human capital investment, and the IGE would also be much higher for the children of these families. Indeed, there has been enormous evidence that income is positively related to education and a broader conception of cognitive and non-cognitive skills, suggesting the presence of credit constraints. In line with our results, Cunha et al. (2006) find that permanent income, as opposed to transitory income, matters for the accumulation of human capital. Moreover, Carneiro and Heckman (2002) suggest that only 8% of American households are actually credit constrained. This finding is consistent with the fact that we observe a higher IGE only for the families at the very low tail of the distribution.

Unlike that for the lower tail of the parental income distribution, the borrowing constraint for human capital investment becomes less binding for the majority of the population. Instead, other channels may matter more for the transmission of income. Bowles and Gintis (2002) note that wealth bequests as well as inter vivos wealth transfers to children may play a major role at the top of the income distribution, but not for most families. Inter-vivos and bequest transfers of significant magnitude, which are typically only available to children of rich families, can open up more opportunities that are not accessible to others (Grawe 2008).[16] For example, children of very rich families may be more likely to become self-employed, because upon receipt of wealth inheritance or transfer they can afford substantial start-up capital required for such ventures, while others may be "locked out of self employment due to lack of fund".[17] Entrepreneurship plays an important role in generating a high concentration of wealth. For example, Quadrini (1999) finds that about 50 percent of households in the very top of the wealth distribution are self-employed. The connection between income, wealth transfer and bequest, and self-employment implies a much greater persistence in income, especially at the upper tail of the income distribution. Indeed, Bowles and Gintis (2002) find that roughly one-third of the IGE can be explained by child wealth. Similarly, Björklund et al. (2012) find that inherited wealth, instead of cognitive and non-cognitive skills, is

[15] Black and Devereux (2011) note that this is not the only approach to classify the families that are credit constrained. Indeed, Grawe (2004) and Mulligan (1997) have considered other approaches to define credit-constrained families.

[16] Few individuals receive inheritances of significant magnitude (see, e.g., Mulligan (1997) for evidence)

[17] Grawe (2008) cites evidence that individuals receiving inheritances are twice as likely to become self-employed, whose business prospects also increase by about 20 percent.

the most likely mechanism to explain the strong persistence at the top of the income distribution.

A recent paper by Cavaglia (2015) provides another explanation for our results. Her explanation is based on the differing role of social network in determining a child's educational and labor market outcomes. Both high- and low-income families have relatively homogeneous social networks, and generally cannot help their children outside of their own networks. Also, considering the job opportunities facing their children, high-income families are more likely to invest in their children, while low-income families not. As a result, the children of these families tend to find the jobs with similar skill levels as their parents, which strengthens the intergenerational persistence. By contrast, middle-income families have relatively diverse networks (perhaps due to their similar costs in investing in highly skilled and unskilled friends), and the role of social networks is also less determinant for medium-skilled positions. Therefore, the children of middle-class families will be less likely to end up in the jobs with similar skill levels as their parents, which weakens the intergenerational persistence. This explanation is consistent with our finding, and Cavaglia (2015) provides some empirical evidence supporting this explanation.

7 Further Implications of the Main Result

The nonlinearity uncovered here seems to be a robust finding and has some important implications for the literature estimating the IGM. First, our results reassure that the results obtained using the transition-matrix in the previous literature are not an artifact of the floor-ceiling effect; the nonlinear pattern – the higher probabilities for the workers whose parents are at the very top and very bottom of the income distribution to stay the same income category – is not due to the fact that people in the tails of the distribution are restricted from further upward or downward movement, but indeed a result of strong persistence in income associations among them. On the other hand, our results do indicate that mobility is a continuous measure, and that some at the extremes of the income distribution can continue to move upward further, which is completely masked by the approaches based on ranks or percentiles. Our nonparametric results thus provide more detailed characterization of the magnitude of the persistence across population.

Second, our nonparametric estimates can accommodate a wide range of IGEs found in the literature. As noted in Durlauf and Shaorshadze (forthcoming), estimates of the IGE in the U.S. for similar cohorts range from 0.2 (Behrman and Taubman, 1985), to .4 (Solon, 1992) and to .6 (Mazumder, 2005). As is now well known, in the presence of nonlinearity, OLS estimator is a weighted average of marginal effects over the distribution of covariates (see Løken et al. (2012) for more detail). The underlying weighting scheme gives more weight to marginal effects close to the sample median of the covariate (parental income in this case). As a result, estimated IGE using OLS can vary with different samples and aggregation methods, even if it is free of life-cycle bias.

This implies that the underlying marginal effects should cover at least these observed average IGEs, and our estimates indeed do.

Finally, in addition to documenting IGEs for a particular cohort, there has also been more attention paid to the changes in IGEs across cohorts. Although there has been some evidence that the IGEs are relatively stable over time (Corak, 2013; Chetty et al., 2014b), it is nevertheless important to apply our approach to investigate this issue. Due to lack of data, we cannot pursue this issue further and leave it for future research. However, we do want to point out a subtle implication that may not be immediately obvious for studies on IGEs over time or across countries. Again, as discussed above, OLS estimator is a weighted average of marginal effects over the distribution of covariates. It is possible that even if the marginal effects are the same, we may find different IGEs, which can lead to completely different conclusions about the cross-cohort or -country analysis.

8 Conclusions

Economists have long been interested in estimation of intergenerational mobility elasticities (IGE) of children's income with respect to parental permanent income, a good indication of equality of opportunity in society. This paper provides a new framework to estimate the IGE. Our approach confronts two important issues surrounding the estimations that have yet to be addressed together in a systematic fashion, namely the heterogeneity/nonlinearity in IGE and the latent nature of the permanent income.

Using the proposed method, we find that there exists substantial heterogeneity in IGEs. More importantly, we observe a U-shape pattern with twin peaks for the IGE. This finding indicates substantial persistence in income across generations for both rich and poor families, but lack of persistence for the majority of the population. This finding reconciles the divergent views that motivate our paper in the introduction. The "twin peaks" is surprisingly close to the perception by the general public of income mobility at both extremes of the income distribution in the U.S.. At the same time, a modest connection between parental and child economic success for most of the families confirms that the U.S. was indeed the "land of opportunity" to live the "American Dream", just not for everyone! We believe that this is an interesting finding. In the paper we attempt to offer an explanation for the main result, but it is nevertheless speculative at best. Future research should focuse on shedding light on the potential mechanisms giving rise to this pattern, which could in turn help us better understand the income evolution over generations in the U.S.. It is also our hope that our approach can be applied to explore within-group heterogeneity in IGEs when datasets with a larger number of individuals for each type of characteristics exist, to examine whether such nonlinearity exists in other countries, and to examine other important topics with similar issues.

References

Becker, G.S., "Family Economics and Macro Behavior," *American Economic Review*, 1988, *78* (1), 1–13.

_ **and N. Tomes**, "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 1986, *4* (3), S1–S39.

Behrman, J.R. and P. Taubman, "Intergeerational Earnings Mobility in the United States: Some Estimates and a Test of Becker's Intergenerational Endowments Model," *Review of Economics and Statistics*, 1985, *67* (1), 144–151.

Bhattacharya, D. and B. Mazumder, "A Nonparametric Analysis of Black-White Differences in Intergerational Income Mobility in the United States," *Quantitative Economics*, 2011, *2*, 335–379.

Björklund, A. and M. Jäntti, *Oxford Handbook of Economic Inequality* number 20, Oxford: Oxford University Press,

_ **, J. Roine, and D. Waldenstrom**, "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?," *Journal of Public Economics*, 2012, *96*, 474–484.

Black, Dan A and Jeffrey A Smith, "Estimating the returns to college quality with multiple proxies for quality," *Journal of Labor Economics*, 2006, *24* (3), 701–728.

Black, S.E. and P.J. Devereux, *Handbook of labor economics*, Vol. 4, Elsevier,

Bloome, Deidre, "Income Inequality and Intergenerational Income Mobility in the United States," *Social Forces*, 2015, *93* (3), 1047–1080.

Böhlmark, Anders and Matthew J Lindquist, "Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden," *Journal of Labor Economics*, 2006, *24* (4), 879–896.

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

Carneiro, Pedro and James J. Heckman, "The Evidence on Credit Constraints in Post-Secondary Schooling," *Economic Journal*, 2002, *112* (482), 705–734.

Cavaglia, Chiara, "Intergenerational mobility and social networks: theory and evidence," *Unpublished Manuscript*, 2015.

Chadwick, Laura and Gary Solon, "Intergenerational Income Mobility Among Daughters," *American Economic Review*, 2002, *92* (1), 335–344.

Chetty, R., N. Hendren, P. Kline, and E. Saez, "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

_ , _ , _ , _ , and N. Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review Papers and Proceedings*, 2014, *104* (5), 141–147.

Corak, M., *The Economics of Inequality, Poverty, and Discrimination in the 21st Century*, Praeger,

_ and A. Heisz, "The Intergenerational EArnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data," *Journal of Human Resources*, 1999, *34* (3), 504–533.

Cunha, F., J.J. Heckman, and S. Schennach, "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 2010, *78* (3), 883–931.

_ , _ , L.J. Lochner, and D.V. Miasterov, *Handbook of the Economics of Education*, Amsterdam: North-Holland: Elsevier,

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

Durlauf, S.N. and I. Shaorshadze, *Emerging Trends in the Social and Behavioral Sciences*, Sherman Oaks: Sage Publishing, forthcoming.

Fagereng, A., M. Mogstad, and M. Rønning, "Why Do Wealthy Parents Have Wealthy Children?," *Statistics Norway, Discussion Papers*, June 2015, *813.*

Grawe, N., "Lifecycle Bias in Estimates of Intergernational Earnings Persistence," *Labour Economics*, 2006, *13* (5), 551–570.

_ , "Wealth and Economic Mobility," *PEW Report on Economic Mobility Project*, 2008.

Grawe, Nathan D., "Reconsdering the Use of Nonlinearities in Intergenerational Earnings Mobility as a Test for Credit Constraints," *Journal of Human Resources*, 2004, *39*, 813–827.

Haider, Steven J and Gary Solon, "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, 2006, *96* (4), 1308–1320.

Hertz, T., *Unequal Chances: Family Background and Economic Success*, Princeton, N.J.: Princeton University Press,

Hu, Y., "Identification and Estimation of Nonlinear Models with Misclassification Error Using Instrumental Variables: A General Solution," *Journal of Econometrics*, 2008, *144*, 27–61.

_ **and G. Ridder**, "On Deconvolution as a First Stage Nonparametric Estimator," *Econometric Reviews*, 2008, *29* (4), 1–32.

_ **and S. Schennach**, "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 2008, *76*, 195–216.

Hu, Yingyao and Yuya Sasaki, "Closed-Form Estimation of Nonparametric Models with Non-Classical Measurement Errors," *Journal of Econometrics*, 2014, *185* (2), 392–408.

Lee, C. and G. Solon, "Trends in Intergenerational Income Mobility," *Review of Economics and Statistics*, 2009, *91* (4), 766–772.

Lefgren, L., M.J. Lindquist, and D. Sims, "Rich Dad, Smart Dad: Decomposing the Intergenerational Transmission of Income," *Journal of Political Economy*, 2012, *120* (2), 268–303.

Lemieux, Thomas, "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 2006, *96* (3), 461–98.

Li, Tong, "Robust and consistent estimation of nonlinear errors-in-variables models," *Journal of Econometrics*, 2002, *110* (1), 1–26.

Løken, Katrine V, Magne Mogstad, and Matthew Wiswall, "What linear estimators miss: The effects of family income on child outcomes," *American Economic Journal: Applied Economics*, 2012, *4* (2), 1–35.

Mayer, Susan E. and Leonard M. Lopoo, *Generational Income Mobility in North America and Europe*, Cambridge: Cambridge University Press,

_ **and** _ , "Has the Intergenerational Transmission of Economic Status Changed?," *Journal of Human Resources*, 2005, *40* (1), 169–285.

Mazumder, Bhashkar, "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data," *Review of Economics and Statistics*, 2005, *87* (2), 235–255.

Mulligan, Casey B., *Parental Priorities and Economic Inequality*, Chicago, IL: University of Chicago Press, 1997.

Nilsen, Øivind A., Kjell Vaage, Arild Aakvik, and Karl Jacobsen, "Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings," *Scandinavian Journal of Economics*, 2012, *114* (1), 1–23.

O'Neill, D., O. Sweetman, and D. Van de gaer, "The Effects of Measurement Error and Omitted Variables when Using Transition Matrices to Measure Intergenerational Mobility," *Journal of Economic Inequality*, 2007, *5*, 159–178.

Quadrini, Vincenzo, "The Importance of Entrepreneurship for Wealth Concentration and Mobility," *Review of Incoe and Wealth*, 1999, *45* (1), 1–19.

Rilstone, P and A Ullah, "Nonparametric estimation of response coefficients," *Communications in Statistics-Theory and Methods*, 1989, *18* (7), 2615–2627.

Schennach, Susanne M, "Estimation of nonlinear models with measurement error," *Econometrica*, 2004, pp. 33–75.

_ , "Nonparametric regression in the presence of measurement error," *Econometric Theory*, 2004, *20* (06), 1046–1093.

_ **and Yingyao Hu**, "Nonparametric identification and semiparametric estimation of classical measurement error models without side information," *Journal of the American Statistical Association*, 2013, *108* (501), 177–186.

Solon, G., "Intergenerational Income Mobility in the United States," *American Economic Review*, 1992, *82* (3), 393–408.

_ , *Handbook of Labor Economics*, Vol. III, Amsterdam: North-Holland,

Zimmerman, D.J., "Regression Towards Mediocrity in Economic Stature," *American Economic Review*, 1992, *82*, 409–429.

Appendix

A Proofs

Proof of Lemma 2. We only prove the identification of f_{X^*}. The proof for the density f_{Y^*} follows a symmetric argument.

First consider the two measures of X^*

$$X_1 = X^* + V_1, \;\; X_2 = \alpha_2 X^* + V_2,$$

with α_2 being a known. First define the joint ch.f. of X_1 and X_2 as:

$$\phi_{X_1,X_2}(s_1, s_2) \equiv E[e^{is_1X_1+is_2X_2}].$$

Take the derivate of the ch.f. above with respect to s_1 and evaluate it at $s_1 = 0$:

$$\begin{aligned} \left.\frac{\phi_{X_1,X_2}(s_1, s_2)}{\partial s_1}\right|_{s_1=0} &= E[i(X^* + V_1)e^{is_2X_2}] \\ &= E[iX^*e^{is_2X_2}] + E[iV_1e^{is_2X_2}] \\ &= E[iX^*e^{is_2\alpha_2X^*}]E[e^{is_2V_2}], \end{aligned}$$

where $E[iV_1e^{is_2X_2}] = 0$ is due to the assumption $E[V_1|X_2] = 0$, and the third equality holds because of the independence of V_2 from X^*. Similarly, we have $\phi_{X_2}(s_2) = E[e^{is_2\alpha_2X^*}]E[e^{is_2V_2}]$. Consider that

$$\frac{\left.\frac{\phi_{X_1,X_2}(s_1,s_2)}{\partial s_1}\right|_{s_1=0}}{\phi_{X_2}(s_2)} = \frac{E[iX^*e^{is_2\alpha_2X^*}]}{E[e^{is_2\alpha_2X^*}]} = \left.\frac{\partial \ln E[e^{irX^*}]}{\partial r}\right|_{r=\alpha_2s_2}.$$

Thus, the ch.f. of f_{X^*}, ϕ_{X^*} can be recovered as

$$\phi_{X^*}(s) = \exp\left(\int_0^s \frac{\left.\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_2/\alpha_2)\right|_{s_1=0}}{\phi_{X_1,X_2}(s_2/\alpha_2)|_{s_1=0}} ds_2\right).$$

The density f_{X^*} can be recovered using inverse Fourier transform

$$\begin{aligned} f_{X^*}(x^*) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*}\phi_{X^*}(s)ds \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*}\exp\left(\int_0^s \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1,s_2/\alpha_2)\Big|_{s_1=0}}{\phi_{X_1,X_2}(s_2/\alpha_2)\big|_{s_1=0}}ds_2\right)ds. \end{aligned} \tag{A.1}$$

■

Proof of Proposition 1. The proof follows Hu and Sasaki (2014) and we present here for completeness. Once f_{X^*} is identified, the identification of the mobility function $g(\cdot)$ can be obtained by using the joint distribution of only one measurement for both generations. Thus, one can identify the mobility function using any combination of children and parental observed incomes, i.e., $\{X_1, Y_1\}$, $\{X_1, Y_2\}$, $\{X_2, Y_1\}$, or $\{X_2, Y_2\}$. We illustrate the identification using $\{X_2, Y_2\}$ where the life-cycle coefficients are not equal to 1.

We first define the joint ch.f. of X_2 and Y_2 in the following:

$$\phi_{X_2,Y_2}(s, v) \equiv E[e^{isX_2+ivY_2}].$$

Note that $Y_2 = \delta_2 Y^* + U_2 = \delta_2(g(X^*) + \epsilon) + U_2 = \delta_2 g(X^*) + \delta_2\epsilon + U_2$.

Taking derivative of the ch.f. above with respect to v and evaluate the objective at $v = 0$.

$$\begin{aligned} \frac{\partial}{\partial v}\phi_{X_2,Y_2}(s,0) &= iE\left[(\delta_2 g(X^*) + \delta_2\varepsilon + U_2)e^{is(\alpha_2X^*+V_2)}\right] \\ &= iE\left[(\delta_2 g(X^*)e^{is(\alpha_2X^*+V_2)}\right] + iE[(\delta_2\varepsilon + U_2)e^{isX_2}] \\ &= iE\left[(g(X^*)e^{i\alpha_2 sX^*}\right]\delta_2 E[e^{isV_2}] \\ &= iE\left[(g(X^*)e^{i\alpha_2 sX^*}\right]\delta_2\frac{\phi_{X_2}(s)}{\phi_{X^*}(\alpha_2 s)}, \end{aligned}$$

where $\phi_{X_2}(s) \equiv E[e^{isX_2}]$ and $\phi_{X^*}(\alpha_2 s) \equiv E[e^{i\alpha_2 sX^*}]$ are ch.fs of X_2 and X^* evaluated at s and $\alpha_2 s$, respectively. The third equation holds because of Assumption 4, ($E(\epsilon|X_2) = 0$ and $E(U_2|X_2) = 0$). As a result, the equation above enables us to obtain:

$$E\left[g(X^*)e^{i\alpha_2 sX^*}\right] = \phi_{X^*}(\alpha_2 s)\frac{\frac{\partial}{\partial v}\phi_{X_2,Y_2}(s,0)}{i\delta_2\phi_{X_2}(s)}.$$

It can be equivalently expressed as,

$$E\left[g(X^*)e^{isX^*}\right] = \int e^{isX^*}g(X^*)f_{X^*}(x^*)dx^* = \phi_{X^*}(s)\frac{\frac{\partial}{\partial v}\phi_{X_2,Y_2}(s/\alpha_2,v)\big|_{v=0}}{i\delta_2\phi_{X_2}(s/\alpha_2)}.$$

Using the inverse Fourier transform,

$$g(X^*)f_{X^*} = \frac{1}{2\pi\delta_2}\int_{-\infty}^{+\infty} e^{-isX^*}\phi_{X^*}(s)\frac{\frac{\partial}{\partial v}\phi_{X_2,Y_2}(s/\alpha_2, v)\big|_{v=0}}{i\phi_{X_2}(s/\alpha_2)}ds.$$

Consequently, the mobility function $g(X^*)$ can be identified as

$$g(X^*) = \frac{1}{2\pi\delta_2 f_{X^*}}\int_{-\infty}^{+\infty} e^{-isX^*}\phi_{X^*}(s)\frac{\frac{\partial}{\partial v}\phi_{X_2,Y_2}(s/\alpha_2, v)\big|_{v=0}}{i\phi_{X_2}(s/\alpha_2)}ds, \tag{A.2}$$

where ϕ_{X^*} and f_{X^*} are identified in Lemma 2. ■

Proof of Theorem 1. The effect of inappropriately assuming δ_1 and α_1 to be one can be analyzed by comparing the mobility function $g(\cdot)$ based on the inappropriate normalization with that of δ_1 and α_1 are known to the researcher.

When both α_1 and α_2 are known, a direct application of Lemma 2 implies that the density of X^* can be identified as

$$f_{X^*}(x^*) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*}\exp\left(\int_0^s \frac{1}{\alpha_1}\frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_2/\alpha_2)\Big|_{s_1=0}}{\phi_{X_2}(s_2/\alpha_2)}ds_2\right)ds.$$

Note that this closed-form identification degenerates to (A.1) when α_1.

Let $\widetilde{\delta}_2$ and $\widetilde{\alpha}_2$ denote the identified coefficients in model (1) if the normalization is mis-specified, i.e., $\alpha_1 \neq 1$ and $\delta_1 \neq 1$. Then

$$\widetilde{\delta}_2 = \frac{\delta_2}{\delta_1} \neq \delta_2, \;\; \widetilde{\alpha}_2 = \frac{\alpha_2}{\alpha_1} \neq \alpha_2.$$

When α_1 is mistakenly assumed to be one, we have $\widetilde{\alpha}_2 = \frac{\alpha_2}{\alpha_1}$. Following Lemma 2, the density function f_{X^*} is obtained by plugging α_2 as $\widetilde{\alpha}_2$ and α_1 as 1. Thus, we may obtain an incorrect density function as follows:

$$
\begin{aligned}
\widetilde{f}_{X^*}(x^*) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^s \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_2/\widetilde{\alpha}_2)\Big|_{s_1=0}}{\phi_{X_2}(s_2/\widetilde{\alpha}_2)} ds_2\right) ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^s \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, \alpha_1 s_2/\alpha_2)\Big|_{s_1=0}}{\phi_{X_2}(\alpha_1 s_2/\alpha_2)} ds_2\right) ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^{\alpha_1 s} \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_3/\alpha_2)\Big|_{s_1=0}}{\alpha_1\phi_{X_2}(s_3/\alpha_2)} ds_3\right) ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\alpha_1 s\frac{x^*}{\alpha_1}} \exp\left(\int_0^{\alpha_1 s} \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_3/\alpha_2)\Big|_{s_1=0}}{\alpha_1\phi_{X_2}(s_3/\alpha_2)} ds_3\right) ds \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{\alpha_1} e^{-iv\frac{x^*}{\alpha_1}} \exp\left(\int_0^{v} \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_3/\alpha_2)\Big|_{s_1=0}}{\alpha_1\phi_{X_2}(s_3/\alpha_2)} ds_3\right) dv \\
&= \frac{1}{\alpha_1} f_{X^*}\left(\frac{x^*}{\alpha_1}\right). \qquad \text{(A.3)}
\end{aligned}
$$

A similar relationship holds for the density of Y^*, i.e., $\widetilde{f}_{Y^*}(y^*) = \frac{1}{\delta_1} f_{Y^*}\left(\frac{y^*}{\delta_1}\right)$.

To analyze the effects of inappropriate normalization on the mobility function, we compare the incorrect mobility function (denoted by $\widetilde{g}(\cdot)$) with the one that identified when all the coefficients are known. With slight abuse of notation, let $g(x^*; \delta_1, \alpha_1)$ be the mobility function identified correctly under Assumption 6. Following Proposition 1, the true mobility function is identified as follows.[18]

$$
\begin{aligned}
g(x^*; \alpha_1, \delta_1) &= \frac{1}{2\pi\delta_2 f_{X^*}(x^*)}\int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^s \frac{1}{\alpha_1} \frac{\frac{\partial}{\partial s_1}\phi_{X_1,X_2}(s_1, s_2/\alpha_2)\Big|_{s_1=0}}{\phi_{X_1,X_2}(s_2/\alpha_2)|_{s_1=0}} ds_2\right) \\
&\times \frac{\frac{\partial}{\partial v}\phi_{X_2,Y_2}(s/\alpha_2, v)\Big|_{v=0}}{i\phi_{X_2}(s/\alpha_2)} ds. \qquad \text{(A.4)}
\end{aligned}
$$

In contrast, when we mistakenly pick the normalization period, again the life-cycle coefficient for

[18] Note that when $\alpha_1 = \delta_1 = 1$, $g(x^*; \alpha_1, \delta_1)$ becomes the expression we obtain in Proposition 1, when δ_1 does not affect the identification because only $\{X_1, X_2, Y_2\}$ are used for identifying the mobility function.

period 2 is mis-identified, i.e., $\widetilde{\alpha}_2 = \frac{\alpha_2}{\alpha_1}$ and $\widetilde{\delta}_2 = \frac{\delta_2}{\delta_1}$. Consequently, the mobility function becomes

$$\begin{aligned}
& \widetilde{g}(x^*; \alpha_1, \delta_1) \\
= & \frac{1}{2\pi \widetilde{\delta}_2 \widetilde{f}_{X^*}(x^*)} \int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^s \frac{\frac{\partial}{\partial s_1} \phi_{X_1,X_2}(s_1, s_2/\widetilde{\alpha}_2)\Big|_{s_1=0}}{\phi_{X_2}(s_2/\widetilde{\alpha}_2)} ds_2 \right) \frac{\frac{\partial}{\partial w} \phi_{X_2,Y_2}(s/\widetilde{\alpha}_2, w)\big|_{w=0}}{i\phi_{X_2}(s/\widetilde{\alpha}_2)} ds \\
= & \frac{\delta_1 \alpha_1}{2\pi \delta_2 f_{X^*}(x^*/\alpha_1)} \int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^s \frac{\frac{\partial}{\partial s_1} \phi_{X_1,X_2}(s_1, \alpha_1 s_2/\alpha_2)\Big|_{s_1=0}}{\phi_{X_2}(\alpha_1 s_2/\alpha_2)} ds_2 \right) \frac{\frac{\partial}{\partial w} \phi_{X_2,Y_2}(\alpha_1 s/\alpha_2, w)\big|_{w=0}}{i\phi_{X_2}(\alpha_1 s/\alpha_2)} ds \\
= & \frac{\delta_1 \alpha_1}{2\pi \delta_2 f_{X^*}(x^*/\alpha_1)} \int_{-\infty}^{\infty} e^{-isx^*} \exp\left(\int_0^{\alpha_1 s} \frac{\frac{\partial}{\partial s_1} \phi_{X_1,X_2}(s_1, v/\alpha_2)\Big|_{s_1=0}}{\alpha_1 \phi_{X_2}(v/\alpha_2)} dv \right) \frac{\frac{\partial}{\partial w} \phi_{X_2,Y_2}(\alpha_1 s/\alpha_2, w)\big|_{w=0}}{i\phi_{X_2}(\alpha_1 s/\alpha_2)} ds \\
= & \frac{\delta_1 \alpha_1}{2\pi \delta_2 f_{X^*}(x^*/\alpha_1)} \int_{-\infty}^{\infty} \frac{1}{\alpha_1} e^{-iux^*/\alpha_1} \exp\left(\int_0^{u} \frac{\frac{\partial}{\partial s_1} \phi_{X_1,X_2}(s_1, v/\alpha_2)\Big|_{s_1=0}}{\alpha_1 \phi_{X_2}(v/\alpha_2)} dv \right) \frac{\frac{\partial}{\partial w} \phi_{X_2,Y_2}(u/\alpha_2, w)\big|_{w=0}}{i\phi_{X_2}(u/\alpha_2)} du \\
= & \frac{\delta_1}{2\pi \delta_2 f_{X^*}(x^*/\alpha_1)} \int_{-\infty}^{\infty} e^{-iux^*/\alpha_1} \exp\left(\int_0^{u} \frac{\frac{\partial}{\partial s_1} \phi_{X_1,X_2}(s_1, v/\alpha_2)\Big|_{s_1=0}}{\alpha_1 \phi_{X_2}(v/\alpha_2)} dv \right) \frac{\frac{\partial}{\partial w} \phi_{X_2,Y_2}(u/\alpha_2, w)\big|_{w=0}}{i\phi_{X_2}(u/\alpha_2)} du \\
= & \delta_1 g(x^*/\alpha_1; \alpha_1, \delta_1).
\end{aligned} \tag{A.5}$$

■

Proof of Corollary 1. From Theorem 1, $\widetilde{g}(x^*) = \delta_1 g(x^*/\alpha_1)$, so the IGM satisfies the following relationship

$$\widetilde{g}'(x^*) = \frac{\delta_1}{\alpha_1} g'(x^*/\alpha_1).$$

If the mobility function is linear, i.e., $g(x^*) = \beta_0 + \beta_1 x^*$ as widely used in the literature, then

$$\widetilde{g}'(x^*) = \frac{\delta_1 \beta_1}{\alpha_1},$$

which is also a constant for any x^*. If $\delta_1 = \alpha_1$, then the IGM becomes

$$\widetilde{g}'(x^*) = \beta_1. \tag{A.6}$$

As a result, if we have a linear mobility function, mis-specifying the normalization does not affect the estimation of the IGM given the life-cycle coefficient is the same across generations. ■

Proof of Theorem 2. From Theorem 1, we know that $g(x^*) = \frac{1}{\delta_1} \widetilde{g}(\alpha_1 x^*)$, naturally, we can bound the mobility function through

$$\min_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1],\delta_1\in[\underline{\delta}_1,\bar{\delta}_1]}\left(\frac{1}{\delta_1}\widetilde{g}(\alpha_1 x^*)\right) \leq g(x^*) \leq \max_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1],\delta_1\in[\underline{\delta}_1,\bar{\delta}_1]}\left(\frac{1}{\delta_1}\widetilde{g}(\alpha_1 x^*)\right)$$

$$\Rightarrow \quad \frac{1}{\bar{\delta}_1}\min_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1]}\left(\widetilde{g}(\alpha_1 x^*)\right) \leq g(x^*) \leq \frac{1}{\underline{\delta}_1}\max_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1]}\left(\widetilde{g}(\alpha_1 x^*)\right) \tag{A.7}$$

Similarly, we can bound the IGM elasticities as follows:

$$\frac{\underline{\alpha}_1}{\bar{\delta}_1}\min_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1]}\left(\widetilde{g}'(\alpha_1 x^*)\right) \leq g'(x^*) \leq \frac{\bar{\alpha}_1}{\underline{\delta}_1}\max_{\alpha_1\in[\underline{\alpha}_1,\bar{\alpha}_1]}\left(\widetilde{g}'(\alpha_1 x^*)\right), \tag{A.8}$$

■

B Estimation

Since the identification results in the previous section are constructive, we follow the identification argument in estimation. Given that an observed sample of $\{X_{jt}, Y_{jt}\}_{j=1}^N, t = 1,2$, the mobility function $g(x^*)$ is estimated in three steps. First, the coefficients α_2 and δ_2 can be directly estimated using the sample analog of equation (3).

$$\widehat{\delta}_2 = \frac{\frac{1}{N}\sum_j X_{1j}Y_{2j} - \frac{1}{N}\sum_j X_{1j}\frac{1}{N}\sum_j Y_{2j}}{\frac{1}{N}\sum_j X_{1j}Y_{1j} - \frac{1}{N}\sum_j X_{1j}\frac{1}{N}\sum_j Y_{1j}}. \tag{B.1}$$

α_2 can be estimated analogously. Based on the estimate of α_2 and δ_2, together with the closed-form provided in Lemma 2, we estimate the densities of permanent income f_{X^*} and f_{Y^*} in the second step. Such an approach has been widely used, e.g, Li (2002).

$$\widehat{f}_{X^*}(x^*) = \frac{1}{2\pi}\int_{-S_N}^{S_N} e^{-isx^*}\exp\left(i\int_0^s \frac{\sum_{j=1}^N X_{1j}\exp(iX_{2j}v/\widehat{\alpha}_t)}{\sum_{j=1}^N \exp(iX_{2j}v/\widehat{\alpha}_2)}dv\right)\phi_K(s/S_N)ds,$$

where $i \equiv \sqrt{-1}$, $\exp\left(i\int_0^s \frac{\sum_{j=1}^N X_{1j}\exp(iX_{2j}v/\widehat{\alpha}_t)}{\sum_{j=1}^N \exp(iX_{2j}v/\widehat{\alpha}_t)}dv\right)$ is an estimate of $\phi_{X^*}(s)$, the characteristic function of f_{X^*}, and $\phi_K(\cdot)$ is the Fourier transform of a kernel function $K(\cdot)$ with a bandwidth $1/S_N$. The smoothing parameter S_N depends on the sample size N. To assure that the estimate of $\phi_{X^*}(s)$ uniformly converges to its true function over $[-S_N, S_N]$ at a geometric rate with respect to the sample size N, Hu and Ridder (2008) suggest a form of S_N

$$S_N = O\left(\frac{N}{\log N}\right)^\gamma \text{ for } \gamma \in \left(0, \frac{1}{2}\right).$$

Table 1: Estimate of α and δ

	Parameter=(True Value)†			
	$\alpha_2 = 1.105$		$\delta_2 = 1.184$	
	Mean	Std. Dev.	Mean	Std. Dev.
Mobility function: $g(X^*) = 0.6X^*$				
$N = 100$	1.112	0.123	1.189	0.155
$N = 300$	1.110	0.068	1.189	0.083
$N = 500$	1.108	0.054	1.186	0.063
Mobility function: $g(X^*) = -0.2X^{*2} + X^*$				
$N = 100$	1.110	0.108	1.184	0.094
$N = 300$	1.110	0.059	1.186	0.049
$N = 500$	1.108	0.047	1.185	0.038

† The coefficients at period $t = 1$ are $\alpha_1 = \delta_1 = 1$ by construction.

In the last step, we estimate the mobility function using the closed-form in Proposition 1:

$$\widehat{g}(x^*) = \frac{1}{2\pi\widehat{\delta}_2\widehat{f}_{X^*}} \times \\ \int_{-S_N}^{S_N} e^{-isx^*} \exp\left(i \int_0^s \frac{\sum_{j=1}^N X_{1j} \exp(iX_{2j}v/\widehat{\alpha}_2)}{\sum_{j=1}^N \exp(iX_{2j}v/\widehat{\alpha}_2)} dv\right) \frac{\sum_{j=1}^N Y_{2j} \exp(isX_{2j}/\widehat{\alpha}_2)}{\sum_{j=1}^N \exp(isX_{2j}/\widehat{\alpha}_2)} \phi_K(s/S_N) ds. \tag{B.2}$$

If we are further interested in the derivative of the mobility function, several methods are available for the estimation. In our paper, we employ the simple kernel method proposed in Rilstone and Ullah (1989).

$$\widehat{\frac{dg(x^*)}{dx^*}} = \frac{\widehat{g}(x^* + h_N) - \widehat{g}(x^* - h_N)}{2h_N}, \tag{B.3}$$

where h_N is a bandwidth and $h_N \to 0$ as $N \to \infty$.

The results in Theorem 2 allows us to estimate the bounds of the mobility function $g(\cdot)$ when δ_1 and α_1 are known to be in certain closed interval while their exact values are unknown. The procedure is similar to the steps above. ■

Figure 1: Estimate of income densities, $N = 300$

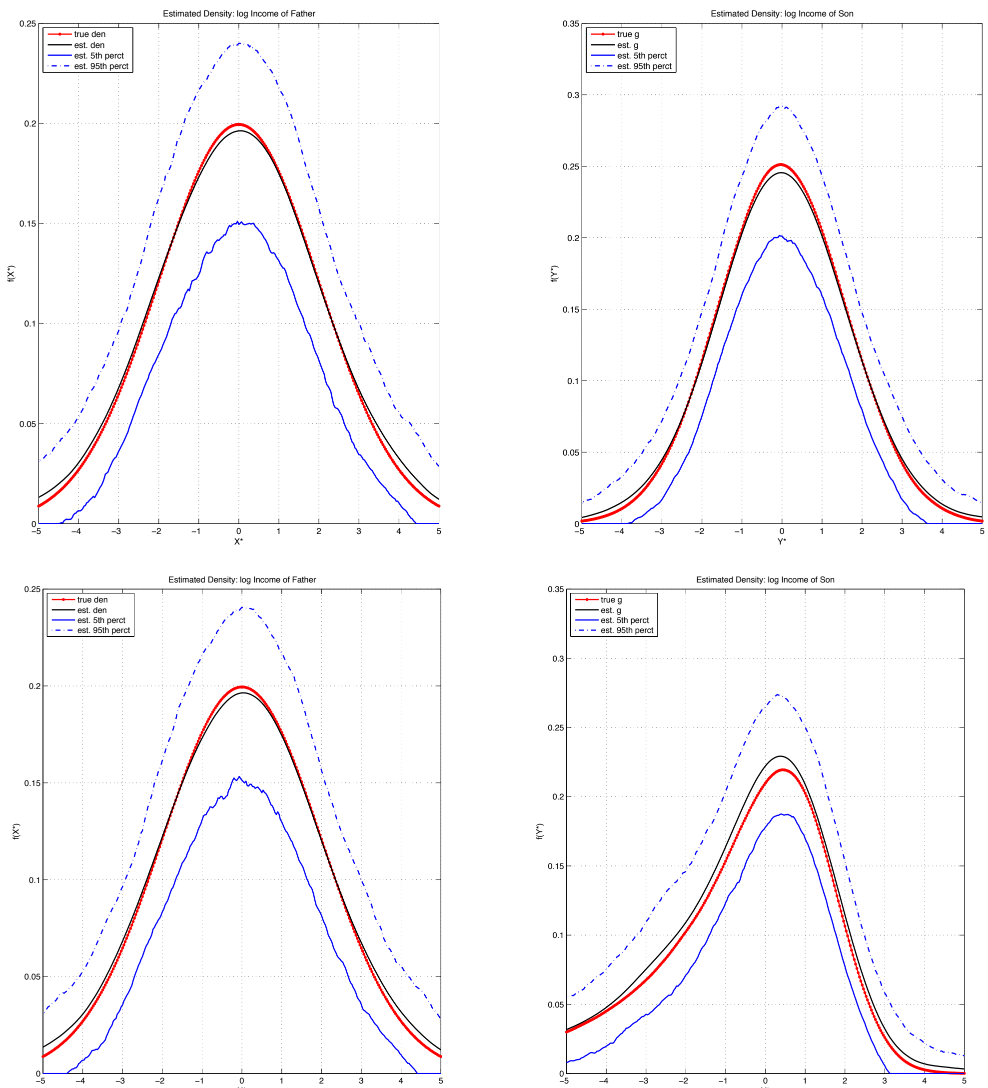


Figure 2: Estimate of mobility function, derivatives and bounds: $N = 300$

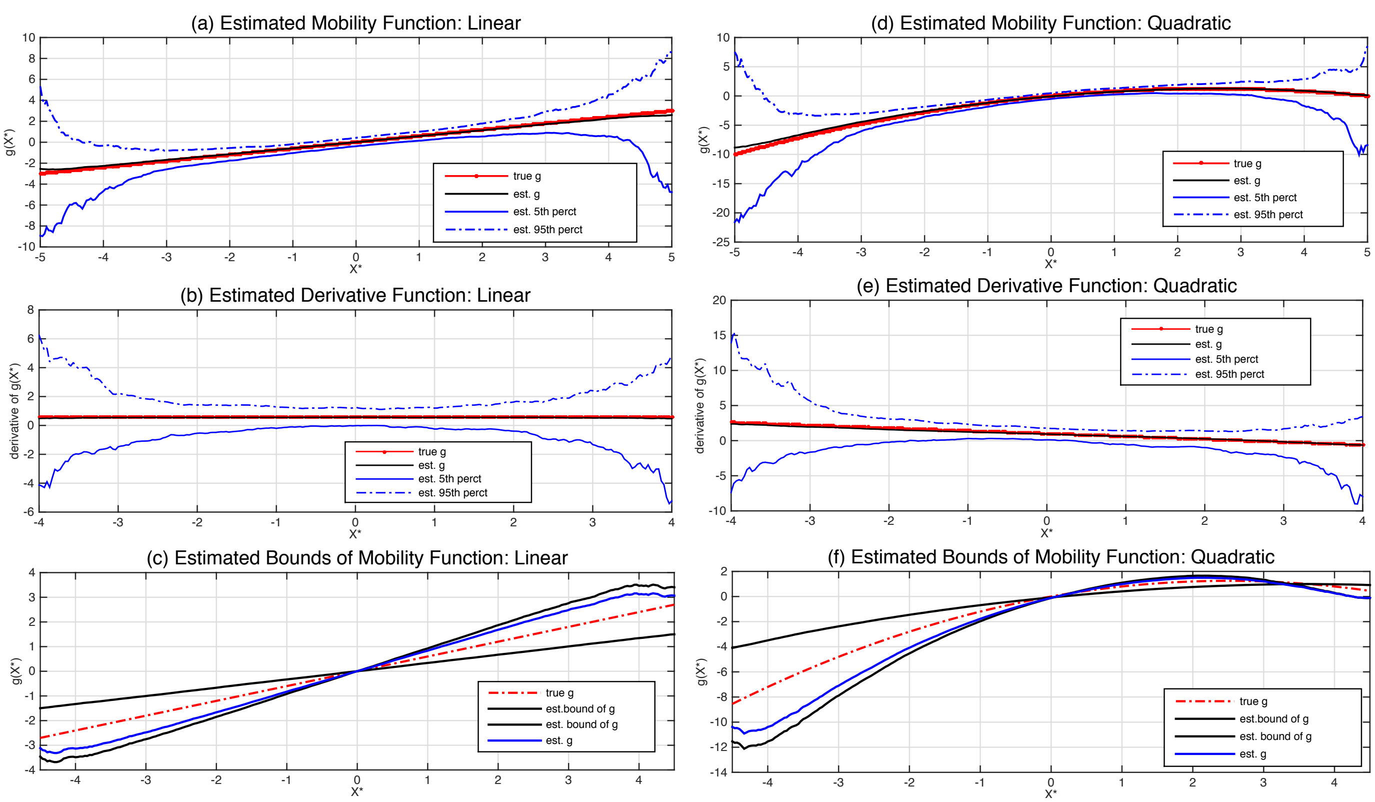


Table 2: Summary statistics

Var.		Son					Father				
		Obs.	Mean	Std.	Min	Max	Obs.	Mean	Std.	Min	Max
Gender		1839	1.508	.5	1	2	1839	1	0	1	1
Region		1685	1.631	.68	1	3	1685	1.631	.68	1	3
Birth		1839	1973.016	9.517	1945	1985	1839	1945.31	10.156	1917	1966
Race		1839	1.092	.289	1	2	1839	1.092	.289	1	2
Age	26	1050	8.592	1.033	3.045	12.373	987	10.477	.863	5.724	11.798
	27	995	8.587	1.02	1.609	11.732	1015	10.463	.975	4.394	11.936
	28	953	8.628	1.025	3.871	12.156	1040	10.376	1.159	1.13	12.234
	29	878	8.636	.961	4.625	11.892	1087	10.305	1.253	.831	12.298
	30	839	8.654	.965	2.485	11.686	1123	10.224	1.299	.704	12.561
	31	775	8.729	.874	4.663	12.113	1171	10.156	1.27	.606	12.315
	32	732	8.733	.83	5.565	11.723	1229	10.093	1.314	.938	12.643
	33	683	8.75	.787	5.602	11.69	1219	9.984	1.309	5.762	12.587
	34	632	8.718	.776	5.209	11.438	1216	9.898	1.336	.472	12.524
	35	571	8.755	.817	3.091	11.613	1249	9.893	1.341	5.746	12.489
	36	542	8.773	.795	2.708	11.744	1303	9.778	1.35	.472	13.07
	37	485	8.791	.744	6.48	11.566	1365	9.736	1.34	.704	13.883
	38	454	8.785	.758	4.248	11.624	1394	9.629	1.302	4.554	12.882
	39	401	8.818	.76	4.585	11.491	1406	9.65	1.302	.606	12.456
	40	385	8.812	.763	4.477	11.388	1438	9.643	1.245	6.14	12.603
	41	344	8.805	.792	4.934	11.194	1425	9.629	1.309	.472	12.665
	42	317	8.907	.794	3.434	11.443	1436	9.608	1.347	.472	13.365
	43	283	8.849	.831	4.304	13.092	1407	9.618	1.294	.472	13.11
	44	266	8.953	.684	6.236	11.278	1395	9.587	1.214	5.333	12.874
	45	226	8.903	.828	6.275	11.603	1347	9.534	1.343	3.584	12.939
	46	221	8.937	.724	6.127	11.122	1376	9.594	1.272	4.477	12.97
	47	174	8.869	.879	5.598	12.08	1317	9.524	1.245	4.382	12.927
	48	174	8.905	.929	2.639	11.609	1251	9.507	1.215	4.883	12.701
	49	141	8.938	.862	5.313	11.987	1245	9.525	1.234	3.989	12.796
	50	139	8.988	.794	6.492	11.364	1207	9.442	1.197	5.182	12.73

Gender: male (1) and female (2); region: northeast+south(1), north central (2) and west (3); race: white (2) and nonwhite (2).

Table 3: Sample Size by Ages of Parent and Child

Parent Child	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
26	420	432	461	502	528	560	604	601	606	669	729	686	661	624	626	609	617	597	575	536	568	532	495	486	516
27	408	430	442	474	497	515	557	578	589	642	703	648	618	604	605	593	601	585	559	528	551	506	480	468	505
28	370	377	396	433	457	485	536	542	549	616	661	628	604	581	582	566	574	557	536	487	522	487	451	450	476
29	346	363	380	408	426	440	482	507	514	569	622	576	548	544	550	534	541	530	516	481	501	453	438	423	468
30	292	304	320	356	382	400	447	457	462	525	564	545	519	503	496	497	495	483	459	425	446	429	383	400	406
31	287	308	324	350	363	369	403	426	435	488	532	495	468	470	465	465	465	468	448	423	433	401	382	374	402
32	215	228	248	285	306	318	358	376	382	433	467	450	422	415	412	414	413	406	387	358	377	360	330	337	348
33	232	252	264	287	296	297	335	358	364	410	446	417	393	394	388	392	391	394	379	362	367	341	312	321	336
34	173	179	194	226	246	257	289	309	315	353	379	364	339	332	333	338	335	336	315	297	306	301	274	284	289
35	173	187	196	218	227	230	264	279	284	322	348	329	307	301	299	306	305	313	301	288	290	274	249	263	267
36	137	145	157	182	198	209	241	259	263	293	312	299	275	267	268	276	281	280	263	246	261	251	236	239	247
37	132	145	156	171	182	187	213	228	233	259	278	262	243	236	234	238	244	252	244	237	236	226	207	218	223
38	102	107	120	144	160	167	194	207	213	241	254	244	223	216	218	222	230	229	215	201	215	209	195	202	201
39	94	104	112	127	140	147	164	178	179	195	214	204	190	179	174	177	187	197	187	183	185	181	169	172	181
40	65	72	83	103	117	127	151	161	174	198	202	197	180	175	177	176	179	179	170	160	171	160	155	162	158
41	69	79	86	101	114	119	136	149	155	167	179	171	161	149	145	150	154	161	154	149	157	151	140	146	154
42	45	49	58	76	89	96	118	130	142	161	164	159	148	145	145	145	149	148	136	130	141	131	124	133	128
43	44	54	57	70	83	88	106	118	125	134	141	137	132	121	114	117	119	128	122	113	120	115	108	117	122
44	29	33	42	56	69	75	96	108	119	136	137	132	125	125	128	125	123	121	113	108	115	106	101	111	106
45	18	27	30	44	54	58	73	84	90	99	106	106	104	96	91	95	93	94	88	82	91	85	78	87	89
46	15	18	24	34	45	50	67	78	89	105	106	105	102	100	101	98	95	94	89	83	87	80	75	82	76
47	2	9	11	22	31	33	45	55	59	67	72	72	71	71	68	68	65	63	58	52	60	56	49	56	57
48	2	5	9	16	22	24	39	49	59	73	74	74	74	74	73	72	70	65	63	60	63	53	46	52	47
49	0	2	3	10	16	19	28	34	37	42	46	46	45	45	45	45	43	42	38	33	36	33	29	35	36
50	0	2	2	6	10	12	24	29	38	50	50	50	50	50	50	50	48	46	46	42	42	34	31	34	27

Figure 3: Estimated Pattern of Income

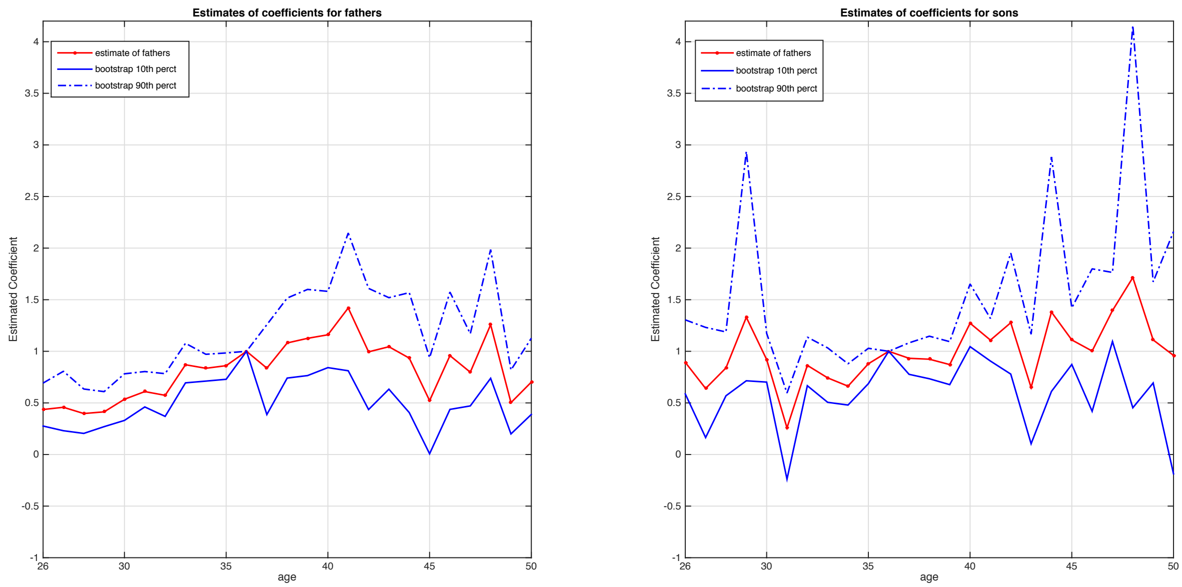
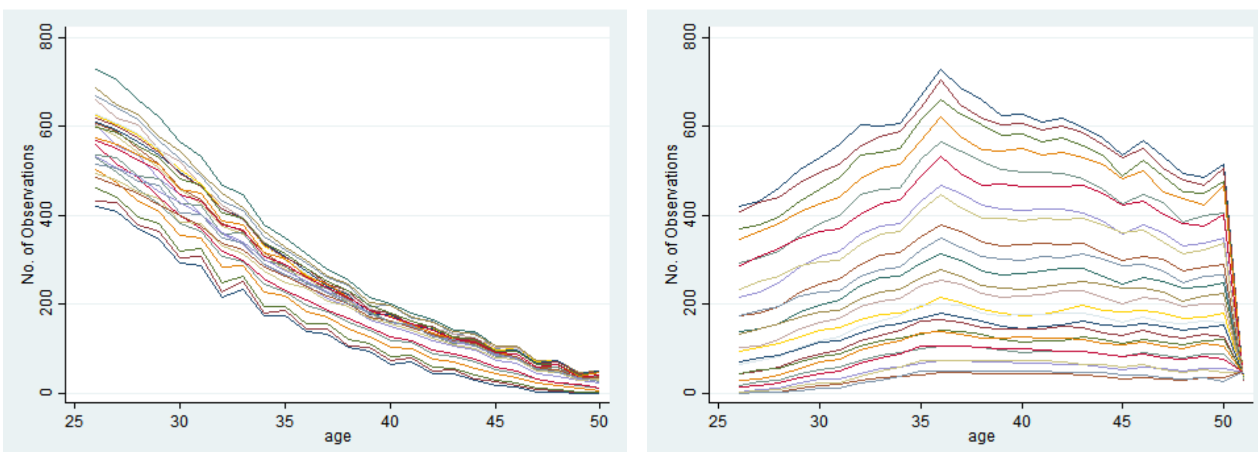


Figure 4: Sample size by ages of parent and child



(a) Child (conditional on Parent's age) (b) Parent (conditional on Child's age)

Figure 5: Estimated derivatives of mobility function

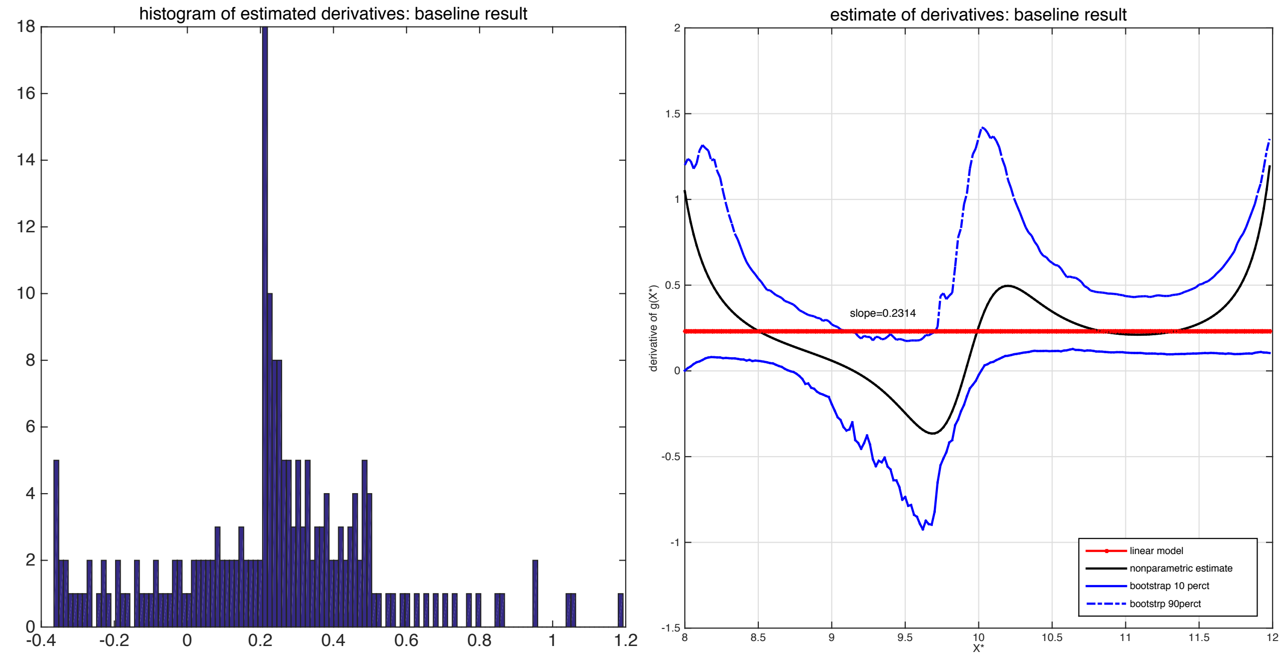


Figure 6: Estimated densities of permanent income

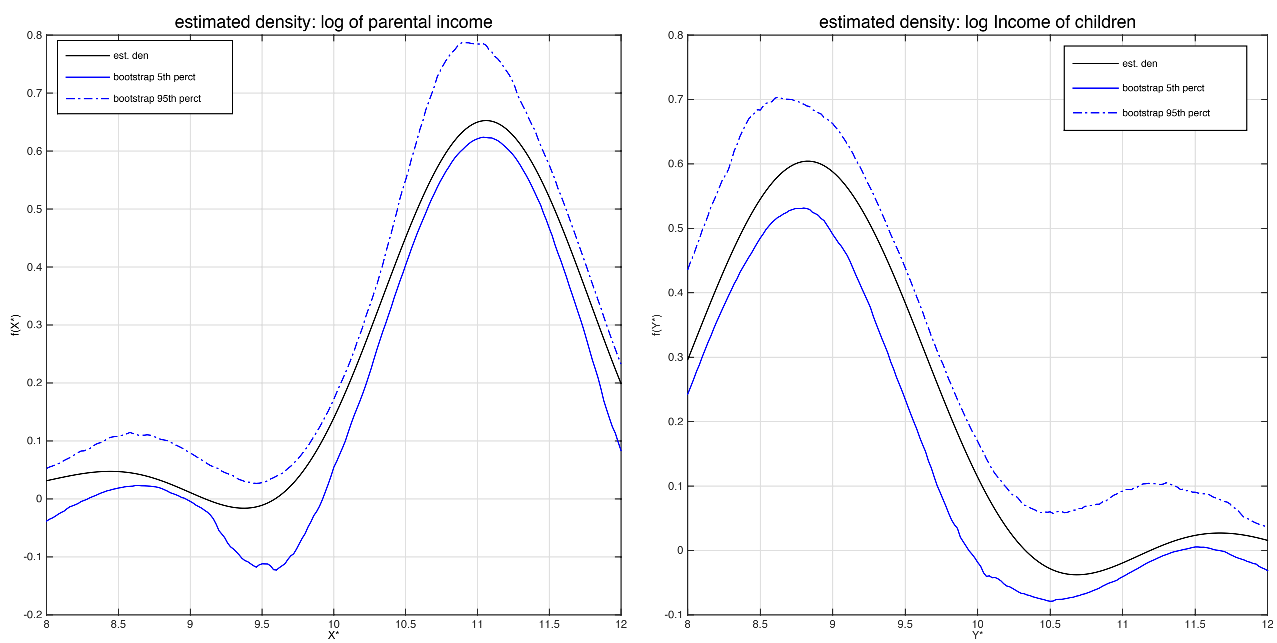
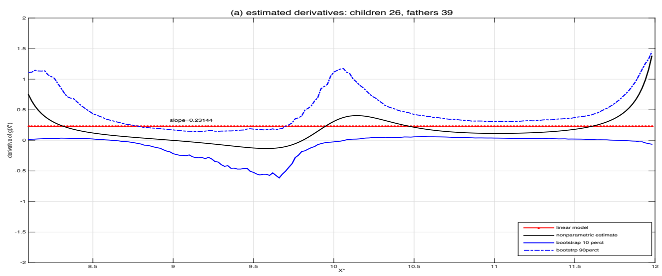


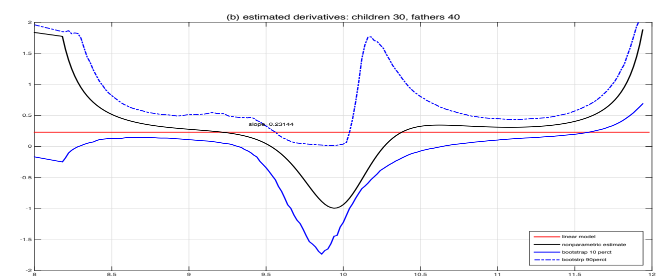
Table 4: Quartiles of Estimated Derivatives

	25-th		50-th		75-th	
	Estimate	90% C.I.	Estimate	90% C.I.	Estimate	90% C.I.
Overall	0.130	[-0.033,0.351]	0.263	[0.124,0.514]	0.483	[0.065,1.140]
White	0.111	[-0.099,0.356]	0.311	[0.111,0.572]	0.618	[0.035,1.478]
Male	0.153	[-0.073,0.454]	0.250	[0.068,0.587]	0.392	[0.140,1.010]

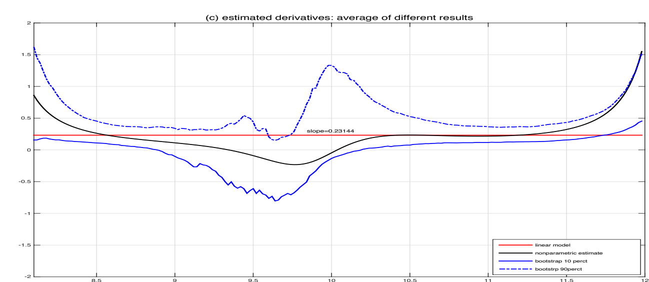
Figure 7: Robustness checks



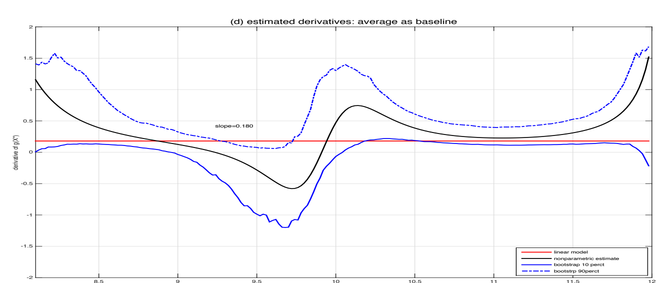
(a) children 26, fathers 39



(b) children 30, fathers 40



(c) average of different results



(d) average as baseline

Figure 8: Bounds of mobility function

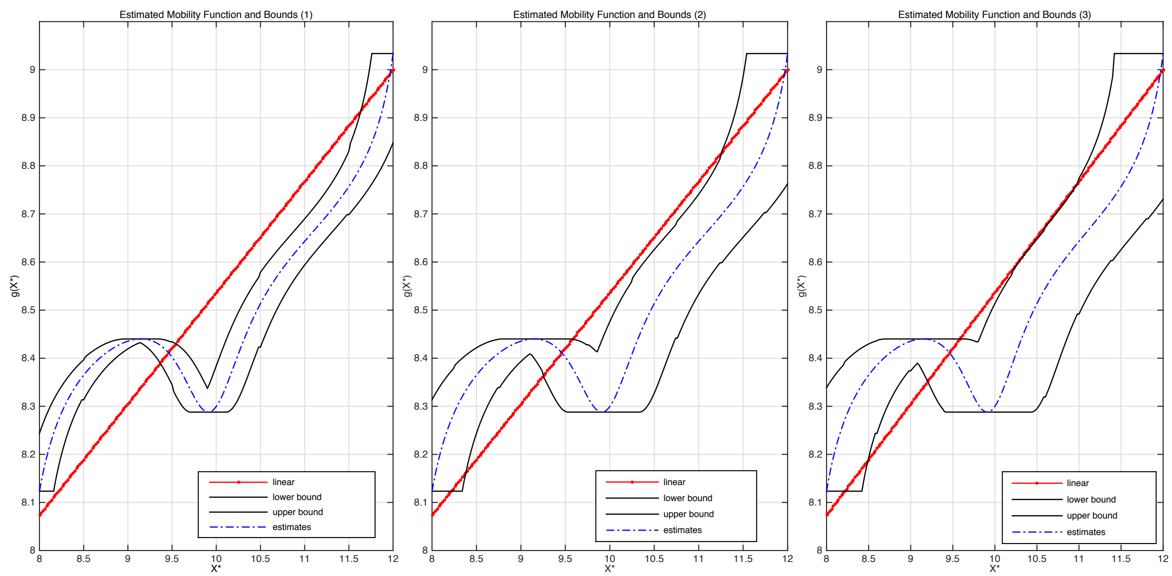


Figure 9: Sub-sample analysis