Bhattacharya, Debopam; Mazumder, Bhashkar

Working Paper

Nonparametric analysis of intergenerational income mobility with application to the United States

Working Paper, No. 2007-12

Provided in Cooperation with:
Federal Reserve Bank of Chicago

Suggested Citation: Bhattacharya, Debopam; Mazumder, Bhashkar (2010) : Nonparametric analysis of intergenerational income mobility with application to the United States, Working Paper, No. 2007-12, Federal Reserve Bank of Chicago, Chicago, IL

This Version is available at:
http://hdl.handle.net/10419/70601

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes. You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Nonparametric Analysis of Intergenerational Income Mobility with Application to the United States

Debopam Bhattacharya and Bhashkar Mazumder

REVISED
March 2010
WP 2007-12
Abstract

This paper concerns the problem of inferring the effects of covariates on intergenerational income mobility, i.e. on the relationship between the incomes of parents and future earnings of their children. We focus on two mobility measures: (i) traditional transition probability of movement across income quantiles and (ii) the probability that an adult child’s relative position exceeds that of the parents. We estimate the effect of possibly continuously distributed covariates on these measures using nonparametric regression. Here, the dependent variables are nonsmooth functions of marginal quantiles and relative ranks and so inference on them requires nontrivial modifications of standard nonparametric regression theory. Using our methods on data from the National Longitudinal Survey of Youth, we document that whites experience greater intergenerational mobility than blacks and that most of this gap can be accounted for by differences in cognitive skills during adolescence. The technical methods developed here have wider applicability to inference for nonparametric regression where the dependent variable is a nonsmooth functional of ranks or quantiles of one or more random variables.

Keywords: intergenerational mobility, upward mobility, nonparametric regression, Hadamard differentiability, black-white mobility gap

JEL codes: C14, D31

*Correspondence address: Department of Economics, Oxford University, Manor Road Building, Manor Road, OX1 3UQ, UK; phone number 44-7503-858289; Fax: +44 1865 271094, email address: debobhatta@gmail.com
1 Introduction

Intergenerational mobility (IGM, henceforth) refers to the extent of movement in economic status across generations. If parents’ economic status has a relatively small effect on children’s eventual economic status then this might suggest that there is room for individuals to succeed or fail irrespective of their circumstances at birth, suggesting greater equality of opportunity. There is clearly reason to think that mobility patterns might differ markedly across subgroups of the population. In the US, for instance, it is particularly interesting to know whether the prospects for upward mobility differ across racial groups, given the legacy of slavery and segregation. From a policy perspective, it is also important to understand the channels through which intergenerational income persistence arises. In this paper, we present and analyze a set of statistics, based on joint distributions of parent and child’s incomes, which contribute towards answering these substantive questions.

The empirical analysis of intergenerational income mobility has traditionally been based on two broad and complementary measures- (i) the intergenerational elasticity (IGE)\(^1\) which is simply the regression coefficient obtained by regressing (log) child’s permanent income on (log) parents’ permanent income and (ii) matrices of transition probabilities which measure the rates of movement across quantiles of the distribution over subsequent generations. The IGE simply relates mean economic status in one generation to economic status in the previous generation,\(^2\) while transition probabilities capture relationships between relative positions across generations including positions far away from the median. A further limitation of the IGE is that one cannot use it to analyze mobility differences across population subgroups with respect to the entire distribution. For example, the IGE for blacks only describes the rate at which earnings among black children regress to the black mean –not the mean of the entire distribution. In contrast, transition probabilities can be used to make statements concerning the movements of blacks across the income distribution of the entire population comprising both blacks and whites.

On the other hand, it is straightforward to measure effects of covariates on the IGE,

\(^{1}\)The intergenerational correlation (IGC) has also been used by many researchers. The IGC is qualitatively similar to the IGE and the two measures are equivalent when the variance in in income is unchanging across generations. See Corak (2006) or Bowles and Gintis (2002) for recent surveys of the empirical literature.

\(^{2}\)One may use quantile regressions to calculate IGE to avoid problems of outliers in and top-coding of earnings data.
one simply needs to include the covariates and their interactions with parents’ income as additional regressors and the statistical theory is straightforward. In contrast, a formal statistical theory for using covariates in transition matrices seems to be lacking. The development of such a theory therefore would combine the advantages of transition matrix based approaches together with the ability to investigate the effects of covariates on mobility. For instance, using PSID data to construct transition probabilities, Hertz (2005) has previously shown that mobility among blacks is low at the low end of the income distribution and high at the high end relative to whites. Conditional transition probabilities will let us infer to what extent these racial differences in mobility can be attributed to differences in human capital. Compared to the IGE, *conditional* transition probabilities allow us to answer a richer set of questions that are of more direct relevance for policy design.

From a methodological viewpoint, one can use standard parametric specifications—e.g. a probit—to model dependence of transition probability on covariates. Such methods are easy to use for estimation. However, this approach is theoretically problematic because it is unclear what type of joint distribution of errors will imply a probit form for transition probabilities; in particular, a bivariate normal error distribution will not. This is unlike probit regressions where normal errors for a latent equation lead to the probit likelihood for a binary outcome. Moreover, in the application, we show that restrictive parametric inference, as an approximation, can produce misleading qualitative conclusions (see fig. 8 and discussion in section 4.2.1, below). We therefore resort to nonparametric techniques.

The distribution theory for estimated *marginal* transition probabilities was previously developed in Formby, Smith and Zheng (2004), henceforth FSZ. When the relevant covariates are discrete, one can simply apply the results of FSZ within each covariate category to conduct inference on conditional transitions. But with continuously distributed covariates, the parameters of interest are infinite-dimensional and thus nonparametric smoothing methods are warranted. Trede (1998) has shown how continuous covariates can be included in analyzing a class of inequality reducing functionals that have been used to measure *intra-generational* mobility but not for measures based on transition probabilities. One relevant covariate of significant interest in the US is the son’s percentile score on Armed Forces Qualifying Test (often taken to be an indicator of cognitive skills in adolescent years—see discussion below) which runs from 1 to 99 in the data. Given the size of typical household surveys, analysis within each percentile will be very imprecise,
particularly when broken up by racial subgroups. Furthermore, small differences in AFQT percentiles, unlike differences in race, are unlikely to imply big changes in functional relationships. Therefore, treating AFQT score as a continuously distributed covariate will be the natural and correct approach. Our first result in the present paper concerns the relevant distribution theory for estimating conditional transition probabilities.

A practical drawback of using transition probabilities is its overtly disaggregate nature (i.e. there are an infinite number of transition probabilities depending on which quantiles are compared) and a summary measure of mobility across _relative_ income positions is useful for consolidating the information provided in transition matrices. In this paper, we introduce a new and intuitively simple measure of overall upward mobility in relative terms and establish distribution theories for inferring both its level and the effects of covariates on it. Our mobility measure for e.g. blacks is simply the probability that a black son’s income rank in his generation (including both blacks and whites) exceeds his parents’ rank in the prior generation. A key feature of this measure is that it probably captures what most policy-makers actually think of as mobility, namely to what extent are people doing better than their parents? It is a single summary measure, easy to interpret and its value does not depend on arbitrary discretization of income distributions, unlike existing measures of overall mobility (see section 2.2 below for more on this). In our application with US data, we show that black-white differences in mobility are much smaller when based on our measure of upward mobility compared to traditional (upward) transition probabilities. This is because (i) our measure captures small upward movements in relative positions of a large number of blacks which are ignored by transition probabilities and (ii) blacks typically need much larger upward mobility to surpass a common percentile threshold than whites (and thus increase their transition probability) because for almost any common threshold, black incomes below that threshold tend to be smaller than the white incomes below it.

1.1 Contributions

This paper adapts nonparametric regression theory to accommodate outcome variables involving non-smooth functions of initially estimated functionals. For transition probabilities, the outcome variable involves a non-smooth function of marginal quantiles of income; for upward mobility, it involves non-smooth functions of estimated ranks. The analysis of conditional transition probabilities is relatively easier because one can use
stochastic equicontinuity type arguments. So we state regularity conditions and outline the key steps in the proof. The analysis of marginal and conditional upward mobility requires independent analysis because the dependent variable involves estimated ranks and we present formal theorems with proofs, establishing distribution theories for them. To do this, we prove and use Hadamard differentiability of the map from the joint c.d.f. of father and son’s income to upward mobility and control first stage estimation errors using Hoeffding’s inequality.

Although the methods developed here are motivated by nonparametric analysis of IGM, they have more general applicability. For example, one can use these methods to analyze the relation between initial and retirement salaries of a single individual for each major industry and then compare industries by how economically mobile their workforce is. More generally, whenever the parameter of interest is a nonparametric regression or a functional thereof but the dependent variable involves preliminary components estimated from the same dataset, the methods developed here can be utilized to get the respective distribution theories. For instance, consider estimating the effects of covariates like study hours on the probability that a high school senior is in the top 5% of his graduating class. Class rank reveals a student’s relative performance with respect to his peers in the same school and so implicitly controls for inter-school variations in absolute grades. The methodology of this paper can be used to conduct statistical inference on regressions of such relative outcomes on covariates. Note that such relative regressions tell us something very different from quantile regressions. The former captures effects of covariates on the relative position in the marginal distribution of the dependent variable but the latter pertains to relative position in its conditional (on the covariate) distribution.

We also provide a number of novel substantive contributions to the empirical literature on IGM. We are the first to use a large and nationally representative sample of blacks and whites to estimate interracial differences in IGM.3 We document that there are sizable differences in upward transition probabilities between blacks and whites. We apply nonparametric methods to investigate how a continuous measure of cognitive skill, the Armed Forces Qualifying Test (AFQT), affects the transition probabilities for blacks and whites. AFQT scores have been previously used to account for other aspects of black-white inequality (Neal and Johnson, 1996; Cameron and Heckman, 2001) but, to our knowledge,

3Hertz (2005) also shows large black-white differences using the PSID but for reasons we discuss below it is unclear whether the PSID sample of blacks is representative.
have not been used in studies of IGM. Interestingly, we show that this variable can account for most of the gap between blacks and whites in their ability to rise from the bottom quintile of the income distribution. We also find similar results with respect to inter-racial differences in our new measure of upward mobility. We show that alternative parametric methodology, such as running probit regressions for conditional transitions on our data lead to misleading conclusions. The reader may note that in accordance with virtually all of the existing literature on economic mobility including all the papers cited above and below, what we propose here are essentially descriptive measures and one should be cautious in attaching causal interpretations to them. There has indeed been some discussions in the literature that traditional measures of intergenerational mobility may not always reflect equality of opportunity (Van de Gaer et al, 2001) and that it is not always clear which policies may be appropriate for fostering mobility (e.g. Roemer (2004) Swift (2005), Jencks and Tach (2006)). An in-depth investigation of these questions is outside the scope of the present paper and is left to future research.

1.2 Plan of the paper

The plan of the paper is as follows. Section 2 describes the parameters of interest, section 3 discusses the asymptotic distribution theory, section 4 presents the application using NLSY data. Finally, section 5 concludes. All proofs are collected in section 6. In the statements of the theorems and in the proofs, $c$ will denote a generic positive constant not always having the same value and whenever derivatives (or Lebesgue densities) are defined, they are implicitly assumed to exist.

2 Parameters of interest

We first describe the parameters of interest based on transition probabilities and then those related to our new measure of upward mobility.

2.1 Conditional transition probabilities

Let $\zeta_1$ denote the bottom $s$th quantile and $F_0(\cdot)$ the c.d.f. of the overall income distribution for sons and let $(\zeta_{01}, \zeta_{02})$ be the quantiles corresponding to percentiles $0 < t_1 < t_2 < 1$ of fathers and $F_1(\cdot)$ the corresponding c.d.f., respectively. Then, the transition probability
measures the probability that a son is at or above $\zeta_1$, conditional on his father being between $\zeta_{01}$ and $\zeta_{02}$ i.e.

$$
\theta (s, (t_1, t_2)) = \frac{\Pr [F_1 (Y_1) \geq s, t_1 \leq F_0 (Y_0) \leq t_2]}{\Pr [t_1 \leq F_0 (Y_0) \leq t_2]} = \frac{\Pr [Y_1 \geq \zeta_1, \zeta_{01} \leq Y_0 \leq \zeta_{02}]}{\Pr [\zeta_{01} \leq Y_0 \leq \zeta_{02}]}.
$$

(1)

Now, $\theta (s, (t_1, t_2))$ can be decomposed by level of discrete and continuous covariates $X$ such as age and education of the father and/or the son as

$$
\theta (s, (t_1, t_2)) = \int \frac{\Pr [Y_1 \geq \zeta_1, \zeta_{01} \leq Y_0 \leq \zeta_{02} | X = x]}{\Pr [\zeta_{01} \leq Y_0 \leq \zeta_{02}]} dF (x)
= \int \theta (x; s, (t_1, t_2)) dF (x),
$$

where

$$
\theta (x; s, (t_1, t_2)) = \frac{\Pr [Y_1 \geq \zeta_1, \zeta_{01} \leq Y_0 \leq \zeta_{02} | X = x]}{\Pr [\zeta_{01} \leq Y_0 \leq \zeta_{02}]}.
$$

(2)

So when $X$ denotes a son’s educational attainment, then the difference in $\theta (x; s, (t_1, t_2))$ and $\theta (x'; s, (t_1, t_2))$ represents the change in the son’s probability of exceeding the $s$th quantile $\zeta_1$ when $X$ changes from $x$ to $x'$, where the aggregate population of interest is all families (and not only those whose sons’ values of $X$ are $x$) whose fathers’ incomes were between $\zeta_{01}$ and $\zeta_{02}$. When $X$ denotes the father’s education, $\theta (x; s, (t_1, t_2))$ can be used to infer how effective is a highly educated father in improving his son’s condition relative to a less educated father who was in the same quantile of income as him. Note also that by differencing the transition probabilities defined above, one can get the ones defined in, e.g. FSZ (2004) or Shorrocks (1978).

One can immediately define a measure of conditional mobility, based on our definition above and use it to measure the black-white difference in transition at each value of $x$ by

$$
\theta_B (x; s, (t_1, t_2)) - \theta_W (x; s, (t_1, t_2)),
$$

4Note that the denominator is not conditioned on $X$ because we want to see how the numerator (and not the denominator) of (1) can be decomposed by $X$. In particular,

$$
\neq \int \frac{\Pr [Y_1 \leq \zeta_1, \zeta_{01} \leq Y_0 \leq \zeta_{02} | X = x]}{\Pr [\zeta_{01} \leq Y_0 \leq \zeta_{02} | X = x]} dF (x).
$$


where, e.g.,

\[
\theta_{B}(x; s, (t_1, t_2)) = \frac{Pr[Y_1 \geq \zeta_1, \zeta_{01} \leq Y_0 \leq \zeta_{02} | X = x, Black = 1]}{Pr[\zeta_{01} \leq Y_0 \leq \zeta_{02} | Black = 1]}.
\]

### 2.2 Upward mobility

We now formally introduce our new measure of upward mobility. We first present the analytic expressions and then discuss the substantive features which make our measure both intuitively appealing and analytically different from measures based on transition probabilities.

Our direct measure of upward mobility is simply the probability that the son’s percentile rank in the overall income distribution of his generation exceeds that of his parents’ in the income distribution of the parents’ generation. We believe that this measure more closely conforms to what most people think of as economic mobility. Indeed much of the recent attention in the popular press concerning IGM has been couched in terms of the prospects for upward mobility for those starting in the bottom of the distribution.\(^5\)

Policy-makers also tend to be more concerned about mobility with respect to what it signifies about the prospects for economic gains among the poor and disadvantaged groups and whether government interventions are necessary to foster greater upward mobility.

Let \(Y_0, Y_1\) denote parent and son’s income with respective marginal c.d.f.’s \(F_0\) and \(F_1\). Then for fixed \(0 < s_1 < s_2 < 1\), we define upward mobility for families between percentiles \(s_1\) and \(s_2\) by an extent \(\tau \in [0, 1 - s_2]\) as

\[
v(\tau, s_1, s_2) = \Pr(F_1(Y_1) - F_0(Y_0) > \tau | s_1 \leq F_0(Y_0) \leq s_2) = 1 - \frac{\Pr(F_1(Y_1) - F_0(Y_0) < \tau, s_1 \leq F_0(Y_0) \leq s_2)}{s_2 - s_1} = 1 - \frac{1}{s_2 - s_1} \int_{F_0^{-1}(s_1)}^{F_0^{-1}(s_2)} \int_1^{F_1^{-1}(F_0(y_0) + \tau)} f(y_0, y_1) dy_1 dy_0. \tag{3}
\]

For the purpose of this paper, the leading case of interest is where \(\tau = 0\) which gives the probability that the son’s relative position exceeds that of the father. But we develop inference theory for a general \(\tau\). One can alternatively define upward mobility by conditioning on \(F_0(Y_0) = s\) rather than \(s_1 \leq F_0(Y_0) \leq s_2\), which might be easier to interpret.

\(^5\)For example, a front page article in the Wall Street Journal described class mobility as follows:

"...The promise that a child born in poverty isn’t trapped there remains a staple of America’s self-portrait." (Wessel, 2005).
when we compare upward mobility between blacks and whites below. The reason is that
given any interval \((s_1, s_2)\) for \(F_0(Y_0)\), blacks are likely to be concentrated in the lower
tail of \((s_1, s_2)\), leading us to compare them with relatively wealthier whites. However,
conditioning on \(F_0(Y_0) = s\) would entail additional smoothing since \(F_0(Y_0)\) would be
continuously distributed. By making the length of the interval \((s_1, s_2)\) small, we would
both avoid this smoothing and yet make blacks and whites comparable to start with.

Introducing covariates \(X\) into the analysis, define conditional upward mobility at val-
ues of \(X = x\) as

\[
v_c(\tau, s_2, s_1; x) = \frac{\Pr(F_1(Y_1) - F_0(Y_0) > \tau, s_1 \leq F_0(Y_0) \leq s_2|X = x)}{\Pr(s_1 \leq F_0(Y_0) \leq s_2)}
\]

\[
= \frac{\Pr(F_1(Y_1) - F_0(Y_0) > \tau, s_1 \leq F_0(Y_0) \leq s_2|X = x)}{s_2 - s_1}.
\] (4)

This measure is analogous to (2) above. The idea is that we start with all families
where the father was between the \(s_2\)th and \(s_1\)th percentile. This ensures that all the
corresponding sons have equal "space to move up". With these families constituting our
population, we evaluate the extent of upward mobility for children at various values \(x\) of
\(X\).

Below, we will derive the statistical distribution theory for estimates of \(v(\tau, s_2, s_1)\) and
\(v_c(\tau, s_2, s_1; x)\). In the application, we will contrast overall upward mobility among blacks
versus whites and then analyze how inclusion of relevant covariates affects this difference.

It is useful to note that one can alternatively define overall mobility based on transition
matrices after incorporating effects of covariates. Consider a transition matrix based
on an arbitrary M-class discretization of the marginal distributions of \(Y_0\) and \(Y_1\): 
\(\tilde{\Theta} = \{\tilde{\theta}(j, k)\}_{j, k = 1, \ldots, M}\). Then Shorrock's measure of unconditional mobility is given by

\[
M_1 = \frac{K - \text{trace}(\tilde{\Theta})}{K - 1} = 1 - \frac{\sum_{j=1}^{K} \tilde{\theta}(j, j) - 1}{K - 1}.
\]

One can incorporate covariates into the above formula and define

\[
M_1(x) = 1 - \frac{\sum_{j=1}^{K} \tilde{\theta}(j, j; x) - 1}{K - 1} \] (5)

where

\[
\tilde{\theta}(j, j; x) = \frac{\Pr(\xi_j \leq Y_1 \leq \xi_{j+1}, \xi_j \leq Y_0 \leq \xi_{j+1}|X = x)}{\Pr(\xi_j \leq Y_0 \leq \xi_{j+1})},
\]
and $\zeta_j, \xi_j$ denote the $j$th marginal quantiles of $(Y_1, Y_0)$ respectively. Given the simple linear relation (5), inference on $M_1(x)$ will follow straightforwardly from inference on $\tilde{\theta}(j, j; x)$. However, this measure will depend crucially on the discretization employed which is clearly an undesirable feature. Altering the above formulas to allow for a continuous transition matrix seem complicated and we leave that to future research. Instead, we focus on our measure $v(0, s_1, s_2)$, which, we believe, is much closer to what is commonly understood as mobility and whose enhancement appears to be a stated goal of liberal policy-making. This single summary measure does not employ any discretization, has an immediate intuitive interpretation and, unlike the IGE, is based on a direct comparison of the relative positions between fathers and son.

Another important feature of $v(0, s_1, s_2)$ is that it counts small upward movements in relative positions which are ignored by transition probabilities. Comparing

$$v(0, s_1, s_2) = \Pr(F_1(Y_1) - F_0(Y_0) > 0, s_1 \leq F_0(Y_0) \leq s_2)$$

and

$$1 - \theta(s_2, (s_1, s_2)) = \Pr[F_1(Y_1) > s_2 | s_1 \leq F_0(Y_0) \leq s_2],$$

one can see that unlike $1 - \theta(s, (s_1, s_2))$, $v(0, s_1, s_2)$ is counting all sons whose ranks exceeded their fathers’ but did not exceed $s_2$. In our application, this makes a substantial impact on black-white differences in mobility (see figure 4 below). We find that whites appear to be much more upwardly mobile relative to blacks when measured by the transition probability of moving out of a given quantile. The difference between whites and blacks is much smaller when measured in terms of our upward mobility index. The first reason for this is that many black sons make relatively small upward movements which are missed by $\theta(s_2, (s_1, s_2))$ but captured by $v(0, s_1, s_2)$. The second reason is that incomes of white fathers tend to be larger than that of black fathers (figure 5) under any fixed threshold. Therefore, for any overall percentile $F_0^{-1}(s)$, sons born to black fathers below $F_0^{-1}(s)$ need a larger increase in absolute income to surpass it compared to sons of white fathers below $F_0^{-1}(s)$. This suggests that even if rates of upward mobility are similar across groups, transition probabilities are likely to be much larger for whites.

---

6 The problem is that $\int_0^1 \theta(s, s) ds$ is not a probability, unlike $\sum_{j=1}^{K} \tilde{\theta}(j, j)$ and one needs to replace $\theta(s, s)$ with a density type analog before integrating. What that analog should be is not obvious.
2.3 Measurement Error

Researchers working on earnings mobility have paid particular attention to measurement error in sons’ and fathers’ earnings in the context of intergenerational regressions (c.f., Haider and Solon, 2006).\(^7\) It is interesting to note that all our measures of mobility are based on the relative positions of individuals in the population. So if ranks of individuals are preserved despite measurement errors, then our measures will not be affected by the fact that we have erroneous earnings measures. One specific example of this is where reported earnings are a monotone function of true earnings—i.e. if for two people 1 and 2, true incomes satisfy \(y_1^* > y_2^*\), then their reported incomes satisfy \(y_1 > y_2\). This can be easily consistent with non-classical measurement error, i.e., \(y - y^*\) being negatively correlated with true earnings \(y^*\) (Bound et al (1994)). In this case, all our measures based on reported \(y\) will be identical to those based on \(y^*\). With more general types of measurement error, using time averaged incomes or earnings, as is common in the IGM literature, can partially mitigate the effect of purely random measurement error in addition to providing more reliable estimates for permanent income. This is the approach we follow in the application below. Finally our measures of son’s earnings are taken around the age of 40 when life-cycle bias (Haider and Solon (2006)) is minimized.

2.4 Related Parameters

Based on the parameters (2) and (4), one can define the corresponding marginal effects by differentiating them with respect to the regressor values. One can also consider a summary measure of marginal effects by calculating density-weighted average derivatives, a la Powell, Stoker and Stock (1989). For brevity, we do not pursue these quantities here.

3 Estimation and distribution theory

We now turn to estimation of the parameters and derivation of their asymptotic properties. Note that we have defined 4 parameters above, viz., (1), (2), (3), (4). FSZ (2004) had analyzed only (1) and so, in what follows, we will derive the distribution theory for the

\(^7\)We are aware of only one study that has examined the effect of measurement error in the context of transition probabilities. O’Neill et al (2007) show that measurement error can induce a modest bias in transition probabilities compared to regressions and that this bias may vary at different points of the distribution.
other three. The analysis of (2) requires slight modification of standard kernel regression theory. We only provide an outline of proof by pointing out the modifications needed. But the estimators of (3) and (4) are fundamentally harder to analyze owing to the presence of \( \hat{F}_1(\cdot) \) and \( \hat{F}_0(\cdot) \) in the definition of the dependent variables. Our analysis of these will rely crucially on the idea of Hadamard differentiability and we will use Hoeffding’s inequality to control the errors involved in the estimation of \( \hat{F}_1(\cdot) \) and \( \hat{F}_0(\cdot) \). These results are formally stated as the two theorems in the paper and their proofs are presented in the appendix.

Also, without loss of generality, we will assume that the supports of \( Y_0 \) and \( Y_1 \) are subsets of \([1, \infty)\). Note that all our transition probability measures are based on quantiles and so fixed location shifts in either variable does not affect any of the measures.

### 3.1 Conditional transition probability

Recall from (2) the definition of conditional transition probabilities. We will first state the distribution theory for

\[
\theta(x; s, t) = \frac{\Pr \{ Y_1 \leq \zeta_1, Y_0 \leq \zeta_0 | X = x \}}{\Pr \{ Y_0 \leq \zeta_0 \}} \tag{6}
\]

where \( \zeta_0 \) denotes the \( t \)th quantile of \( Y_0 \) and \( \zeta_1 \) denotes the \( s \)th quantile of \( Y_1 \). One can obtain (2) from (6) by simple subtraction. Now, (6) can be estimated by

\[
\hat{\theta}(x; s, t) = \frac{\frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) 1 \left( Y_{1i} \leq \hat{\zeta}_1, Y_{0i} \leq \hat{\zeta}_0 \right)}{\left( \frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) \right) \times \left( \frac{1}{n} \sum_{i=1}^{n} 1 \left( Y_{0i} \leq \hat{\zeta}_0 \right) \right)}, \tag{7}
\]

where \( K(\cdot) \) is a \( d \)-dimensional kernel and \( \sigma_n \) is a sequence of bandwidths. Let \( f(y_0, y_1|x) \) denote the density of \( (Y_0, Y_1) \) conditional on \( X = x \) and define

\[
\phi(x, \zeta_0, \zeta_1) = \Pr \{ Y_1 \leq \zeta_1, Y_0 \leq \zeta_0 | X = x \} = \int_1^{\zeta_1} \int_1^{\zeta_0} f(y_0, y_1|x) \, dy_0 \, dy_1
\]

\[
\hat{\phi}(x, \zeta_0, \zeta_1) = \frac{\frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) 1 \left( Y_{1i} \leq \hat{\zeta}_1, Y_{0i} \leq \hat{\zeta}_0 \right)}{\left( \frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) \right) \times \left( \frac{1}{n} \sum_{i=1}^{n} 1 \left( Y_{0i} \leq \hat{\zeta}_0 \right) \right)}
\]

\[
\phi_1(x, \zeta_0, \zeta_1) = \int_1^{\zeta_0} f(y_0, \zeta|x) \, dy_0, \quad \phi_0(x, \zeta_0, \zeta_1) = \int_1^{\zeta_1} f(\zeta_0, y_1|x) \, dy_1
\]

\[
v^2(x) = \phi(x, \zeta_0^0, \zeta_1^0) \times (1 - \phi(x, \zeta_0^0, \zeta_1^0))
\]
The asymptotic distribution for conditional (on covariates) transition matrices is given by the following proposition. We state this and subsequent propositions and theorems in terms of a $d$-dimensional $X$ all of whose components are continuously distributed. For discrete covariates, the analysis is identical to that for the marginal (i.e. unconditional) measures.

We now state a set of general regularity conditions which will imply zero-mean asymptotic normality for the Nadaraya-Watson estimated regression of the unobserved random variable

$$W \equiv 1 \{ Y_1 \leq \zeta_1, Y_0 \leq \zeta_0 \}$$

on $X$, evaluated at $X = x$. These conditions are standard (for textbook treatments, see Bierens (1994), theorem 10.2.1 or Pagan and Ullah (1999) theorem 3.5, 3.6) but we state them here to make the subsequent proposition and theorem statements self-sufficient.

**Condition NW**

**NW1.** $X$ is a $d$-dimensional continuously distributed random variable with Lebesgue density $f(\cdot)$ which is positive at $x$.

**NW2.** The data $(X_i, Y_{1i}, Y_{0i})$ are i.i.d.

**NW3.** $K(\cdot)$ is a Borel-measurable, bounded and real-valued kernel function with $d$-dimensional argument, satisfying (i) $\int K(a)da = 1$, $\int aK(a)da = 0$, $\int a_i^2K(a)da < \infty$ for $i = 1, \ldots, d$, (ii) $\int |K(a)|da < \infty$, (iii) for some $\delta > 0$, $\int |K(a)|^{2+\delta}da < \infty$.

**NW4.** The bandwidth sequence $\sigma_n$ satisfies $\lim_{n \to \infty} \sigma_n = 0$, $\lim_{n \to \infty} n\sigma_n^d = \infty$ and $\lim_{n \to \infty} \sigma_n^2(\sigma_n^d)^{1/2} = 0$; $\frac{1}{\sigma_n^d} K\left(\frac{a-b}{\sigma_n}\right)$ is uniformly bounded for $a, b \in \text{support}(X)$.

**NW5.** The functions $f(\cdot)$ and $f(\cdot) \times \phi(\cdot; \zeta_0, \zeta_1)$ and their derivatives up to order 2 are continuous and uniformly bounded.

**Proposition 1** Suppose that conditions **NW1-5** are satisfied. Assume further that for $X = x$, $(Y_0, Y_1)$ admits a nonnegative joint density w.r.t. the Lebesgue measure everywhere on the joint support. Further, the function $\phi(x, \cdot, \cdot)$ is Lipschitz with respect to the Euclidean norm $\| \cdot \|$: 

$$\| \phi(x, \zeta_0, \zeta_1) - \phi(x, \tau_0, \tau_1) \| \leq \| (\zeta_0, \zeta_1) - (\tau_0, \tau_1) \| \delta(x)$$

with $\delta(\cdot)$ uniformly bounded on the support of $X$. Then we have

$$\left( n\sigma_n^d \right)^{1/2} \left( \hat{\phi}(x, \hat{\zeta}_0, \hat{\zeta}_1) - \phi(x, \zeta_0^0, \zeta_1^0) \right) \xrightarrow{d} N \left( 0, \frac{v^2(x)}{f(x)} \right) \int K^2(u)du$$

where $v^2(x) = \phi(x, \zeta_0^0, \zeta_1^0) \times (1 - \phi(x, \zeta_0^0, \zeta_1^0))$. 

13
Proof. We provide an outline of the proof in the appendix. ■

The noteworthy point is that the distribution of \( \hat{\phi} (x, \hat{\zeta}_0, \hat{\zeta}_1) \) and that of the infeasible estimator \( \hat{\phi} (x, \zeta_0, \zeta_1) \) are identical because \( \left( \hat{\zeta}_0, \hat{\zeta}_1 \right) \) converges at the parametric \( \sqrt{n} \) rate but \( \hat{\phi} (x, \zeta_0, \zeta_1) \) converges to \( \phi (x, \zeta_0, \zeta_1) \) slower than \( \sqrt{n} \)-rate and at the same time an equicontinuity argument can be used to handle the non-smoothness of \( 1 (Y_1 \leq \zeta_1, Y_0 \leq \zeta_0) \) in the \( \zeta \)'s. Note also that we have used an "undersmoothed" estimator to achieve bias reduction and omitted bounded moment assumptions on the errors because the dependent variable and \( (x; \zeta_0, \zeta_1) \) lie in \([0; 1]\).

Returning to (7), by the standard delta method we get that for demographic group \( k \),

\[
(n\sigma_n^d)^{1/2} \left\{ \hat{\theta}_k (x; s, (t_1, t_2)) - \theta_k^0 (x; s, (t_1, t_2)) \right\} = (n\sigma_n^d)^{1/2} \left( \frac{\hat{\phi}_k (x, \hat{\zeta}_0, \hat{\zeta}_1) - \phi_k (x, \zeta_0^0, \zeta_1^0)}{\pi_k} \right) + o_p (1),
\]

where

\[
\phi_k (x, \zeta_0, \zeta_1) = \Pr \{ Y_i \leq \zeta_1, Y_0i \leq \zeta_0 | X_i = x, k \}, \quad \pi_k = \Pr [ Y_0 \leq \zeta_0 | k ].
\]

This follows because the estimator of \( \pi_k \) will converge at parametric rates and will not affect the distribution of \( \hat{\theta}_k (x; s, (t_1, t_2)) \). One thus gets that

\[
(n\sigma_n^d)^{1/2} \left\{ \hat{\theta}_k (x; s, (t_1, t_2)) - \theta_k^0 (x; s, (t_1, t_2)) \right\} \xrightarrow{d} N \left( 0, \frac{\sigma_k^2 (x)}{\pi_k f_k (x)} \int K^2 (u) du \right) \quad (9)
\]

where \( \sigma_k^2 (x) = \phi_k (x, \zeta_0^0, \zeta_1^0) \times (1 - \phi_k (x, \zeta_0^0, \zeta_1^0)) \) and \( f_k (x) \) is simply the density of \( X \) at \( x \) for demographic group \( k \).

3.2 Distribution of the gap in levels

In this paper, our substantive interest is focused around the black-white gap in conditional transition probabilities. In that context, it is interesting to note that the estimators \( \hat{\theta}_B (x; s, (t_1, t_2)) \) for blacks and \( \hat{\theta}_W (x; s, (t_1, t_2)) \) for whites will be asymptotically uncorrelated. To see this formally, let \( y_i = 1 (Y_0i \leq \zeta_0) \) and let

\[
\hat{y}_B = \frac{1}{n} \sum_{i=1}^n B_i y_i \equiv \bar{y}_B, \quad g_B = \frac{E (\bar{y}_B)}{E (\bar{d}_B)} = \frac{\mu_B}{\pi_B}
\]
Asymptotically,
\[
\sqrt{n} (\hat{g}_B - g_B) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (B_i y_i - \mu_B) - \frac{\mu_B}{\pi_B^2} \sum_{i=1}^{n} (B_i y_i - \mu_B) + o_p(1)
\]
\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} B_i y_i - \frac{\mu_B}{\pi_B^2} \sum_{i=1}^{n} B_i y_i + o_p(1)
\]
\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} B_i \left( y_i - \frac{\mu_B}{\pi_B} \right) + o_p(1).
\]
So the asymptotic covariance is given by,
\[
E \left[ \frac{B_i}{\pi_B} \left( y_i - \frac{\mu_B}{\pi_W} \right) \right] W_i \left( y_i - \frac{\mu_W}{\pi_W} \right) = 0
\]
since \(B_i W_i \equiv 0\) and
\[
E \left[ \frac{W_i}{\pi_W} \left( y_i - \frac{\mu_W}{\pi_W} \right) \right] = E \left[ \frac{W_i y_i}{\pi_W} - \frac{\mu_W}{\pi_W} W_i \right] = E \left[ \frac{W_i y_i}{\pi_W} \right] - \frac{\mu_W}{\pi_W} = 0.
\]
Thus, we get that
\[
(n \sigma_n^2)^{1/2} \left\{ \hat{\theta}_B (x; s, (t_1, t_2)) - \theta^0_B (x; s, (t_1, t_2)) - \hat{\theta}_W (x; s, (t_1, t_2)) + \theta^0_W (x; s, (t_1, t_2)) \right\}
\]
d\rightarrow N \left( 0, \left\{ \frac{\sigma_B^2 (x)}{\pi_B f_B (x)} + \frac{\sigma_W^2 (x)}{\pi_W f_W (x)} \right\} \right) \int K^2 (u) du
\]
(10)
where \(\sigma_W^2 (x) = \phi_W (x, \zeta_0^0, \zeta_0^0) \times (1 - \phi_W (x, \zeta_0^0, \zeta_0^0))\) and \(f_W (x)\) is simply the density of \(X\) for whites at \(x\). This result simplifies the calculation of asymptotic variances of the black-white gap in conditional transition probabilities since no covariance term needs to be calculated.

### 3.3 Marginal upward mobility

We now consider the notationally simpler version of \(v (\tau, s)\) defined in (3)
\[
v (\tau, s) = \Pr (F_1 (Y_1) - F_0 (Y_0) > \tau | F_0 (Y_0) \leq s)
\]
(11)
which can be estimated by
\[
\hat{v} (\tau, s) = 1 - \frac{1}{n} \sum_{i=1}^{n} 1 \left( \hat{F}_1 (y_{1i}) \geq \hat{F}_0 (y_{0i}) + \tau \right) 1 \left( \hat{F}_0 (y_{0i}) \leq s \right)
\]
(12)

---

---

---

---
where

\[ \hat{F}_1(y_{1i}) = \frac{1}{n} \sum_{j \neq i} 1(y_{1j} \leq y_{1i}). \]

We will now derive the asymptotic distribution of \( \hat{v}(\tau, s) \). Let \( F(., .) \) denote the joint c.d.f. of \((Y_0, Y_1)\) with corresponding joint density \( f(., .) \). Then for fixed \( s, \tau \), one may view \( v(\tau, s) \) as a functional \( v(F) \). We can therefore estimate it by \( v(\hat{F}) \), where \( \hat{F} \) denotes the usual empirical c.d.f.. We will obtain a large sample distribution of \( v(\hat{F}) \) using the functional delta method by showing that the functional \( F \mapsto v(F) \) is Hadamard-differentiable.\(^9\)

If one assumes that the joint density of \((Y_0, Y_1)\) is bounded away from zero on a compact support, then the proof of Hadamard differentiability is considerably simpler. This assumption may be tenable if the population of interest excludes families with "abnormally" high and low earnings in either generation which is typically where the density will be close to zero. However, for the sake of greater generality, we dispense with this assumption and establish Hadamard differentiability under more general tail conditions on the joint density and its partial derivatives.

Since all our measures are robust to monotone transformation of the income variables, we will continue assume that the support of the income variables is contained in \([1, \infty)\).

We are now ready to state a formal theorem. Let \( f(y_0, y_1) \) and \( f^{(0)}(y_0, y_1) \) denote respectively the joint density of \((Y_0, Y_1)\) and its derivative w.r.t. the first argument, evaluated at the point \((y_0, y_1)\). Let \( f_1(y_1) \) denote the marginal density of \( Y_1 \) and let \( c \) denote a generic positive constant. We now introduce a set of conditions concerning the tail behavior of the densities and their derivatives evaluated at the marginal quantiles. The exact forms of these conditions arise in the proof of Hadamard differentiability of the relevant functionals in the following theorem and, as such, should be viewed as "high-level" ones. As mentioned above, if one is willing to assume that the joint density of \((Y_0, Y_1)\) is bounded away from zero on a compact support, then these conditions are not necessary for proving theorem 1.

**Condition :** (Ai) for some \( \alpha > 1 \), we have \( f_1(x) \geq \frac{c}{x^\alpha} \) for \( x \) large enough which

\(^9\)The concept of Hadamard differentiability has been used before in the context of analyzing features of univariate income distributions, c.f. Bhattacharya (2007) and Barrett and Donald (2000). The results obtained here involve more complicated functionals of bivariate distribution functions and are not related to the results in the above papers.
also implies that \(F_1^{-1}(u) < c(1-u)^{\frac{1}{\alpha}}\), \(f^{(0)}(y_0, F_1^{-1}(F_0(y_0) + \tau)) \leq \frac{\sigma}{y_0^{1+\alpha}}\) for some \(\alpha_0 > 0\), \((\text{Aiii})\) for some \(\varepsilon > 0\), \(1 - F_0(y_0) > cy_0^{(1+\varepsilon-\alpha_0)(\alpha-1)}\) and \((\text{Aiv})\)

\[
\int_1^\infty (1 - F_0(y_0))^{\frac{\alpha}{c}} f(y_0, F_1^{-1}(F_0(y_0) + \tau)) \, dy_0 < \infty.
\]

It is also interesting to note that if the tails of \((Y_0, Y_1)\) have a joint Pareto distribution, then all of these conditions are automatically satisfied. To see this, assume that \((Y_0, Y_1)\) satisfy

\[
\Pr(Y_0 \geq y_0, Y_1 \geq y_1) = \frac{1}{(1 + (y_0 - 1) + (y_1 - 1))^\gamma}
\]

for all \(y_0, y_1 \geq 1\) for some \(\gamma > 0\). Then their joint density is given by

\[
f(y_0, y_1) = \frac{\gamma (\gamma + 1)}{(1 + (y_0 - 1) + (y_1 - 1))^{\gamma+2}}.
\]

Then one may verify that conditions \(\text{A}(i)-\text{A}(iv)\) are satisfied with \(\alpha = \gamma + 1\), \(\alpha_0 = \gamma + 2\) and \(\varepsilon = 1 + \gamma + \gamma (\gamma + 1)\).

An exactly symmetric set of conditions are assumed to hold for the marginal density \(f_0(\cdot)\) of \(Y_0\) as well.

\((\text{Bi})\) for some \(\beta > 1\), we have \(f_0(x) \geq \frac{c}{x^\beta}\) for \(x\) large enough which also implies that \(F_0^{-1}(u) < c(1-u)^{ \frac{1}{\beta}}\), \((\text{Bii})\) \(f^{(0)}(F_0^{-1}(s), y_1) \leq \frac{\sigma}{y_1^{\gamma+1}}\) for some \(\beta_0 > 0\), \((\text{Biii})\) for some \(\delta > 0\), we have \(1 - F_1(y_1) > cy_1^{(1+\delta-\beta_0)(\beta-1)}\) and \((\text{Biv})\)

\[
\int_1^\infty (1 - F_1(y_1))^{\frac{\beta}{c}} f(F_0^{-1}(s), y_1) \, dy_1 < \infty.
\]

To show that the map \(F \mapsto v(F)\) is Hadamard-differentiable, let \(\tilde{D}[1,\infty)\) denote the space of bivariate c.d.f.’s on \([1,\infty)\), satisfying conditions \((\text{Ai})-(\text{Aiv})\) and \((\text{Bi})-(\text{Biv})\). Denote by \(D_0\) the space of sample paths corresponding to the composite Brownian bridge \([G_\lambda \circ F]\) where \(G_\lambda\) is a standard Brownian bridge and \(F\) is any c.d.f. in \(\tilde{D}[1,\infty)\). Let \(D = \tilde{D}[1,\infty) \cup D_0\), equipped with the supremum norm. We want to show Hadamard differentiability of the map \(F \mapsto v(F)\) as a map from the normed vector space \(D\) to \(\mathbb{R}\).

Consider perturbations \(F_t(y_0, y_1) = F(y_0, y_1) + tH_t(y_0, y_1) \in \tilde{D}[1,\infty)\) with \(H_t \rightharpoonup H \in D_0\), uniformly as \(t \to 0\). We want to show that

\[
\left| \frac{v(F_t) - v(F)}{t} - v'_F(H) \right| \to 0 \quad \text{as} \quad t \to 0
\]

for a linear functional \(v'_F(\cdot)\) which is a map from \(\tilde{D}[1,\infty)\) to \(\mathbb{R}\).

\footnote{Condition \((\text{A(iv)})\) is like a moment condition; recall that for a positive random variable \(X\) with marginal c.d.f. \(G(\cdot)\) and support \(A\), the quantity \(\int_A (1 - G(x)) \, dx\) equals \(E(X)\).}
Theorem 1 Under conditions (Ai)-(Aiv) and (Bi)-(Biv), the map $F \mapsto v(F)$ from $D \to \mathbb{R}$, defined as

$$v(F) = \int_{1}^{F^{-1}(s)} \int_{1}^{F^{-1}(F_0(y_0) + \tau)} f(y_0, y_1) dy_1 dy_0$$

for any fixed $s, \tau \in (0, 1)$ is Hadamard differentiable at $F$ tangentially to $D_0$. The derivative at $F$ in the direction $H$ is given by the linear functional $v'_F(.)$ defined as

$$v'_F(H) = \left( \frac{H_0[F^{-1}(s)]}{f_0[F^{-1}(s)]} \right) \int_{1}^{F^{-1}(F_0(y_0) + \tau)} f[F^{-1}(s), y_1] dy_1$$

$$+ \int_{1}^{F^{-1}(s)} \frac{H_0(y_0) - H_1[F^{-1}(F_0(y_0) + \tau)]}{f_1[F^{-1}(F_0(y_0) + \tau)]} f(y_0, F^{-1}(F_0(y_0) + \tau)) dy_0$$

$$+ \int_{1}^{F^{-1}(s)} \int_{1}^{F^{-1}(F_0(y_0) + \tau)} dH(y_0, y_1),$$

where

$$H_0(a) = \lim_{x \to \infty} H(a, x) \text{ and } H_1(a) = \lim_{x \to \infty} H(x, a).$$

\textbf{Proof.} in Appendix \textit{■}

One can now derive the asymptotic distribution of $\hat{v} = v(\hat{F})$ as follows. Let $\sim$ denote standard weak convergence of distribution functions and define $\mathcal{G}$ by $\sqrt{n}(\hat{F} - F) \sim \mathcal{G}$. Then from theorem 1 and the functional delta method, we have that

$$\sqrt{n}(\hat{v} - v_0) \overset{d}{\to} v'_F(\mathcal{G})$$

with

$$v'_F(\mathcal{G}) = \left( \frac{\mathcal{G}_0[F^{-1}(s)]}{f_0[F^{-1}(s)]} \right) \int_{1}^{F^{-1}(F_0(y_0) + \tau)} f[F^{-1}(s), y_1] dy_1$$

$$+ \int_{1}^{F^{-1}(s)} \frac{\mathcal{G}_0(y_0) - \mathcal{G}_1[F^{-1}(F_0(y_0) + \tau)]}{f_1[F^{-1}(F_0(y_0) + \tau)]} f(y_0, F^{-1}(F_0(y_0) + \tau)) dy_0$$

$$+ \int_{1}^{F^{-1}(s)} \int_{1}^{F^{-1}(F_0(y_0) + \tau)} d\mathcal{G}(y_0, y_1),$$

where $\mathcal{G}_0$ and $\mathcal{G}_1$ are stochastic processes defined from $\mathcal{G}$, analogous to (13), e.g., $\mathcal{G}_0(a)$ is a univariate normal with mean zero and variance $F_0(a) \times [1 - F_0(a)]$.

It is well-known that the bootstrap provides consistent approximations to the asymptotic distribution of the sample c.d.f. process $\sqrt{n} (\hat{F} - F)$.$^{11}$ Using the Hadamard

$^{11}$For a textbook treatment, see theorem 3.6.1 part (iii) in van der Vaart and Wellner (1996) and its discussion on page 346 of the same text.
differentiability result of our theorem 1, it follows, via the functional delta method for the bootstrap in probability (c.f., van der Vaart and Wellner (1996), theorem 3.9.11), that bootstrapping will lead to consistent approximation of the distribution of the estimator of $v(\tau, s)$ and hence of $1 - \frac{1}{s}v(\tau, s)$. In the application discussed below, we use the bootstrap to approximate standard errors for the marginal upward mobility by race (table 2) and for mobility by race and parent income (table 4). We also provide a histogram for the bootstrap distribution (figure 3) and summarize some descriptive measures pertaining to the distribution, such as moments, skewness and kurtosis (table 3). Standard tests fail to reject normality of the distribution, as is to be expected, given the Gaussian form of the ingredients of $v'_{p}(G)$.

### 3.4 Conditional upward mobility

Recall from (4) that conditional upward mobility is given by

$$v_c(\tau, s; x) = \frac{\Pr(F_1(Y_1) - F_0(Y_0) > \tau, F_0(Y_0) \leq s | X = x)}{s}.$$  

Then, $v_c(\tau, s; x)$ is estimated by

$$\hat{v}_c(\tau, s; x) = \frac{1}{n\sigma_n^2} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) \frac{1}{s} \left( \frac{\hat{F}_1(Y_{1i}) - \hat{F}_0(Y_{0i}) > \tau, \hat{F}_0(Y_{0i}) \leq s}{s} \right),$$

where $\hat{F}_1(Y_{1i}) = \frac{1}{n-1} \sum_{j \neq i} 1(Y_{ij} \leq Y_{1i})$ and $K(.)$ is a $d$-dimensional kernel function with a bandwidth sequence $\sigma_n$, satisfying the NW conditions specified above. Therefore,

$$\hat{v}_c(\tau, s; x) - v_c(\tau, s; x) = \left\{ \frac{1}{n\sigma_n^2} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) 1(F_1(Y_{1i}) - F_0(Y_{0i}) > \tau, F_0(Y_{0i}) \leq s) - \frac{1}{n\sigma_n^2} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) \right\} s \times \frac{1}{n\sigma_n^2} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) \left[ 1(F_1(Y_{1i}) - \hat{F}_0(Y_{0i}) > \tau, \hat{F}_0(Y_{0i}) \leq s) - 1(F_1(Y_{1i}) - F_0(Y_{0i}) > \tau, \hat{F}_0(Y_{0i}) \leq s) \right]$$

$$\equiv T_{1n} + \frac{T_{2n}}{f(x)}, \text{ say.}$$

We will show that $\frac{T_{2n}}{f(x)}$ has a smaller order of magnitude than $T_{1n}$. This will imply that asymptotically, the distribution of $\hat{v}_c(\tau, s; x)$ will be $\frac{1}{s}$ times that of a Nadaraya-Watson regression function of the unobserved random variable $W \equiv 1(F_1(Y_1) - F_0(Y_0) > \tau, F_0(Y_0) \leq s)$.
on $X$, evaluated at $X = x$. Zero-mean normality of the Nadaraya-Watson regression of $W$ on $X$ will follow from the assumptions NW1-4 plus the following modification of NW5:

**Assumption NW5’.** The functions $f(x)$ and $f(x) \times E(W|X = x)$ and their derivatives up to order 2 are continuous and uniformly bounded.

**Theorem 2** Suppose the data $(X_i, Y_{1i}, Y_{0i})$ for $i = 1, \ldots, n$ are i.i.d. and assumptions NW1-4, NW5’ hold. Then, we have that

\[
\left( n\sigma^4_n \right)^{1/2} \left( \hat{v}_c(\tau, s; x) - v_c(\tau, s; x) \right) \xrightarrow{d} N \left( 0, \frac{\sigma^2(x; \tau, s)}{s^2 f(x)} \int K^2(u) \, du \right),
\]

where $\sigma^2(x; \tau, s) = s \times v_c(\tau, s; x) \times (1 - s \times v_c(\tau, s; x))$ and $f(\cdot)$ is the marginal density of $X$.

**Proof.** Appendix

4 Application

In this section we produce empirical estimates and standard errors of IGM for black and white men using the two measures described in the previous sections: transition probabilities and upward mobility. For each measure we show two sets of results: unconditional estimates and conditional on AFQT test scores. For data, we use a sample of 2766 individuals from the National Longitudinal Survey of Youth (NLSY) who were between the ages of 14 and 22 in 1979. We measure the average family income of these individuals when they were living at home with their parents in 1978, 1979 and 1980. We also measure their average annual earnings as adults in 1996, 1998, 2000 and 2002. Specifically, we use the time average over any years during the relevant time period in which data are available. This allows us to include individuals even if data is missing in some years. The time averaging also provides a better measure of permanent income in both generations. All income variables are deflated to 1978 dollars using the CPI-U.

4.1 Marginal probabilities

4.1.1 Upward transition probabilities

We begin by showing estimates for upward IGM using transition probabilities. These represent the probability that son’s income, $(Y_{1i})$ surpasses a given quantile conditional on
parent income \((Y_0)\) having been at or below the same quantile in the parent generation (i.e. \(s = t\) in (1)). To facilitate comparisons with the upward mobility measure we have introduced in this paper, we also consider transition probabilities where the son must surpass the quantile by the amount \(\tau\), viz.,

\[
\theta_\tau (s) = \frac{\Pr [Y_1 > \zeta_1 + \tau, Y_0 \leq \zeta_0]}{\Pr [Y_0 \leq \zeta_0]}.
\]

Confidence intervals for \(\theta_\tau (s)\) are calculated using analogs of the analytical formulae in FSZ.

The results are shown in Table 1. In the first set of three columns we produce separate estimates for whites, blacks, and the white-black difference for the baseline case where \(\tau = 0\). In the subsequent sets of columns we allow \(\tau\) to vary from 0.1 to 0.3. In each row we condition on parent income being below the \(s\) percentile where \(s\) goes from 5 to 50 in increments of 5. It is immediately evident that the white-black differences are dramatic. For example, the baseline transition probability out of the bottom quartile is 71 percent for whites but only 45 percent for black, or a 26 percentage point difference. We plot the transition probabilities for whites and blacks along with the pointwise 95 percent confidence intervals in Figure 1. The figure plots the results in intervals of 5 percentile points as parent percentile varies from 5 to 50. As is evident in the chart, except for those at the very bottom of the distribution (below the 5th percentile), blacks are significantly less likely to surpass the quantile thresholds.

This is an important finding because most previous research on IGM has used measures such as the intergenerational elasticity, which do not allow for comparisons of group differences in mobility with respect to the entire population. We are only aware of one previous study, Hertz (2005) that has documented differences between blacks and whites in intergenerational transition probabilities.\(^{12}\) However, Hertz relied on PSID data where there is some concern about the representativeness of intergenerational samples to identify black-white differences.\(^{13}\)

---

\(^{12}\)Previous studies such as Corcoran and Adams (1997) and Datcher (1981) have described general patterns in intergenerational mobility between blacks and whites but have not utilized transition probabilities.\(^{13}\)Lee and Solon (2008), for example note that due to technical errors there are concerns about the representativeness of the sampling frame of the oversample of poorer households in the PSID. In addition, there has been significant attrition among black families in the nationally representative portion of the PSID since the sample began in 1968 (Solon, 1992). Attrition rates are considerably smaller in the NLSY
Interestingly, the white-black difference in the transition probability out of the bottom quartile does not change very much as we allow \( \tau \) to vary. For example the racial gap in the probability of rising from the bottom quartile to at least the 45th percentile (i.e. \( \tau = 0.2 \)) is 23 percentage points. When we condition on parents that are at or below the median and allow \( \tau \) to be large (0.2 to 0.3) then we find that the interracial mobility gap begins to narrow to a smaller, but still significant, 10 percentage point difference.

### 4.1.2 Upward mobility

We now show an analogous set of estimates of our measure of upward mobility \( v \) (defined in (11)) for whites and blacks and the white-black difference in Table 2. We now find much smaller racial differences in our baseline case (\( \tau = 0 \)). For example, among white men whose family income during their youth was below the 25th percentile, 84 percent achieved a higher percentile than their parents. The comparable figure for black men is 76 percent implying a difference of about 8 percentage points. The results are plotted in intervals of 5 percentile points along with pointwise 95% confidence bands in Figure 2. To calculate pointwise confidence intervals for mobility \( v \), we compute the sample analog \( \hat{v} \) and then draw 200 bootstrap resamples from our sample. The use of the bootstrap is justified via the functional delta method, discussed just after theorem 1 above.\(^{14}\) For each bootstrap resample, we calculate the corresponding estimate, \( v^* \) and the statistic \( t^* = \sqrt{n} |v^* - \hat{v}| \).

We then calculate \( z^* \), the 95th percentile of \( t^* \). and use \( \left( \hat{v} - \frac{s^*}{\sqrt{n}}, \hat{v} + \frac{s^*}{\sqrt{n}} \right) \) as our confidence interval. We calculate the standard errors shown in the table, \( \sigma_v \), by taking the standard deviation of \( v^* \). The histogram for the bootstrap distribution of \( (v^* - \hat{v}) \) is plotted in figure 3 for the case of whites where \( s = 0.25 \) and \( \tau = 0.2 \). We report the summary statistics (e.g. skewness, kurtosis and tests for normality) for various values of \( s \) and \( \tau \) in Table 3. Because the histograms do not look perfectly symmetric, we also calculated equal-tailed confidence intervals. Since we found no consistent pattern in the relative size of the confidence intervals between the symmetric and the equal-tailed ones, we chose to report the symmetric ones.

---

\( s \) and while there may be some concern about the representativeness of the NLSY (MacCurdy et al., 1998) there is no known problem that puts into question the initial selection of households in the sampling frame.

\( ^{14} \) While studentization may be preferable before bootstrapping for higher order refinements, it is quite challenging to simulate the distribution stated in theorem 1 and so we simply use the unstudentized version.
As figure 2 makes clear, aside from those whose family income was at or below the fifth percentile, whites appear to experience greater upward mobility than blacks but not nearly as much as implied by the difference in the transition probabilities. The gap in most cases, however, is statistically significant as is shown in figure 4 where we plot the white minus black difference for both the transition probability and the upward mobility along with confidence bands.

Clearly, among poorer families there are many blacks who exceed their parents rank in the distribution but do not surpass them by enough to move across specific quantiles. As discussed in section 2, the fact that the white distribution of parent income lies to the right of blacks over most of the support, makes it more likely that whites will surpass the quantile thresholds more easily. This is illustrated in Figure 5 which plots the c.d.f.'s of the parent income distribution for both blacks and whites. This implies that if blacks and whites below the threshold experienced equal sized percentile gains, then the transition probabilities would generally be higher for whites. 15

The remaining columns of Table 2 show the comparable results as \( \tau \) varies from 0.1 to 0.3. In each case, the magnitude of the black white difference is generally between 15 and 25 percentage points and does not change too much as \( s \) changes. These results are comparable to the upward transition probability results in Table 1 and suggest that the two measures produce roughly similar results for larger values of \( \tau \).

Thus far the IGM measures presented have used progressively larger samples that have added more families as \( s \) is increased. This "cumulative" approach could obscure patterns that might arise if we focused more finely on upward economic mobility for individuals coming from specific parts of the income distribution. In addition, the fact that the white distribution lies to the right of the black distribution suggests that blacks may have a built in advantage with respect to upward mobility using cumulative samples since they have more "room" to rise. To address this we recalculated measures by using non-overlapping ranges (\( s_1 \) to \( s_2 \)) for parent income that move progressively up the income distribution. Table 4 which presents these results, demonstrates that much of the rapid upward mobility experienced by blacks is concentrated at the very bottom of the distribution. For example, among those whose parents were between the 21st and 25th percentile, upward mobility is 28 percentage points more rapid for whites than blacks. Overall, these results suggest

15 However, in other results (not shown) we also find that the magnitude of the percentile gains for blacks are actually much lower than for whites.
that by most measures the extent of upward mobility among blacks is vastly lower than for whites.

4.2 Conditional probabilities

The underlying mechanisms by which economic status is transmitted across generations is not yet well understood and is clearly a question of great importance. Estimates of IGM conditional on key covariates can potentially shed light on this question. Understanding the source of the black-white mobility gap in particular, is of great policy interest.

Previous studies using the NLSY have taken advantage of the availability of scores on the Armed Forces Qualifying Test (AFQT) as measure of cognitive skills to identify this as a source of interracial inequality.\footnote{All individuals in the NLSY were given the AFQT test in 1980 as part of the renorming of the test. Following Neal and Johnson (1996) we use the 1989 version of the percentile score. The U.S. military views the AFQT score as "a general measure of trainability and predictor of on-the-job performance". (http://www.defenselink.mil/prhome/poprep2002/chapter2/c2_recruiting.htm)} For example, Neal and Johnson (1996) have shown that the black-white wage gap among adults can largely be explained by pre-market skills as proxied by AFQT scores during adolescence. O’Neill, Sweetman and Van de Gaer (2006) show that equalization of cognitive skill gaps does not fully account for the black-white gap at the low end of the distribution. Cameron and Heckman (2001) have shown that the sizable gap in college enrollment between whites and blacks can largely be accounted for by AFQT scores. These previous findings suggest that it is possible that the average black-white IGM gap might be accounted for by inclusion of AFQT scores but that there might be differences that remain at the low end. However, it is important to note that unlike these previous studies, our measures of mobility capture movements in the distribution relative to the parent generation, so it is not obvious whether mobility gaps will be eliminated the same way that level gaps are. Below, we produce estimates of upward transition probabilities and our measure of upward mobility for black and white men where we now include AFQT scores nonparametrically as a covariate.

We employ Nadaraya-Watson regressions. To do so, we first normalized the regressor to lie between 0 and 1, using maximum and minimum possible values of the AFQT viz.,.99 and 1 and then estimated the regressions at 100 points with spacing of 0.01. We used an Epanechnikov kernel and chose the bandwidth $\sigma_n$ in accordance with assumption NW4 above where $d = 1$. We experimented with bandwidths around the range $n^{-1/4}$.
(moving from \( n^{-1/5} \) to \( n^{-1/3} \)), where \( n \) denotes the size of the effective sample (this size varies depending on which parent percentile and race are conditioned on). Our results for conditional mobility were quite stable over this range and so we report the results at the \( \sigma_n = n^{-1/4} \) value. For inference, we calculated uniform bands using the analytical formulas from Hardle (1990, theorem 4.3.1), which are based on Bickel and Rosenblatt (1973). The latter steps are reproduced in the appendix.

4.2.1 Conditional Transition Probabilities

We estimate the effect of AFQT scores on upward transition probabilities separately by race. Our dependent variable is the probability of leaving the bottom quintile. Figure 6 shows the result of this exercise. We find that conditional on AFQT scores, whites have only slightly higher likelihood of exiting the bottom quintile and that this gap does not vary a great deal across the AFQT distribution. For example at the 25th interval of our normalized AFQT scores, the transition probability for whites is 0.63 and for blacks is 0.61, or a difference of just 2 percentage points. At the 10th interval the gap is about 7 percentage points and at the 75th interval the gap is about 15 percentage points. At no point in the AFQT distribution can we reject the hypothesis that the transition probabilities are the same.

The shape of the regression lines are also similar between blacks and whites for the bottom half of the distribution. In the upper half of the AFQT distribution, however, the slopes differ and the lines fan apart. It is important to note however, that there is relatively little data for blacks in the upper end of the AFQT distribution as is evidenced by the widening confidence intervals.

This finding of similar point estimates of conditional transition probabilities using AFQT scores can be contrasted with results using years of education. In figure 7, we do a similar exercise where we instead use 20 intervals of the sons’ years of completed schooling (normalized) as a covariate in estimating the transition probabilities by race. Here we find sharp differences in the transition probability even conditional on years of schooling throughout much of the distribution. For example among those in the 10th interval, with roughly 10 years of schooling, the transition probability out of the bottom quintile for whites is 67 percent while for blacks it is just 45 percent. At the higher end of the education distribution, however, the racial gap converges and at the very top of the distribution, black mobility is actually higher. Our confidence intervals are quite large,
however, so although the differences are quite big over much of the distribution, they are not statistically significant. In similar exercises using measures of parent education (not shown) we find broadly similar results. Therefore, like Hertz (2005), we find that parent education cannot explain the black-white mobility gap for most of the distribution. However, we find that accounting for AFQT scores does appear to account for the gap.

Finally, we also find that using our nonparametric approach produces some important substantive differences compared to simply estimating a probit with AFQT as a covariate, i.e.,

$$\Pr [Y_1 \leq \zeta_1, Y_0 \leq \zeta_0 | X = x] = \Phi (\beta_0 + \beta_1 \times AFQT).$$

This is particularly true for whites at the bottom of the distribution and for blacks at the top of the distribution. In figure 8 we compare the transition probability results for whites in the bottom of the distribution with the results from simply using a probit. As the chart shows, moving from the first percentile of the AFQT distribution to the median nearly doubles the transition probability of leaving the bottom quintile for whites from 0.43 to 0.85 when using the probit. In contrast, the non-parametric estimator implies an increase of only about 27 percentage points from 0.52 to 0.79. This is not surprising because the probit estimate at a point is affected by the outcome at far-off regressor values unlike the nonparametric estimates.

4.2.2 Conditional Upward Mobility

We also estimate the effect of AFQT scores on our measure of upward mobility separately by race. For this exercise, we condition on parent income being at or below the 20th percentile and set $\tau = 0$. The results are shown in Figure 9. In this case the effects on the black-white gap are even more striking as the point estimates imply that upward mobility is virtually identical for blacks and whites in the bottom half of the distribution.

4.2.3 Discussion of Results

We wish to be careful to point out that we do not think that the finding that AFQT scores can account for the black-white IGM gap lends itself to any simple interpretation or any obvious policy remedy. The development of cognitive skills that we measure in adolescence can be due to a range of factors including health endowments, parental investment, peer influences or school quality. Understanding the formation of the black-white skills gap
has been, and will likely continue to be, an area of intense research activity. Our results suggest that whatever the underlying causes of the gap in cognitive skills it appears to translate into significant differences in IGM.

5 Conclusion

In this paper, we have developed some analytic tools for investigating levels of intergenerational mobility and effects of covariates on it, based on sample data. We have focused on nonparametric regression of transition probabilities and a new direct measure of upward mobility on continuously distributed covariates. Available statistical techniques cannot be used to derive the sampling distribution theory of these estimates because the dependent variables are nonsmooth functions of a set of initially estimated functionals. Therefore, we have developed the relevant asymptotic distribution theory which allows us to investigate the difference in the nature and determinants of intergenerational mobility across population subgroups using survey data.

Applying our techniques to micro data from the National Longitudinal Survey of Youth from the United States, we have demonstrated that most of the black-white difference in ability to rise out of the bottom quintile can be accounted for by differences in cognitive skills during adolescence as measured by the AFQT score.

Although our analytical methods are applied in the context of intergenerational mobility here, they are applicable to any problem involving nonparametric regression where the dependent variable involves nonsmooth functions of preliminary finite-dimensional estimates or estimated ranks.

Our analysis of mobility is essentially descriptive and thus in line with the existing research on the topic. In order to draw causal inference here, one would need to develop an instrumental variable model for the mobility outcome. There are two complications involved. First, one needs a structure for the joint determination of \((Y_0, Y_1, X)\) with the IV’s satisfying appropriate moment or quantile restrictions and secondly, the mobility measures are nonlinear maps from the joint c.d.f. of \((Y_0, Y_1)\) to the unit interval. Such IV-based analysis is outside the scope of the present paper and it would be interesting to explore it in future research.
### Table 1: Transition Probability Estimates by Race

\[ \theta = \text{Prob}(Y_i > \zeta + \tau, Y_o < \zeta_o) / \text{Prob}(Y_o < \zeta_o) \]

<table>
<thead>
<tr>
<th>( \zeta )</th>
<th>Whites ( \tau=0 )</th>
<th>Blacks ( \tau=0 )</th>
<th>W-B ( \tau=0 )</th>
<th>Whites ( \tau=0.1 )</th>
<th>Blacks ( \tau=0.1 )</th>
<th>W-B ( \tau=0.1 )</th>
<th>Whites ( \tau=0.2 )</th>
<th>Blacks ( \tau=0.2 )</th>
<th>W-B ( \tau=0.2 )</th>
<th>Whites ( \tau=0.3 )</th>
<th>Blacks ( \tau=0.3 )</th>
<th>W-B ( \tau=0.3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.978</td>
<td>0.891</td>
<td>0.087</td>
<td>0.849</td>
<td>0.579</td>
<td>0.270</td>
<td>0.704</td>
<td>0.407</td>
<td>0.297</td>
<td>0.593</td>
<td>0.280</td>
<td>0.312</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.025)</td>
<td>(0.041)</td>
<td>(0.057)</td>
<td>(0.043)</td>
<td>(0.073)</td>
<td>(0.070)</td>
<td>(0.044)</td>
<td>(0.084)</td>
<td>(0.084)</td>
<td>(0.040)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>10</td>
<td>0.917</td>
<td>0.702</td>
<td>0.216</td>
<td>0.760</td>
<td>0.458</td>
<td>0.302</td>
<td>0.632</td>
<td>0.340</td>
<td>0.292</td>
<td>0.555</td>
<td>0.249</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.027)</td>
<td>(0.042)</td>
<td>(0.046)</td>
<td>(0.030)</td>
<td>(0.055)</td>
<td>(0.053)</td>
<td>(0.028)</td>
<td>(0.061)</td>
<td>(0.054)</td>
<td>(0.025)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>15</td>
<td>0.812</td>
<td>0.616</td>
<td>0.196</td>
<td>0.692</td>
<td>0.423</td>
<td>0.269</td>
<td>0.542</td>
<td>0.309</td>
<td>0.232</td>
<td>0.459</td>
<td>0.212</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.026)</td>
<td>(0.042)</td>
<td>(0.035)</td>
<td>(0.026)</td>
<td>(0.045)</td>
<td>(0.037)</td>
<td>(0.023)</td>
<td>(0.046)</td>
<td>(0.038)</td>
<td>(0.020)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>20</td>
<td>0.752</td>
<td>0.524</td>
<td>0.228</td>
<td>0.618</td>
<td>0.389</td>
<td>0.229</td>
<td>0.496</td>
<td>0.281</td>
<td>0.215</td>
<td>0.379</td>
<td>0.192</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td>(0.025)</td>
<td>(0.041)</td>
<td>(0.033)</td>
<td>(0.025)</td>
<td>(0.044)</td>
<td>(0.035)</td>
<td>(0.022)</td>
<td>(0.041)</td>
<td>(0.037)</td>
<td>(0.019)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>25</td>
<td>0.708</td>
<td>0.447</td>
<td>0.261</td>
<td>0.558</td>
<td>0.326</td>
<td>0.232</td>
<td>0.459</td>
<td>0.234</td>
<td>0.225</td>
<td>0.342</td>
<td>0.156</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.024)</td>
<td>(0.036)</td>
<td>(0.026)</td>
<td>(0.021)</td>
<td>(0.035)</td>
<td>(0.027)</td>
<td>(0.019)</td>
<td>(0.034)</td>
<td>(0.029)</td>
<td>(0.017)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>30</td>
<td>0.646</td>
<td>0.403</td>
<td>0.244</td>
<td>0.539</td>
<td>0.290</td>
<td>0.249</td>
<td>0.418</td>
<td>0.200</td>
<td>0.217</td>
<td>0.305</td>
<td>0.131</td>
<td>0.174</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.020)</td>
<td>(0.034)</td>
<td>(0.026)</td>
<td>(0.018)</td>
<td>(0.033)</td>
<td>(0.024)</td>
<td>(0.016)</td>
<td>(0.028)</td>
<td>(0.021)</td>
<td>(0.013)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>35</td>
<td>0.583</td>
<td>0.349</td>
<td>0.234</td>
<td>0.478</td>
<td>0.254</td>
<td>0.224</td>
<td>0.366</td>
<td>0.173</td>
<td>0.193</td>
<td>0.257</td>
<td>0.120</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.020)</td>
<td>(0.034)</td>
<td>(0.023)</td>
<td>(0.018)</td>
<td>(0.032)</td>
<td>(0.022)</td>
<td>(0.015)</td>
<td>(0.028)</td>
<td>(0.020)</td>
<td>(0.013)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>40</td>
<td>0.544</td>
<td>0.311</td>
<td>0.233</td>
<td>0.427</td>
<td>0.223</td>
<td>0.203</td>
<td>0.315</td>
<td>0.148</td>
<td>0.167</td>
<td>0.228</td>
<td>0.105</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.018)</td>
<td>(0.029)</td>
<td>(0.019)</td>
<td>(0.017)</td>
<td>(0.027)</td>
<td>(0.018)</td>
<td>(0.013)</td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.011)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>45</td>
<td>0.494</td>
<td>0.262</td>
<td>0.232</td>
<td>0.372</td>
<td>0.180</td>
<td>0.192</td>
<td>0.264</td>
<td>0.123</td>
<td>0.141</td>
<td>0.190</td>
<td>0.080</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.015)</td>
<td>(0.025)</td>
<td>(0.016)</td>
<td>(0.012)</td>
<td>(0.020)</td>
<td>(0.014)</td>
<td>(0.011)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>50</td>
<td>0.428</td>
<td>0.226</td>
<td>0.202</td>
<td>0.320</td>
<td>0.152</td>
<td>0.168</td>
<td>0.227</td>
<td>0.107</td>
<td>0.119</td>
<td>0.147</td>
<td>0.065</td>
<td>0.082</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.015)</td>
<td>(0.023)</td>
<td>(0.016)</td>
<td>(0.012)</td>
<td>(0.021)</td>
<td>(0.015)</td>
<td>(0.011)</td>
<td>(0.020)</td>
<td>(0.011)</td>
<td>(0.009)</td>
<td>(0.015)</td>
</tr>
</tbody>
</table>

Notes: See text for a description of the estimator. Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. Standard errors are in parentheses.
Table 2: Upward Mobility Estimates by Race

\[ \nu = \text{Prob}(F(Y_1) - F(Y_0) > \tau | F(Y_0) \leq s) \]

<table>
<thead>
<tr>
<th>s</th>
<th>Whites</th>
<th>Blacks</th>
<th>W-B</th>
<th>Whites</th>
<th>Blacks</th>
<th>W-B</th>
<th>Whites</th>
<th>Blacks</th>
<th>W-B</th>
<th>Whites</th>
<th>Blacks</th>
<th>W-B</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.977</td>
<td>0.950</td>
<td>0.027</td>
<td>0.904</td>
<td>0.635</td>
<td>0.270</td>
<td>0.745</td>
<td>0.420</td>
<td>0.325</td>
<td>0.614</td>
<td>0.312</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.033)</td>
<td>(0.047)</td>
<td>(0.044)</td>
<td>(0.066)</td>
<td>(0.065)</td>
<td>(0.045)</td>
<td>(0.083)</td>
<td>(0.073)</td>
<td>(0.040)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>10</td>
<td>0.947</td>
<td>0.883</td>
<td>0.065</td>
<td>0.840</td>
<td>0.574</td>
<td>0.266</td>
<td>0.698</td>
<td>0.377</td>
<td>0.321</td>
<td>0.595</td>
<td>0.288</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.022)</td>
<td>(0.032)</td>
<td>(0.035)</td>
<td>(0.032)</td>
<td>(0.051)</td>
<td>(0.047)</td>
<td>(0.031)</td>
<td>(0.059)</td>
<td>(0.053)</td>
<td>(0.025)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>15</td>
<td>0.909</td>
<td>0.835</td>
<td>0.074</td>
<td>0.786</td>
<td>0.567</td>
<td>0.219</td>
<td>0.629</td>
<td>0.390</td>
<td>0.240</td>
<td>0.519</td>
<td>0.281</td>
<td>0.238</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.020)</td>
<td>(0.029)</td>
<td>(0.031)</td>
<td>(0.027)</td>
<td>(0.042)</td>
<td>(0.040)</td>
<td>(0.025)</td>
<td>(0.047)</td>
<td>(0.040)</td>
<td>(0.025)</td>
<td>(0.048)</td>
</tr>
<tr>
<td>20</td>
<td>0.871</td>
<td>0.796</td>
<td>0.075</td>
<td>0.755</td>
<td>0.556</td>
<td>0.198</td>
<td>0.592</td>
<td>0.387</td>
<td>0.205</td>
<td>0.485</td>
<td>0.285</td>
<td>0.200</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.017)</td>
<td>(0.027)</td>
<td>(0.029)</td>
<td>(0.024)</td>
<td>(0.039)</td>
<td>(0.030)</td>
<td>(0.022)</td>
<td>(0.037)</td>
<td>(0.032)</td>
<td>(0.020)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>25</td>
<td>0.838</td>
<td>0.762</td>
<td>0.076</td>
<td>0.724</td>
<td>0.537</td>
<td>0.187</td>
<td>0.575</td>
<td>0.373</td>
<td>0.202</td>
<td>0.463</td>
<td>0.274</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.019)</td>
<td>(0.030)</td>
<td>(0.024)</td>
<td>(0.024)</td>
<td>(0.038)</td>
<td>(0.028)</td>
<td>(0.024)</td>
<td>(0.036)</td>
<td>(0.028)</td>
<td>(0.019)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>30</td>
<td>0.821</td>
<td>0.734</td>
<td>0.087</td>
<td>0.715</td>
<td>0.521</td>
<td>0.193</td>
<td>0.568</td>
<td>0.360</td>
<td>0.208</td>
<td>0.447</td>
<td>0.262</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.019)</td>
<td>(0.027)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.033)</td>
<td>(0.026)</td>
<td>(0.020)</td>
<td>(0.036)</td>
<td>(0.025)</td>
<td>(0.019)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>35</td>
<td>0.786</td>
<td>0.717</td>
<td>0.069</td>
<td>0.668</td>
<td>0.514</td>
<td>0.154</td>
<td>0.537</td>
<td>0.360</td>
<td>0.178</td>
<td>0.415</td>
<td>0.263</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.017)</td>
<td>(0.026)</td>
<td>(0.020)</td>
<td>(0.023)</td>
<td>(0.030)</td>
<td>(0.021)</td>
<td>(0.019)</td>
<td>(0.031)</td>
<td>(0.023)</td>
<td>(0.016)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>40</td>
<td>0.757</td>
<td>0.704</td>
<td>0.052</td>
<td>0.641</td>
<td>0.506</td>
<td>0.135</td>
<td>0.513</td>
<td>0.357</td>
<td>0.156</td>
<td>0.393</td>
<td>0.254</td>
<td>0.139</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.016)</td>
<td>(0.025)</td>
<td>(0.017)</td>
<td>(0.020)</td>
<td>(0.028)</td>
<td>(0.020)</td>
<td>(0.019)</td>
<td>(0.027)</td>
<td>(0.019)</td>
<td>(0.018)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>45</td>
<td>0.731</td>
<td>0.687</td>
<td>0.044</td>
<td>0.605</td>
<td>0.495</td>
<td>0.110</td>
<td>0.484</td>
<td>0.350</td>
<td>0.134</td>
<td>0.367</td>
<td>0.248</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.017)</td>
<td>(0.024)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td>(0.032)</td>
<td>(0.019)</td>
<td>(0.018)</td>
<td>(0.028)</td>
<td>(0.019)</td>
<td>(0.017)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>50</td>
<td>0.695</td>
<td>0.668</td>
<td>0.028</td>
<td>0.578</td>
<td>0.481</td>
<td>0.097</td>
<td>0.457</td>
<td>0.342</td>
<td>0.115</td>
<td>0.342</td>
<td>0.242</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.018)</td>
<td>(0.025)</td>
<td>(0.016)</td>
<td>(0.020)</td>
<td>(0.028)</td>
<td>(0.016)</td>
<td>(0.018)</td>
<td>(0.025)</td>
<td>(0.017)</td>
<td>(0.015)</td>
<td>(0.024)</td>
</tr>
</tbody>
</table>

Notes: See text for a description of the estimator. Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. Bootstrapped standard errors are in parentheses.
Table 3: Summary statistics of bootstrapped values of $\nu^* - \nu$

<table>
<thead>
<tr>
<th>For $s = 0.25$</th>
<th>mean</th>
<th>median</th>
<th>skewness</th>
<th>kurtosis</th>
<th>p-value skew test</th>
<th>p-value kurt. test</th>
<th>p-value joint (chi sq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 0$</td>
<td>-0.009</td>
<td>-0.010</td>
<td>0.100</td>
<td>2.787</td>
<td>0.550</td>
<td>0.637</td>
<td>0.746</td>
</tr>
<tr>
<td>$\tau = 0.1$</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.091</td>
<td>2.700</td>
<td>0.587</td>
<td>0.417</td>
<td>0.618</td>
</tr>
<tr>
<td>$\tau = 0.2$</td>
<td>0.001</td>
<td>0.001</td>
<td>0.118</td>
<td>2.916</td>
<td>0.483</td>
<td>0.982</td>
<td>0.780</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For $s = 0.5$</th>
<th>mean</th>
<th>median</th>
<th>skewness</th>
<th>kurtosis</th>
<th>p-value skew test</th>
<th>p-value kurt. test</th>
<th>p-value joint (chi sq)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau = 0$</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.283</td>
<td>3.301</td>
<td>0.097</td>
<td>0.297</td>
<td>0.143</td>
</tr>
<tr>
<td>$\tau = 0.1$</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.066</td>
<td>2.638</td>
<td>0.695</td>
<td>0.285</td>
<td>0.519</td>
</tr>
<tr>
<td>$\tau = 0.2$</td>
<td>0.001</td>
<td>0.001</td>
<td>-0.017</td>
<td>2.543</td>
<td>0.918</td>
<td>0.131</td>
<td>0.314</td>
</tr>
</tbody>
</table>

Notes: In all cases $N = 200$. $p$-values are from using "sktest" command in STATA v10.1
Table 4: Upward Mobility Estimates by Race Using Intervals of Parent Income

\[
v = \text{Prob}(F_1(Y_i) - F_0(Y_0) > \tau | F_0(Y_0) > s_1 \ & \ F_0(Y_0) \leq s_2)
\]

<table>
<thead>
<tr>
<th>(s_1 \text{ to } s_2)</th>
<th>(\tau=0)</th>
<th>(\tau=0.1)</th>
<th>(\tau=0.2)</th>
<th>(\tau=0.3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Whites</td>
<td>Blacks</td>
<td>W-B</td>
<td>Whites</td>
</tr>
<tr>
<td>1 to 5</td>
<td>0.977</td>
<td>0.950</td>
<td>0.027</td>
<td>0.904</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.033)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>6 to 10</td>
<td>0.915</td>
<td>0.813</td>
<td>0.102</td>
<td>0.770</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.035)</td>
<td>(0.059)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>11 to 15</td>
<td>0.847</td>
<td>0.708</td>
<td>0.138</td>
<td>0.698</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.051)</td>
<td>(0.070)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>16 to 20</td>
<td>0.780</td>
<td>0.645</td>
<td>0.134</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.048)</td>
<td>(0.079)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>21 to 25</td>
<td>0.751</td>
<td>0.473</td>
<td>0.278</td>
<td>0.645</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.070)</td>
<td>(0.092)</td>
<td>(0.058)</td>
</tr>
<tr>
<td>26 to 30</td>
<td>0.755</td>
<td>0.534</td>
<td>0.221</td>
<td>0.677</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.059)</td>
<td>(0.077)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>31 to 35</td>
<td>0.639</td>
<td>0.495</td>
<td>0.144</td>
<td>0.471</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.073)</td>
<td>(0.104)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>36 to 40</td>
<td>0.613</td>
<td>0.489</td>
<td>0.124</td>
<td>0.510</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.092)</td>
<td>(0.113)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>41 to 45</td>
<td>0.578</td>
<td>0.258</td>
<td>0.320</td>
<td>0.385</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.096)</td>
<td>(0.116)</td>
<td>(0.071)</td>
</tr>
<tr>
<td>46 to 50</td>
<td>0.450</td>
<td>0.311</td>
<td>0.138</td>
<td>0.393</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.080)</td>
<td>(0.094)</td>
<td>(0.064)</td>
</tr>
</tbody>
</table>

Notes: See text for a description of the estimator. Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. Bootstrapped standard errors are in parentheses.
Figure 1: Transition Probabilities Conditional On Parent Percentile

Figure 2: Upward Mobility Conditional On Parent Percentile

Pr (F1 > F0 | F0 ≤ s)

Figure 3: Histogram of v^* - v

Notes for figures 1 to 3: See text for descriptions of the estimators. Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. For upward mobility estimates, bootstrapped 95% pointwise confidence intervals are shown as bands.
Notes for figures 4 to 6: Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. Uniform confidence intervals shown as dashed lines. For details of computation see text.
Notes for figures 7 to 9: Data are from the NLSY. We use multiyear averages of son's income over 1996-2002 and parent income measured over 1978-1980. For figures 7 and 9, uniform confidence intervals shown as dashed lines. For details of computation see text.
6 Appendix with Proofs

Proof outline for proposition 1:

Proof. Consider the expression

\[
\tilde{m} (\zeta, x) = \frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{X_i - x}{\sigma_n} \right) 1 \left( Y_{1i} \leq \zeta_1, Y_{0i} \leq \zeta_0 \right)
\]

whose expectation is given by

\[
\tilde{m}^* (\zeta, x) = \mathbb{E}_X \left[ \frac{1}{\sigma_n^d} K \left( \frac{X_i - x}{\sigma_n} \right) \phi (X_i, \zeta_0, \zeta_1) \right] \\
= \int K (u) f (x + u\sigma_n) \phi (x + u\sigma_n, \zeta_0, \zeta_1) \, du \\
= f (x) \phi (x, \zeta_0, \zeta_1) + O \left( \sigma_n^2 \right). \tag{15}
\]

So

\[
\tilde{m}^* (\tilde{\zeta}, x) = f (x) \phi (x, \tilde{\zeta}_0, \tilde{\zeta}_1) + O \left( \sigma_n^2 \right) \\
= f (x) \left[ \phi (x, \zeta_0^0, \zeta_1^0) + \phi_1 (x, \tilde{\zeta}_0, \tilde{\zeta}_1) \left( \tilde{\zeta}_0 - \zeta_0 \right) + \phi_1 (x, \tilde{\zeta}_0, \tilde{\zeta}_1) \left( \tilde{\zeta}_1 - \zeta_1 \right) \right] + O \left( \sigma_n^2 \right),
\]

where \(\tilde{\zeta}_1\) denote value intermediate between \(\hat{\zeta}_1\) and \(\zeta_1^0\) and similarly, \(\tilde{\zeta}_0\). Now,

\[
\hat{\phi} (x, \tilde{\zeta}_0, \tilde{\zeta}_1) - \phi (x, \zeta_0^0, \zeta_1^0) \\
= \frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{X_i - x}{\sigma_n} \right) 1 \left( Y_{1i} \leq \tilde{\zeta}_1, Y_{0i} \leq \tilde{\zeta}_0 \right) - \phi (x, \zeta_0^0, \zeta_1^0) \\
= \frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{X_i - x}{\sigma_n} \right) \left\{ 1 \left( Y_{1i} \leq \tilde{\zeta}_1, Y_{0i} \leq \tilde{\zeta}_0 \right) - 1 \left( Y_{1i} \leq \zeta_1, Y_{0i} \leq \zeta_0 \right) \right\} \\
\leq \underbrace{\frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{X_i - x}{\sigma_n} \right) 1 \left( Y_{1i} \leq \zeta_1, Y_{0i} \leq \zeta_0 \right)}_{\equiv \tilde{S}_n, \text{say}} - \phi (x, \zeta_0^0, \zeta_1^0) \\
\leq \left\{ \tilde{m} (\tilde{\zeta}, x) - \tilde{m}^* (\tilde{\zeta}, x) \right\} + \left\{ \tilde{m} (\zeta_0^0, x) - \tilde{m}^* (\zeta_0^0, x) \right\} \\
= \left\{ \tilde{m} (\tilde{\zeta}, x) - \tilde{m}^* (\tilde{\zeta}, x) \right\} + S_n
\]

\[
\int \tilde{f} (x) \left[ \tilde{m} (\tilde{\zeta}, x) - \tilde{m}^* (\tilde{\zeta}, x) \right] \\
+ \int \tilde{f} (x) \left[ \tilde{m}^* (\zeta_0^0, x) - \tilde{m}^* (\zeta_0^0, x) \right] \\
+ \int \tilde{f} (x) \left[ \tilde{m}^* (\tilde{\zeta}, x) - \tilde{m}^* (\zeta_0^0, x) \right] + S_n
\]

\[
\leq \underbrace{\tilde{T}_{1n}}_{\equiv \tilde{S}_n, \text{say}} + \underbrace{\tilde{T}_{2n}}_{\equiv \tilde{S}_n} + \underbrace{\tilde{T}_{3n}}_{\equiv \tilde{S}_n} \tag{16}
\]
Now \((n\sigma_n^d)^{1/2} S_n\), under the assumptions NW, will be \(O_p(1)\) and zero-mean normal (c.f., Bierens (1994), theorem 10.2.1), viz.,

\[
(n\sigma_n^d)^{1/2} S_n \overset{d}{\to} N \left(0, \frac{v^2(x)}{\hat{f}(x)} \int K^2(u) \, du \right). \tag{17}
\]

Next, using (15) and the fact that \(\hat{\phi}_0 (0) - \phi_0 (0)\) and \(\hat{\phi}_1 (1) - \phi_1 (1)\) have parametric convergence rates, we get that

\[
(n\sigma_n^d)^{1/2} T_{2n} = \left[ \phi_0 (x; \tilde{\zeta}_0, \tilde{\zeta}_1) \times (n\sigma_n^d)^{1/2} \times (\hat{\zeta}_0 - \zeta_0) \right] \overset{O_p(1)}{\to} \phi_0 (x; \tilde{\zeta}_0, \tilde{\zeta}_1) \times (n\sigma_n^d)^{1/2} \times (\hat{\zeta}_1 - \zeta_1)
\]

\[
+ O_p \left( \frac{(n\sigma_n^d)^{1/2} \times \sigma_n^2}{o_p(1)} \right), \tag{18}
\]

The nonstandard term in (16) is \(T_{1n}\) and we now demonstrate a stochastic equicontinuity property for it. Letting

\[
v_n (\zeta, x) = (n\sigma_n^d)^{1/2} \left\{ \bar{m} (\zeta, x) - \bar{m}^* (\zeta, x) \right\}
\]

the first term satisfies

\[
(n\sigma_n^d)^{1/2} T_{1n} = \frac{v_n (\hat{\zeta}, x) - v_n (\zeta_0, x)}{\hat{f}(x)}, \tag{19}
\]

and we now show that the numerator of (19) is \(o_p(1)\), for each \(x\). Now, for fixed \(x\), the class of functions

\[
g (W_n, \zeta) \equiv \frac{1}{\sigma_n^d} K \left( \frac{X - x}{\sigma_n} \right) \mathbf{1} (Y_1 \leq \zeta_1, Y_0 \leq \zeta_0)
\]

form a type IV class (c.f. Andrews (1994), equation 5.3) with \(p = 2\). This follows from the Lipschitz property (8) and the uniform boundedness of \(\frac{1}{\sigma_n^d} K \left( \frac{X - x}{\sigma_n} \right)\). This, in turn, implies that the sequence \(v_n (\zeta, x)\) is stochastically equicontinuous. Now, using the same steps as Andrews (1994) leading to his equation (3.8), we conclude that

\[
v_n (\hat{\zeta}, x) - v_n (\zeta_0, x) = o_p(1).
\]
Put together (16), (17), (18) and (19) together with \( \hat{f}(x) = O_p(1) \) to conclude.

**Proof of theorem 1:**

**Proof.** Consider perturbations \( F_t(y_0, y_1) = F(y_0, y_1) + tH_t(y_0, y_1) \) with \( F_{0t}(y_0) = F_0(y_0) + tH_{0t}(y_0) \) and \( F_{1t}(y_0) = F_1(y_1) + tH_{1t}(y_1) \) denoting the corresponding marginals. Let \( H_t \rightarrow H \) uniformly as \( t \to 0 \) and let \( H_0 \) and \( H_1 \) denote its marginals. We want to show that for a linear functional \( v'_F(.) \),

\[
\left| \frac{v(F_t) - v(F)}{t} - v'_F(H) \right| \to 0 \text{ as } t \to 0.
\]  

(20)

Define

\[
z_1(y_0) = F_1^{-1}(F_0(y_0) + \tau), \quad z_{1t}(y_0) = F_{1t}^{-1}(F_{0t}(y_0) + \tau), \\
z_0 = F_0^{-1}(s), \quad z_{0t} = F_{0t}^{-1}(s).
\]

So we need to show

\[
\left| \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} f_t(y_0, y_1) - \int_1^{z_0} \int_1^{z_1(y_0)} f(y_0, y_1) dy_1 dy_0 \right| - v'_F(H) \to 0 \text{ as } t \to 0.
\]

Note that the first term inside \(|.|\) can be expanded as

\[
= \left[ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} \frac{z_{1t}(y_0) - z_1(y_0)}{t} f(z_{1t}(y_0), y_0) dy_0 + \frac{z_{0t} - z_0}{t} \times \int_1^{z_1(y_0)} f(z_{0t}, y_1) dy_1 \\
+ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} dH_t(y_0, y_1) \right]_{T_{3t}} + \left[ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} \frac{z_{1t}(y_0) - z_1(y_0)}{t} f(z_{1t}(y_0), y_0) dy_0 \\
+ \frac{z_{0t} - z_0}{t} \times \int_1^{z_1(y_0)} f(z_{0t}, y_1) dy_1 \right]_{T_{3t}}
\]

\[
= \left[ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} \frac{z_{1t}(y_0) - z_1(y_0)}{t} f(z_0, y_1) dy_1 + \frac{z_{0t} - z_0}{t} \times \int_1^{z_1(y_0)} f(z_0, y_1) dy_1 \\
+ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} dH_t(y_0, y_1) - \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} dH(y_0, y_1) \right]_{T_{3t}}
\]

\[
+ \left[ \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} dH(y_0, y_1) - \int_1^{z_{0t}} \int_1^{z_{1t}(y_0)} dH(y_0, y_1) \right]_{T_{5t}}
\]
We will show that as \( t \to 0 \),

Step 1: \( |T_{1t}| \to 0 \)

Step 2:

\[
T_{2t} \to \frac{H_0 (z_0)}{f_0 (z_0)} \int_1^{z_1 (y_0)} f (z_0, y_1) \, dy_1
\]

Step 3:

\[
T_{3t} \to \int_1^{z_0} \left( \frac{H_0 (y_0) - H_1 (z_1 (y_0))}{f_1 (z_1 (y_0))} \right) f (y_0, z_1 (y_0)) \, dy_0
\]

Step 4: \( |T_{4t}| \to 0 \), \( |\bar{T}_{5t}| \to 0 \).

Then we will have shown (20) with

\[
u'_{F} (H) = \frac{H_0 (F_0^{-1} (s))}{f_0 (F_0^{-1} (s))} \int_1^{F_0^{-1} (F_0 (y_0) + \tau)} f (F_0^{-1} (s), y_1) \, dy_1
\]

\[
+ \int_1^{F_0^{-1} (s)} \frac{H_0 (y_0) - H_1 (F_0^{-1} (F_0 (y_0) + \tau))}{f_1 (F_0^{-1} (F_0 (y_0) + \tau))} \right) f (y_0, F_0^{-1} (F_0 (y_0) + \tau)) \, dy_0
\]

\[
+ \int_1^{F_0^{-1} (s)} \int_1^{F_0^{-1} (F_0 (y_0) + \tau)} dH (y_0, y_1)
\]

which is linear in \( H \).

For steps 1 and 2, we will need the following derivation.

\[
F_1 (z_{1t} (y_0)) + tH_{1t} (z_{1t} (y_0)) = F_{1t} (z_{1t} (y_0)) = F_{0t} (y_0) + \tau
\]

\[
= F_0 (y_0) + \tau + tH_{0t} (y_0) = F_1 (z_1 (y_0)) + tH_{0t} (y_0)
\]

implying that

\[
tH_{1t} (z_{1t} (y_0)) - tH_{0t} (y_0) = F_1 (z_1 (y_0)) - F_1 (z_{1t} (y_0))
\]

\[
= [z_1 (y_0) - z_{1t} (y_0)] f_1 (z_{1t} (y_0))
\]

where for any \( y_0 \) and \( t \), \( z_{1t} (y_0) \) lies in between \( z_1 (y_0) \) and \( z_t (y_0) \). Therefore,

\[
\frac{z_{1t} (y_0) - z_1 (y_0)}{t} = \frac{H_0 (y_0) - H_{1t} (z_{1t} (y_0))}{f_1 (z_{1t} (y_0))}.
\]

(21)

Similarly, \( F_0 (z_0) = s = F_{0t} (z_{0t}) = F_0 (z_{0t}) + tH_{0t} (z_{0t}) \), whence

\[
\frac{z_{0t} - z_0}{t} = \frac{H_{0t} (z_{0t})}{f_0 (\tilde{z}_{0t})}.
\]

(22)
Below, \( c \) will denote a generic constant, not always of the same value.

**Step 1:** By a mean-value theorem argument,

\[
|T_{10t}| \leq \int_1^{2ot} \left| \frac{z_{1t}(y_0) - z_1(y_0)}{t} [f(y_0, \bar{z}_{1t}(y_0)) - f(y_0, z_1(y_0))] \right| dy_0 \leq \int_1^{2ot} \left| \frac{z_1(y_0) - z(y_0)}{t} f^{(1)}(y_0, \bar{z}_{1t}(y_0)) \right| dy_0
\]

where \( f^{(1)}(\cdot, \cdot) \) denotes derivative w.r.t. the second argument and \( \bar{z}_{1t}(y_0) \) lies in between \( z(y_0) \) and \( z_1(y_0) \). Using (21), we get

\[
|T_{10t}| \leq t \int_1^{\infty} \left| \frac{H_{0t}(y_0) - H_{1t}(\bar{z}_{1t}(y_0))}{f^{(1)}(y_0, \bar{z}_{1t}(y_0))} \right| dy_0.
\]

We will show that (i) \( [H_{0t}(y_0) - H_{1t}(\bar{z}_{1t}(y_0))]^2 \) is uniformly bounded, (ii) \( f^{(1)}(\bar{z}_{1t}(y_0)) \geq \frac{c}{\bar{z}_{1t}(y_0)^{\alpha}} \) for \( y_0 \) large enough and \( t \) small enough and (iii) \( f^{(1)}(y_0, \bar{z}_{1t}(y_0)) \leq \frac{c}{y_0} \) for some \( \alpha > 1 \). Then we will have

\[
|T_{10t}| \leq ct \int_1^{\infty} \frac{1}{y_0^{\alpha} (1 - F_0(y_0))^{2\alpha}} dy_0 \leq ct \int_1^{\infty} \frac{1}{y_0^{1+\varepsilon}} dy_0 \to 0,
\]

by A(iii).

To see (i), note that \( \{[H_{0t}(y_0) - H_{1t}(\bar{z}_{1t}(y_0))] - [H_{0}(y_0) - H_1(\bar{z}_{1t}(y_0))]\} \) converges uniformly to 0 and \( H_0(.) \) and \( H_0(.) \) are uniformly bounded.

Next,

\[
\begin{align*}
\bar{z}_{1t}(y_0) &= F_{1t}^{-1} \left( F_{0t}(y_0) + \tau \right) \leq c \left( 1 - F_{0t}(y_0) - \tau \right)^{\frac{1}{1-\alpha}} = c \left( 1 - F_0(y_0) - tH_{0t}(y_0) - \tau \right)^{\frac{1}{1-\alpha}} \\
&\leq c^2 \left( 1 - F_0(y_0) - \tau \right)^{\frac{1}{1-\alpha}} \\
&\leq c \left( 1 - F_0(y_0) \right)^{\frac{1}{1-\alpha}} \quad \text{(23)}
\end{align*}
\]

for small enough \( t \), since \( \alpha > 1 \) and \( tH_{0t}(\cdot) \) converges uniformly to 0. Inequality (1) comes from condition Ai. Similarly,

\[
\bar{z}_1(y_0) \leq c \left( 1 - F_0(y_0) \right)^{\frac{1}{1-\alpha}} \quad \text{(24)}
\]

and therefore (ii) follows. Finally (iii) follows from (23), (24) and condition Aii.
Next, for $|T_{11t}|$, we have that

$$|T_{11t}| \leq \left| \frac{z_{0t} - z_0}{t} \times \int_1^{z_1(y_0)} [f(z_{0t}, y_1) - f(z_0, y_1)] dy_1 \right|$$

$$\leq \left| \frac{[z_{0t} - z_0]^2}{t} \times \int_1^{z_1(y_0)} f^{(0)}(z_{0t}, y_1) dy_1 \right|$$

$$\leq t \left| \frac{H_{0t}(z_{0t})}{f_0(z_{0t})} \right|^2 \times \int_1^{z_1(y_0)} f^{(0)}(z_{0t}, y_1) dy_1 \right| \leq c t \int_1^{z_1(y_0)} \frac{1}{y_0^{1+\delta}} dy_1 \leq c t \int_1^{\infty} \frac{1}{y_0^{1+\delta}} dy_1$$

for $t$ small enough and some $\delta > 0$. Inequality (2) follows from conditions (Bi), (Bii), (Biii) using arguments analogous to those for $T_{10t}$. This implies that $|T_{11t}| \rightarrow 0$.

**Step 2:**

$$\left| \frac{H_{0t}(z_0)}{f_0(z_{0t})} \int_1^{z_1(y_0)} f(z_0, y_1) dy_1 \right|$$

$$\leq \left| \int_1^{z_{0t}} \frac{z_{1t}(y_0) - z_1(y_0)}{t} f(z_0, y_1) dy_0 - \frac{H_{0t}(z_0)}{f_0(z_{0t})} \int_1^{z_1(y_0)} f(z_0, y_1) dy_1 \right|$$

$$\leq \int_1^{z_{0t}} \left| \frac{H_{0t}(z_{0t})}{f_0(z_{0t})} - \frac{H_{0t}(z_0)}{f_0(z_0)} \right| f(z_0, y_1) dy_1$$

$$\leq c \int_1^{z_{0t}} |H_{0t}(z_{0t}) - H_0(z_0)| (1 - F_1(y_1))^{\beta} f(z_0, y_1) dy_1$$

$$\leq c \left( \sup_u |H_{0t}(u) - H_0(u)| + |H_0(z_{0t}) - H_0(z_0)| \right) \int_1^{\infty} (1 - F_1(y_1))^{\beta} f(z_0, y_1) dy_1$$

$$\rightarrow 0$$

as $t \rightarrow 0$, by B(iv). Inequality (1) is a consequence of B(i)-B(iii).
Step 3:

\[
\left| T_{3t} - \int_1^{z_{1t}} \frac{H_0(y_0) - H_1(z_1(y_0))}{f_1(z_1(y_0))} f(y_0, z_1(y_0)) \, dy_0 \right|
\]

\[
= \left| \int_1^{z_{1t}} \frac{z_{1t}(y_0) - z_1(y_0)}{t} f(z_1(y_0), y_0) \, dy_0 - \int_1^{z_{1t}} \frac{H_0(y_0) - H_1(z_1(y_0))}{f_1(z_1(y_0))} f(y_0, z_1(y_0)) \, dy_0 \right|
\]

\[
\leq \int_1^{z_{1t}} \left| \frac{z_{1t}(y_0) - z_1(y_0)}{t} - \frac{H_0(y_0) - H_1(z_1(y_0))}{f_1(z_1(y_0))} \right| f(y_0, z_1(y_0)) \, dy_0
\]

\[
= \int_1^{z_{1t}} \left| \frac{H_0(y_0) - H_1(z_1(y_0))}{f_1(z_1(y_0))} \right| f(y_0, z_1(y_0)) \, dy_0
\]

\[
\leq c \int_1^{z_{1t}} \left\{ \sup_{y_0} \left| H_0(y_0) - H_1(z_1(y_0)) \right| \times (1 - F_0(y_0))^{\alpha-1} \right\} \, dy_0
\]

which goes to zero if \( \int_1^{z_{1t}} (1 - F_0(y_0))^{\alpha-1} f(z_1(y_0), y_0) \, dy_0 < \infty \), which is condition (Aiv).

Note that the inequality \( \leq \) follows from step (ii) in the proof of Step 1, above. Finally, since \( \int_1^{z_{1t}} \frac{H_0(y_0) - H_1(z_1(y_0))}{f_1(z_1(y_0))} f(y_0, z_1(y_0)) \, dy_0 \) is continuous in \( z_{1t} \), the conclusion follows.

Step 4:

\( T_{4t} \to 0 \) since \( H_t \to H \) uniformly and \( T_{5t} \) goes to zero by the continuous mapping theorem since paths of an \( F \)-Brownian bridge are everywhere continuous with probability 1.

Proof of theorem 2:

Proof. Recall display (14). First, under assumptions NW1-4 and NW5’, apply e.g., Bierens (1994), theorem 10.2.1, to conclude that

\[
(n\sigma_n^d)^{1/2} \left\{ \frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) W_i - s \cdot \frac{1}{n\sigma_n^d} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) - v_c(\tau, s; x) \right\} \overset{d}{\to} N \left( 0, \frac{\sigma^2(x; \tau, s)}{s^2 f(x)} \int K^2(u) \, du \right),
\]

where

\[
\sigma^2(x; \tau, s) = s \times v_c(\tau, s; x) \times (1 - s \times v_c(\tau, s; x)).
\]

We will now show that \( E \left( \sqrt{n\sigma_n^d} |T_{2n}| \right) \to 0 \) which will imply that \( |T_{2n}| = o_p \left( (n\sigma_n^d)^{-1/2} \right) \) and thus establish the theorem.
First observe that

$$
\mathbb{E} \left[ \begin{array}{c}
1 & \left( \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \right) \\
-1 & \left( F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s \right)
\end{array} \right] \\
= \Pr \left[ \begin{array}{c}
1 & \left( \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \right) \\
-1 & \left( F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s \right)
\end{array} \right] 
\neq 0

= \Pr \left[ \begin{array}{c}
\{ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \} \\
\cap (F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s)^c
\end{array} \right]

+ \Pr \left[ \begin{array}{c}
\{ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \} \\
\cap (F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s)
\end{array} \right]

\leq \Pr \left[ \left\{ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \right\} \cap (F_1 (Y_{1i}) - F_0 (Y_{0i}) \leq \tau) \right]

+ \Pr \left[ \left\{ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, \hat{F}_0 (Y_{0i}) \leq s \right\} \cap (F_0 (Y_{0i}) > s) \right]

+ \Pr \left[ \left\{ F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s \right\} \cap (\hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) \leq \tau) \right]

+ \Pr \left[ \left\{ F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau, F_0 (Y_{0i}) \leq s \right\} \cap (\hat{F}_0 (Y_{0i}) > s) \right]

\leq \Pr \left[ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) > \tau, F_1 (Y_{1i}) - F_0 (Y_{0i}) \leq \tau \right]

+ \Pr \left\{ \hat{F}_1 (Y_{1i}) - \hat{F}_0 (Y_{0i}) \leq \tau, F_1 (Y_{1i}) - F_0 (Y_{0i}) > \tau \right\}

+ \Pr \left[ \hat{F}_0 (Y_{0i}) \leq s, F_0 (Y_{0i}) > s \right]

+ \Pr \left[ F_0 (Y_{0i}) \leq s, \hat{F}_0 (Y_{0i}) > s \right].
Therefore,

\[
E \left( \sum \frac{1}{\sqrt{n_\sigma_n^d}} \sum_{i=1}^{n} E \left( K \left( \frac{X_i - x}{\sigma_n} \right) \left| 1 \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) \leq \tau, \hat{F}_0(Y_{0i}) \leq s \right) \right| -1 \left( F_1(Y_{i1}) - F_0(Y_{0i}) \leq \tau, F_0(Y_{0i}) \leq s \right) \right) \right)
\]

\[
= \frac{1}{\sqrt{n_\sigma_n^d}} \sum_{i=1}^{n} \left( K \left( \frac{X_i - x}{\sigma_n} \right) \left| 1 \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) \leq \tau, \hat{F}_0(Y_{0i}) \leq s \right) \right| -1 \left( F_1(Y_{i1}) - F_0(Y_{0i}) \leq \tau, F_0(Y_{0i}) \leq s \right) \right) \left( \frac{X_i - x}{\sigma_n} \right)
\]

\[
\leq \frac{1}{\sqrt{n_\sigma_n^d}} \sum_{i=1}^{n} \left( K \left( \frac{X_i - x}{\sigma_n} \right) \left| 1 \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) > \tau, F_0(Y_{0i}) \leq s \right) \right| -1 \left( F_1(Y_{i1}) - F_0(Y_{0i}) > \tau, F_0(Y_{0i}) \leq s \right) \right) \left( \frac{X_i - x}{\sigma_n} \right)
\]

\[
\quad \quad \quad + \frac{1}{\sqrt{n_\sigma_n^d}} \sum_{i=1}^{n} \left( K \left( \frac{X_i - x}{\sigma_n} \right) \left| 1 \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) \leq \tau, F_1(Y_{i1}) - F_0(Y_{0i}) > \tau \right) \right| -1 \left( F_1(Y_{i1}) - F_0(Y_{0i}) \leq s \right) \right) \left( \frac{X_i - x}{\sigma_n} \right)
\]

\[
\quad \quad \quad + \frac{1}{\sqrt{n_\sigma_n^d}} \sum_{i=1}^{n} \left( K \left( \frac{X_i - x}{\sigma_n} \right) \left| 1 \left( \hat{F}_1(Y_{i1}) > \tau, F_0(Y_{0i}) \leq s \right) \right| -1 \left( F_1(Y_{i1}) - F_0(Y_{0i}) \leq s \right) \right) \left( \frac{X_i - x}{\sigma_n} \right)
\]

\[
\equiv S_{1n} + S_{2n} + S_{3n} + S_{4n}, \text{ say.} \tag{25}
\]

We will show that \( S_{1n} \to 0 \) and an exactly analogous proof will show that \( S_{2n}, S_{3n}, S_{4n} \) are also \( o(1) \).

Now, for fixed \( X_i, Y_{0i}, Y_{1i} \) and the fact that e.g. \( \hat{F}_1(Y_{i1}) = \frac{1}{n-1} \sum_{j \neq i} 1(Y_{ij} \leq Y_{i1}) \), we have that

\[
\Pr \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) > \tau, F_1(Y_{i1}) - F_0(Y_{0i}) < \tau \mid X_i, Y_{0i}, Y_{1i} \right)
\]

\[
= \Pr \left( \hat{F}_1(Y_{i1}) - \hat{F}_0(Y_{0i}) - (F_1(Y_{i1}) - F_0(Y_{0i})) > \tau - (F_1(Y_{i1}) - F_0(Y_{0i})), \right.
\]

\[
\left. F_1(Y_{i1}) - F_0(Y_{0i}) < \tau \mid X_i, Y_{0i}, Y_{1i} \right)
\]

\[
\leq \exp \left( -2 (n-1) \frac{\tau}{\tau} \right) \times 1(F_1(Y_{i1}) - F_0(Y_{0i}) < \tau),
\]

by Hoeffding’s inequality (note that conditional on \( Y_{i1} \), \( \hat{F}_1(Y_{i1}) = \frac{1}{n-1} \sum_{j \neq i} 1(Y_{ij} \leq Y_{i1}) \) is an average of independent, binary \((0, 1)\) random variables, thus satisfying the hypothesis.
of Hoeffding’s inequality). Thus, we have that
\[
S_{1n} \leq \frac{1}{\sqrt{n\sigma_n^d}} \sum_{i=1}^{n} E_{X_i, Y_0,Y_1} \left[ K \left( \frac{X_i-x}{\sigma_n} \right) \exp \left( -2(n-1) (\tau - (F_1(Y_{1i}) - F_0(Y_{0i})))^2 \right) \right] \\
= \frac{n}{\sqrt{n\sigma_n^d}} E_{X,Y_0,Y_1} \left[ K \left( \frac{X-x}{\sigma_n} \right) \exp \left( -2(n-1) (\tau - (F_1(Y_1) - F_0(Y_0)))^2 \right) \right] \\
= \frac{n}{\sqrt{n\sigma_n^d}} E_X \left[ K \left( \frac{X-x}{\sigma_n} \right) G_n(X) \right], \text{ where} \\
G_n(x) = E_{Y_0,Y_1|X} \left[ \exp \left( -2(n-1) (\tau - (F_1(Y_1) - F_0(Y_0)))^2 \right) \right].
\]
Continuing with the previous display, we have
\[
S_{1n} \leq \frac{n}{\sqrt{n\sigma_n^d}} E_X \left[ K \left( \frac{X-x}{\sigma_n} \right) G_n(X) \right] \\
= \frac{n\sigma_n^d}{\sqrt{n\sigma_n^d}} \int [K (u) G_n(x+\sigma_n u) f(x+\sigma_n u)] \, du \\
= \sqrt{n\sigma_n^d} \int [K (u) G_n(x) f(x+\sigma_n u)] \, du \\
= f(x) \sqrt{n\sigma_n^d} \int K (u) G_n(x) \, du + \text{terms of smaller order} \\
= f(x) \sqrt{n\sigma_n^d} G_n(x) + \text{terms of smaller order.} \tag{26}
\]
Now, notice that \(G_n(x)\) is of the form
\[
G_n(x) = E_{Z|X} \left[ \exp \left( -2(n-1) Z^2 \right) \times 1(Z > 0) \right]|X = x \\
\leq c \int \exp \left( -2(n-1) z^2 \right) f(z|x) \, dz \\
\leq c' \int \exp \left( -2(n-1) z^2 \right) \, dz \\
= O \left( n^{-1/2} \right) \tag{27}
\]
by the normal (Gaussian) integral formula. From (26) and (27), it follows that
\[
E \left( \sqrt{n\sigma_n^d} |T_{2n}| \right) = O \left( n^{-1/2} \times \sqrt{n\sigma_n^d} \right) = O \left( \sqrt{\sigma_n^d} \right) = o(1).
\]
Together with analogous proofs for $S_{2n}, S_{3n}, S_{4n}$, this implies that in (14), $\sqrt{n\sigma^2} T_{2n} = o_p(1)$. ■

**Uniform confidence bands for conditional mobility (following Hardle (1990), theorem 4.3.1):**

For each sample value $x$ of the conditioning variable $X$, bandwidth $\sigma_n$ and kernel $K(\cdot)$, denote estimated density at $X = x$ by

$$\hat{f}(x) = \frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right).$$

Consider dependent variables $W_i = 1 \left( \hat{F}_1(Y_{1i}) - \hat{F}_0(Y_{0i}) > \tau, \hat{F}_0(Y_{0i}) \leq s \right)$ for upward mobility and $W_i = 1 \left( Y_{1i} \leq \hat{\zeta}_1, Y_{0i} \leq \hat{\zeta}_0 \right)$ for transition probability. Denote regression estimate (predicted value) at $X = x$ by

$$\hat{\mu}(x) = \frac{1}{n\sigma_n} \sum_{i=1}^{n} K \left( \frac{x_i - x}{\sigma_n} \right) W_i \frac{\hat{f}(x)}{\hat{f}(x)}.$$ 

Corresponding to Epanechnikov kernel, set

$$c_K = \int_{-1}^{1} K^2(u) \, du = \int_{-1}^{1} \frac{9}{16} (1 - u^2)^2 \, du = 3/5 = 0.6$$

$$C_2 = \frac{\int_{-1}^{1} \{ K'(u) \}^2 \, du}{2c_K} = 1.25$$

- $$\delta = \sqrt{2 \ln \left( \frac{1}{\sigma_n} \right)}$$

- $$d_n = \sqrt{\left( 2 \ln \left( \frac{1}{\sigma_n} \right) \right) + \frac{1}{2\delta} \ln \left( \frac{C_2}{2\pi^2} \right)}$$

  $$= \delta + \frac{1}{2\delta} \ln \left( \frac{C_2}{2\pi^2} \right)$$

- $$c_\alpha = - \ln (-0.5 \times \ln (1 - .05)) = 3.66$$
\[
\omega^2 (x) = \frac{1}{n\sigma_n \hat{f}(x)} \sum_{i=1}^{n} \{W_i - \hat{\mu}(X_i)\}^2 K \left( \frac{X_i - x}{\sigma_n} \right)
\]

Then lower and upper limit of uniform CI are given by

\[
\begin{align*}
CLO (x) &= \hat{\mu} (x) - \left\{ \frac{c_{\alpha}}{\delta} + d_n \right\} \times \frac{\sqrt{\omega^2 (x)}}{c_K} \\
CUP (x) &= \hat{\mu} (x) + \left\{ \frac{c_{\alpha}}{\delta} + d_n \right\} \times \frac{\sqrt{\omega^2 (x)}}{c_K}.
\end{align*}
\]
References


Working Paper Series

A series of research studies on regional economic issues relating to the Seventh Federal Reserve District, and on financial and economic topics.

Standing Facilities and Interbank Borrowing: Evidence from the Federal Reserve’s New Discount Window
Craig Furfine

Netting, Financial Contracts, and Banks: The Economic Implications
William J. Bergman, Robert R. Bliss, Christian A. Johnson and George G. Kaufman

Real Effects of Bank Competition
Nicola Cetorelli

Finance as a Barrier To Entry: Bank Competition and Industry Structure in Local U.S. Markets?
Nicola Cetorelli and Philip E. Strahan

The Dynamics of Work and Debt
Jeffrey R. Campbell and Zvi Hercowitz

Fiscal Policy in the Aftermath of 9/11
Jonas Fisher and Martin Eichenbaum

Merger Momentum and Investor Sentiment: The Stock Market Reaction To Merger Announcements
Richard J. Rosen

Earnings Inequality and the Business Cycle
Gadi Barlevy and Daniel Tsiddon

Platform Competition in Two-Sided Markets: The Case of Payment Networks
Sujit Chakravorti and Roberto Roson

Nominal Debt as a Burden on Monetary Policy
Javier Díaz-Giménez, Giorgia Giovannetti, Ramon Marimon, and Pedro Teles

On the Timing of Innovation in Stochastic Schumpeterian Growth Models
Gadi Barlevy

Policy Externalities: How US Antidumping Affects Japanese Exports to the EU
Chad P. Bown and Meredith A. Crowley

Sibling Similarities, Differences and Economic Inequality
Bhashkar Mazumder

Determinants of Business Cycle Comovement: A Robust Analysis
Marianne Baxter and Michael A. Kourapisas

The Occupational Assimilation of Hispanics in the U.S.: Evidence from Panel Data
Maude Toussaint-Comeau
Working Paper Series (continued)

Reading, Writing, and Raisinets¹: Are School Finances Contributing to Children’s Obesity?  
Patricia M. Anderson and Kristin F. Butcher  
WP-04-16

Learning by Observing: Information Spillovers in the Execution and Valuation of Commercial Bank M&As  
Gayle DeLong and Robert DeYoung  
WP-04-17

Prospects for Immigrant-Native Wealth Assimilation: Evidence from Financial Market Participation  
Una Okonkwo Osili and Anna Paulson  
WP-04-18

Individuals and Institutions: Evidence from International Migrants in the U.S.  
Una Okonkwo Osili and Anna Paulson  
WP-04-19

Are Technology Improvements Contractionary?  
Susanto Basu, John Fernald and Miles Kimball  
WP-04-20

The Minimum Wage, Restaurant Prices and Labor Market Structure  
Daniel Aaronson, Eric French and James MacDonald  
WP-04-21

Betcha can’t acquire just one: merger programs and compensation  
Richard J. Rosen  
WP-04-22

Not Working: Demographic Changes, Policy Changes, and the Distribution of Weeks (Not) Worked  
Lisa Barrow and Kristin F. Butcher  
WP-04-23

The Role of Collateralized Household Debt in Macroeconomic Stabilization  
Jeffrey R. Campbell and Zvi Hercowitz  
WP-04-24

Advertising and Pricing at Multiple-Output Firms: Evidence from U.S. Thrift Institutions  
Robert DeYoung and Evren Örs  
WP-04-25

Monetary Policy with State Contingent Interest Rates  
Bernardino Adão, Isabel Correia and Pedro Teles  
WP-04-26

Comparing location decisions of domestic and foreign auto supplier plants  
Thomas Klier, Paul Ma and Daniel P. McMillen  
WP-04-27

China’s export growth and US trade policy  
Chad P. Bown and Meredith A. Crowley  
WP-04-28

Where do manufacturing firms locate their Headquarters?  
J. Vernon Henderson and Yukako Ono  
WP-04-29

Monetary Policy with Single Instrument Feedback Rules  
Bernardino Adão, Isabel Correia and Pedro Teles  
WP-04-30
Working Paper Series (continued)

Firm-Specific Capital, Nominal Rigidities and the Business Cycle
*David Altig, Lawrence J. Christiano, Martin Eichenbaum and Jesper Linde*

WP-05-01

Do Returns to Schooling Differ by Race and Ethnicity?
*Lisa Barrow and Cecilia Elena Rouse*

WP-05-02

Derivatives and Systemic Risk: Netting, Collateral, and Closeout
*Robert R. Bliss and George G. Kaufman*

WP-05-03

Risk Overhang and Loan Portfolio Decisions
*Robert DeYoung, Anne Gron and Andrew Winton*

WP-05-04

Characterizations in a random record model with a non-identically distributed initial record
*Gadi Barlevy and H. N. Nagaraja*

WP-05-05

Price discovery in a market under stress: the U.S. Treasury market in fall 1998
*Craig H. Furfine and Eli M. Remolona*

WP-05-06

Politics and Efficiency of Separating Capital and Ordinary Government Budgets
*Marco Bassetto with Thomas J. Sargent*

WP-05-07

Rigid Prices: Evidence from U.S. Scanner Data
*Jeffrey R. Campbell and Benjamin Eden*

WP-05-08

Entrepreneurship, Frictions, and Wealth
*Marco Cagetti and Mariacristina De Nardi*

WP-05-09

Wealth inequality: data and models
*Marco Cagetti and Mariacristina De Nardi*

WP-05-10

What Determines Bilateral Trade Flows?
*Marianne Baxter and Michael A. Kouparitsas*

WP-05-11

Intergenerational Economic Mobility in the U.S., 1940 to 2000
*Daniel Aaronson and Bhashkar Mazumder*

WP-05-12

Differential Mortality, Uncertain Medical Expenses, and the Saving of Elderly Singles
*Mariacristina De Nardi, Eric French, and John Bailey Jones*

WP-05-13

Fixed Term Employment Contracts in an Equilibrium Search Model
*Fernando Alvarez and Marcelo Veracierto*

WP-05-14

Causality, Causality, Causality: The View of Education Inputs and Outputs from Economics
*Lisa Barrow and Cecilia Elena Rouse*

WP-05-15
Competition in Large Markets
Jeffrey R. Campbell
WP-05-16

Why Do Firms Go Public? Evidence from the Banking Industry
Richard J. Rosen, Scott B. Smart and Chad J. Zutter
WP-05-17

Clustering of Auto Supplier Plants in the U.S.: GMM Spatial Logit for Large Samples
Thomas Klier and Daniel P. McMillen
WP-05-18

Why are Immigrants’ Incarceration Rates So Low? Evidence on Selective Immigration, Deterrence, and Deportation
Kristin F. Butcher and Anne Morrison Piehl
WP-05-19

Constructing the Chicago Fed Income Based Economic Index – Consumer Price Index: Inflation Experiences by Demographic Group: 1983-2005
Leslie McGranahan and Anna Paulson
WP-05-20

Universal Access, Cost Recovery, and Payment Services
Sujit Chakravorti, Jeffery W. Gunther, and Robert R. Moore
WP-05-21

Supplier Switching and Outsourcing
Yukako Ono and Victor Stango
WP-05-22

Do Enclaves Matter in Immigrants’ Self-Employment Decision?
Maude Toussaint-Comeau
WP-05-23

The Changing Pattern of Wage Growth for Low Skilled Workers
Eric French, Bhashkar Mazumder and Christopher Taber
WP-05-24

U.S. Corporate and Bank Insolvency Regimes: An Economic Comparison and Evaluation
Robert R. Bliss and George G. Kaufman
WP-06-01

Redistribution, Taxes, and the Median Voter
Marco Bassetto and Jess Benhabib
WP-06-02

Identification of Search Models with Initial Condition Problems
Gadi Barlevy and H. N. Nagaraja
WP-06-03

Tax Riots
Marco Bassetto and Christopher Phelan
WP-06-04

The Tradeoff between Mortgage Prepayments and Tax-Deferred Retirement Savings
Gene Amromin, Jennifer Huang, and Clemens Sialm
WP-06-05

Why are safeguards needed in a trade agreement?
Meredith A. Crowley
WP-06-06
Working Paper Series (continued)

Taxation, Entrepreneurship, and Wealth
Marco Cagetti and Mariacristina De Nardi
WP-06-07

A New Social Compact: How University Engagement Can Fuel Innovation
Laura Melle, Larry Isaak, and Richard Mattoon
WP-06-08

Mergers and Risk
Craig H. Furfine and Richard J. Rosen
WP-06-09

Two Flaws in Business Cycle Accounting
Lawrence J. Christiano and Joshua M. Davis
WP-06-10

Do Consumers Choose the Right Credit Contracts?
Sumit Agarwal, Souphala Chomsisengphet, Chunlin Liu, and Nicholas S. Souleles
WP-06-11

Chronicles of a Deflation Unforetold
François R. Velde
WP-06-12

Female Offenders Use of Social Welfare Programs Before and After Jail and Prison:
Does Prison Cause Welfare Dependency?
Kristin F. Butcher and Robert J. LaLonde
WP-06-13

Eat or Be Eaten: A Theory of Mergers and Firm Size
Gary Gorton, Matthias Kahl, and Richard Rosen
WP-06-14

Do Bonds Span Volatility Risk in the U.S. Treasury Market?
A Specification Test for Affine Term Structure Models
Torben G. Andersen and Luca Benzoni
WP-06-15

Transforming Payment Choices by Doubling Fees on the Illinois Tollway
Gene Amromin, Carrie Jankowski, and Richard D. Porter
WP-06-16

How Did the 2003 Dividend Tax Cut Affect Stock Prices?
Gene Amromin, Paul Harrison, and Steven Sharpe
WP-06-17

Will Writing and Bequest Motives: Early 20th Century Irish Evidence
Leslie McGranahan
WP-06-18

How Professional Forecasters View Shocks to GDP
Spencer D. Krane
WP-06-19

Evolving Agglomeration in the U.S. auto supplier industry
Thomas Klier and Daniel P. McMillen
WP-06-20

Mortality, Mass-Layoffs, and Career Outcomes: An Analysis using Administrative Data
Daniel Sullivan and Till von Wachter
WP-06-21
Working Paper Series (continued)

The Agreement on Subsidies and Countervailing Measures: Tying One’s Hand through the WTO.  
Meredith A. Crowley WP-06-22

How Did Schooling Laws Improve Long-Term Health and Lower Mortality?  
Bhashkar Mazumder WP-06-23

Manufacturing Plants’ Use of Temporary Workers: An Analysis Using Census Micro Data  
Yukako Ono and Daniel Sullivan WP-06-24

What Can We Learn about Financial Access from U.S. Immigrants?  
Una Okonkwo Osili and Anna Paulson WP-06-25

Bank Imputed Interest Rates: Unbiased Estimates of Offered Rates?  
Evren Ors and Tara Rice WP-06-26

Welfare Implications of the Transition to High Household Debt  
Jeffrey R. Campbell and Zvi Hercowitz WP-06-27

Last-In First-Out Oligopoly Dynamics  
Jaap H. Abbring and Jeffrey R. Campbell WP-06-28

Oligopoly Dynamics with Barriers to Entry  
Jaap H. Abbring and Jeffrey R. Campbell WP-06-29

Risk Taking and the Quality of Informal Insurance: Gambling and Remittances in Thailand  
Douglas L. Miller and Anna L. Paulson WP-07-01

Fast Micro and Slow Macro: Can Aggregation Explain the Persistence of Inflation?  
Filippo Altissimo, Benoît Mojon, and Paolo Zaffaroni WP-07-02

Assessing a Decade of Interstate Bank Branching  
Christian Johnson and Tara Rice WP-07-03

Debit Card and Cash Usage: A Cross-Country Analysis  
Gene Amromin and Sujit Chakravorti WP-07-04

The Age of Reason: Financial Decisions Over the Lifecycle  
Sumit Agarwal, John C. Driscoll, Xavier Gabaix, and David Laibson WP-07-05

Information Acquisition in Financial Markets: a Correction  
Gadi Barlevy and Pietro Veronesi WP-07-06

Monetary Policy, Output Composition and the Great Moderation  
Benoît Mojon WP-07-07

Estate Taxation, Entrepreneurship, and Wealth  
Marco Cagetti and Mariacristina De Nardi WP-07-08
Working Paper Series *(continued)*

Conflict of Interest and Certification in the U.S. IPO Market  
*Luca Benzoni and Carola Schenone*  
WP-07-09

The Reaction of Consumer Spending and Debt to Tax Rebates –  
Evidence from Consumer Credit Data  
*Sumit Agarwal, Chunlin Liu, and Nicholas S. Souleles*  
WP-07-10

Portfolio Choice over the Life-Cycle when the Stock and Labor Markets are Cointegrated  
*Luca Benzoni, Pierre Collin-Dufresne, and Robert S. Goldstein*  
WP-07-11

Nonparametric Analysis of Intergenerational Income Mobility with Application to the United States  
*Debopam Bhattacharya and Bhashkar Mazumder*  
WP-07-12