On Subjectivity of Intergenerational Mobility Measures

Esfandiar Maasoumi, Le Wang, Daiqiang Zhang†

Abstract

We study the subjective value judgements implicit in regression-based measures of intergenerational mobility (IGM). These measures can be represented as weighted sums of mobility or elasticities over heterogeneous families and subgroups. We first clarify the implicit decision-theoretic foundation of the two dominant regression-based mobility measures (level and rank regressions). Our analysis clarifies the still under-developed, and different, notions of "mobility" that are inherent in statistical measurements. We suggest alternatives that are equally accessible and computable. Our approach to constructing IGMs is motivated by well founded principles in the literature on inequality, poverty, and cross section mobility. It highlights the near inevitable role of aversion to inequality and poverty in perception of mobility as enhanced wellbeing. Our approach is computationally convenient and can be readily extended to incorporate additional covariates for counterfactual analysis. Using the PSID data, we assess the implications for policy analysis and measurement. We estimate several measures of IGM that demonstrate a nuanced view of mobility, as well as our perspective on geographic differences in mobility and the dynamics of it. These perspectives appear to have been obscured by a veil of overall averages and regression coefficients.

JEL Classifications:

Keywords:

† Corresponding address: We thank participants at XXX

1 Introduction

There has been a growing public concern over rising inequality and the potentially unequal and limited economic opportunity for children from less advantaged families, and their prospects to move up the economic ladder. Estimation and measurement of intergenerational mobility (IGM) is essential for characterizing the level of equality of opportunity. The current debate and interchanges are dominantly informed by partial correlation measures derived from generally linear regression models for income levels and/or ranks of offsprings incomes and their parents. High "correlation" or "dependence" between incomes of generations indicates "immobility". Surveying the field, Mazumder (2018a) notes the estimates range from less than .3 in early studies, to around .5 or higher in a number of papers using the Panel Study of Income Dynamics (PSID), to .34 in the recent, influential study of Chetty et al. (2014) using administrative data. "Where exactly the United States and other countries lie on this spectrum has been a central question in the intergenerational income mobility literature." (Mazumder, 2018a)

There are many econometric challenges surrounding the estimation of mobility, for example, nonlinearity/heterogeneity issues and the life-cycle bias inherent in point-in-time incomes. Here, we focus on an important yet neglected problem of subjective aggregation that is implicit in these regression approaches, and more generally, in any summary measure of a deeply heterogeneous phenomenon as intergenerational mobility.

Due to differences in factors such as incentives to invest in human capital across family incomes, neighborhood or peer effects, education, and a myriad of other factors, the relationship between parent and child incomes is likely heterogenous and nonlinear. As a result, the "correlation" between these incomes would vary with parental income and other factors. Such nonlinearity exists and can differ across groups, as is evidenced in recent empirical studies (e.g., Landersø and Heckman (2017), An et al. (2020)). To be specific, economic theories predict a rather complex relationship connecting child income to parental income (in the absence of other covariates) given by

$$Y = g(X) + \epsilon$$

where Y and X are the logarithm of child income and parental income, respectively, $g(\cdot)$ is an unknown function. The derivative, $g'(x)$ is the slope which measures persistence, interpreted as the inverse of mobility at $X = x$, and ϵ is an error term. In the presence of nonlinearity, income mobility varies with parental income across the population.

An important insight from the literature (e.g., Yitzhaki (1996); Løken et al. (2012)) is that the coefficient from a linear projection/regression ($Y = X\beta + u$) of an underlying nonlinear function in (1) can be shown to be of the following form

$$\beta = \int w(x)g'(x)dx$$

where $w(x)$ is a weighting scheme depending on the parental income, $X = x$. In other words, the commonly used linear regression estimate of the slope is only a *summary* measure or weighted sum of heterogeneous effects across the population subgroups.

Building on this important insight, our first step is to reveal the otherwise implicit, subjective weighting of subgroup mobilities in the current regression measures of IGM. In the context of IGM, β is the intergenerational mobility elasticity (IGE), and it represents a summary measure or a weighted sum of heterogenous correlations between parental and child incomes that vary across families. The weights depend on the parental incomes. Different, low, middle or high income groups and observations, for instance, receive weights to obtain "Least Squares" or satisfy *statistical* orthogonality assumptions. These implicit weights are no less subjective than for any other desired criterion. All observation weights correspond to implicit welfare evaluations of the corresponding individuals or groups. Linear projection weights compete with alternative weights that are consistent with, for example, Lorenz functions and Gini measures. We further extend this insight to the rank-rank regression, another popular regression-based measure of IGE. We show that in fact the rank-rank coefficient can also be represented as a weighted average of the heterogenous IGEs; and that the implicit weights in this context are not even necessarily proper weights. For the first time, we present a systematic analysis of the relation between the rank-rank and level regressions and their weighting schemes. This is useful and necessary for an understanding of the fundamentally different notions of mobility/inequality inherent in these two approaches. [1]

The impact and implications of this first-step in our analysis are profound. Depending on the choice of a summary measure (estimation), one can have drastically different assessments of mobility with different policy implications. Comparative analysis of mobility such as between-groups is subjectively fragile, and can be particularly misleading and

[1] Both approaches are seen to belong to scale invariant notions of status. Doubling of incomes or ranks, clearly a massive degree of "mobility", leaves correlation measures unchanged! The "growth" component of mobility is neglected by IGE regressions, only limited aspects of "exchange mobility" are assessed. See Maasoumi and Mahmoudi (2011?) for a decomposition of these components of any change in inequality and poverty measures.

uninformative about the features of the mobility with which we are generally concerned, for example, for children from the disadvantaged families.[2] Our insights also present a great challenge for mobility-related research. Consider the example of the influential Great Gatsby Curve. The literature typically links regression-based mobility measures to the *Gini Coefficients*, a popular inequality measure. The latter is based on a weighted average of heterogenous incomes with the weighting scheme motivated by the Gini welfare function. The corresponding weighting scheme in the regression/estimation exercise is not compatible, and as shown here, is in contrast to the Gini welfare function. This incompatibility, and the implicitness of regression weights questions the meaning and value implications of using these different weight schemes in the same assessment.

One alternative approach is to avoid aggregation and employ estimation by nonparametric methods, when it is practicable, or use of nonlinear models. Such solution, however, deserves closeutiny. It is implicit in these alternatives that the pattern, although nonlinear, is readily discernible, summarized and compared between groups, particularly for more detailed but easily classified groups such as the lower, middle, and upper tails of the distribution. A nonparametric model can be viewed as having infinitely many parameters, however, and in practice only a few estimates can actually be presented. This does not resolve the challenge of summarization. Moreover, the traditional nonparametric kernel estimation is computationally intensive and infeasible with large datasets (such as administrative records), and perform poorly and thus unreliably in smaller samples (such as Panel Studies of Income Dynamics, PSID) due to the slow rate of convergence; it is even slower and more severe in our context since we are interested in the derivatives of these nonparametric functions, $g'(x)$.[3] [4]

[2] A comparison of *multiple* correlation estimates of mobility between groups, over time, and across space is also required for a better and deeper understanding of mobility, as well as monitoring progress in mobility and assessment of the policies. Indeed, existing empirical studies present such comparisons, from which many influential findings emerge in the literature. For example, by conducting cross-country comparisons, the literature has found that the welfare-state economies such as Denmark and Noway have higher levels of mobility than more market-oriented economies such as the United States (e.g., Landersø and Heckman (2017), Solon (2002)). Cross-country differences in IGMs are also further found to be positively correlated with the differences in inequality, a relationship called "The Great Gatsby Curve" (Krueger (2012)). Chetty and his coauthors document stark differences in IGMs across regions (Chetty et al. (2014)) and between whites and blacks (Chetty et al. (2020)). These influential results have prompted further analysis to investigate the sources of mobility to explain such differences.

[3] See Li and Racine (2007) for the relationship between the rate of convergence and the order of smoothness.

[4] The challenge of summarizing pertains to both nonparametric kernel estimation and assessment of transition matrices which represent a wider view of the conditional distribution that underlies the IGE regressions. These alternatives have been criticized for their "overtly disaggregate nature" (Bhattacharya and Mazumder (2011) and Hertz (2005) points out that with transition matrices, "there is no best way to summarize their content", prompting development of an easier-to-interpret summary measure of mo-

Aggregation is inevitable, especially for between-group comparisons (as for the Great Gatsby curve) and policy metrics. One may consider average derivatives of the nonparametric functions, effectively imposing equal weights across groups. As Sen (1992) notes, however, while such weighting scheme is often considered as a form of egalitarianism, "the effect of ignoring the interpersonal variations can, in fact, be deeply inegalitarian", and "equal consideration for all may demand very unequal treatment" in favor of a particular group.

In this paper, we examine an alternative approach within the popular context of parametric regressions. But we propose computationally convenient ways of incorporating subgroup weighting schemes with explicit, desirable properties consistent with well known social welfare or evaluation functions.

Two estimation approaches are considered here: the first is based on the preference functions underlying the extended Gini family of social welfare functions (similar to Yitzhaki (1996)), while the second is inspired by the Lorenz family of social welfare functions. Both estimators exhibit the property that the income value at which the weight to local IGE is maximized decreases with the degree of inequality aversion. In other words, the higher the inequality aversion is, the more weight is assigned to the lower tail of the income distribution. It also circumvents the need to estimating fully nonparametric models and can be readily implemented in practice with standard statistical softwares such as Stata. The flexibility and computational convenience of our approach thus allows practitioners more likely to focus on the conceptual issues at hand. The tools to flexibly incorporate covariates and perform counterfactual analysis is relatively underdeveloped in the literature of IGM. We also further extend our framework to incorporate covariates in our analysis, which will prove useful for practitioners to perform counterfactual analysis to understand the sources of mobility, as well as the differences across groups.

We illustrate our proposal using the PSID data and reach several main conclusions. These results underscore the subjectivity in a summary measure and highlight their limitations by the classic Arrowian impossibility results. First, our results imply a significant heterogeneity and nonlinearity in the income transmission process. The estimates vary considerably with different methods and inequality-aversion parameters, and hence with their underlying weighting mechanisms. The weighting schemes matter for forming our impression of the mobility in a society. Both of our level and rank-rank regression results

bility. These criticisms are valid if over stated. For instance, there do exist several summary measures of transition matrices, albeit also non consensus ones. These include Bartholomeu and Shorrocks, amongst others.

are generally consistent with prior overall findings in the literature; the level-regression estimates suggest a substantially less mobility than the rank-rank regression estimates]. The results based on both the Gini and Lorenz evaluation functions suggest that with larger level of aversion to inequality, and focusing more on the individuals from disadvantaged families, smaller IGE coefficients are obtained, implying a more mobile society. The Lorenz results suggest that when placing more weights on the richer families, we also obtain smaller IGE coefficients. These results together indicate that children from both disadvantaged and richer families may have a higher level of mobility than those from the "middle class".

Second, the traditional perspectives on geographic disparities in mobility are also challenged when considering different underlying weighting schemes. Both the level and rank-rank regressions suggest a significant disparity in IGMs between the South and the West. However, the former implies a smaller geographic difference than the latter. Varying the inequality aversion parameters revises not only the magnitudes of the geographic disparities in IGMs, but also the patterns. For example, when using the Gini mobility measures and placing more weights on the children from the disadvantaged families, we actually find that the West becomes less and less mobile, relative to the rest of the country, including the South. By contrast, when using the Lorenz measures and placing more weights on the richer families, we observe a larger coefficient for the South. These results may suggest that there may exist "affluence trap" in the South, but for the rest of the country, mobility is higher in both tails than in the middle.

Finally, we also examine the dynamics of the IGMs across cohorts. Both the level and rank-rank regressions suggest that mobility declines for the cohort born before 1954 and the cohort born after 1968. But the dynamics are different. More importantly, when placing more weights on the individuals from the more disadvantaged families, we actually observe that it is actually more mobile for the later cohort born after 1968 than the early cohort born before 1954.

Relationship to the Literature Our paper contributes broadly to two separate literatures other than mobility. First, similar to Maasoumi and Wang (2019), our paper is part of an attempt to connect and integrate the inequality literature more formally with the literature on IGM. On the other hand, our focus differs drastically from the inequality literature in that the latter focuses on univariate distributions, while we deal with the joint distribution of two incomes/outcomes.

Second, our econometric method is related to a growing interest in the treatment effect literature to understand the underlying meaning of various conventional estimators such as instrumental variable estimators (Mogstad and Wiswall (2016)), two-way fixed effects (De Chaisemartin and d'Haultfoeuille (2020)) and difference-in-difference estimators (De Chaisemartin and dHaultfoeuille (2018)) in the presence of heterogeneous treatment effects. Unlike the treatment effects literature, in which the focus is often on a summary measure of treatment effects for a clearly defined subgroup, our paper is based on a summary measure for the entire population with economically or policy motivated weighting schemes.

The rest of the paper is organized as follows. Section 2 provides a numerical example to motivate our analysis. Section 3 exposes the conceptual issues for the traditional regression approaches by analyzing their weighting schemes. Section 4 presents our estimators. We illustrate our proposals using the PSID data in Section 5. Section 6 concludes. Proofs are collected in the Appendix.

2 A Motivating Example

We begin with a simplistic example to highlight the importance of weighting schemes in constructing a summary measure of mobility and motivate our proposed mobility measures. We have two goals: first, we are interested in measuring the extent of mobility in a society or group, say, group A. Second, we are interested in comparing the mobility between this society and another society, say, group B.

In this simplistic world with heterogeneity, there are only two levels of incomes: low-income, l and high-income, h. Suppose the income function for each group is given by

$$Y_A = g_A(X) + \epsilon_A \quad \text{and} \quad Y_B = g_B(X) + \epsilon_B,$$

where $X \in \{l, h\}$ with $0 < l < h < \infty$. The corresponding mobility at $X = x$ is given by $g'_A(l) = 0.6$, $g'_A(h) = 0.3$, $g'_B(l) = 0.9$, $g'_B(h) = 0.2$.[5] In other words, income in the group A is less persistent and more mobile in the lower tail, while income in the group B is less persistent and more mobile in the upper tail. A weighted average measure of mobility for group k is as follow

$$\overline{m}_k = w_l g'_k(l) + w_h g'_k(h)$$

[5] A little bit abuse of notation here about the derivative, but it simply represents the mobility at this particular income level.

where $w_l + w_h = 1$ and $0 < w_l, w_h < 1$. There could be two different weighting schemes (corresponding to two different estimation methods), $w_l^j, w_h^j, j = 1, 2$.

Case 1: Lets see how our impression of the mobility for a particular group/society depends on the weighting schemes. Suppose the first weighting scheme (for a particular estimation method) is $w_l^1 = .95$ and $w_h^1 = .05$. Then, $\overline{m}_A = .585$. However, if we reverse the weighting scheme instead by placing more weights on the richest families with $w_l^2 = .05$ and $w_h^2 = .95$, then $\overline{m}_A = .315$. These two numbers resemble the current debate on the magnitudes of the IGM for the U.S.. The discrepancy is large since the mobility implied by the first weighting scheme is drastically different from that suggested by the second weighting scheme.

Case 2: Now consider how varying weighting schemes may impact the conclusions regarding the between-group comparison of mobility. Under the first weighting scheme when we place more weights on the children from the disadvantaged families with $w_l^1 = .95$ and $w_h^1 = .05$, it follows $\overline{m}_A = .585 < .865 = \overline{m}_B$. The implied mobility of group A is actually *smaller* than that of group B. By contrast, under the second weighting scheme, $w_l^2 = .05$ and $w^2 = .95$, we actually observe the opposite result $\overline{m}_A = .315 > .235 = \overline{m}_B$. In other words, group A is less more mobile than group B.

The above examples may be extreme. In practice, the weighting schemes should depend on the underlying income of each group as well as the social weight attached to that group. Nevertheless, these examples highlight the importance of weighting schemes, and subjectivity in constructing a scalar measure for a society and further between-group comparisons. The key questions are whether there could be any consensus robust to the changes in the weighting schemes, and whether these weighting schemes are indeed sensible and reflect policy goals or our concerns, for example, about individuals from disadvantaged families with lower level of socio-economic statuses.

3 Weighted Average Representations of Traditional Regression-based Mobility Measures

In this section, we characterize the properties of the existing summary measures of mobility based on linear level regressions and rank-rank regressions. Both of these summary measures can be expressed as a weighted average of heterogeneous mobility across the distribution of parents income. These weights are implicit and indeed subjective. The results for level regressions follow closely Yitzhaki (1996), while the results for rank-rank regressions are new to the literature.

Suppose that the true income-transmission process is given by

$$Y = g(X) + \epsilon, \tag{1}$$

where the error term ϵ satisfies $\mathbb{E}(\epsilon|X) = 0$. The implied *rank-rank* relationship is given by

$$V = \mathbb{E}(V|U) + \eta = g_r(U) + \eta, \tag{2}$$

where $U = F_X(X)$ (rank of a child's income), $V = F_Y(Y)$ (rank of a parent's income), and the error term η satisfies $\mathbb{E}(\eta|U) = 0$ by the definition $g_r(u) = \mathbb{E}(F_Y(Y)|F_X(x) = u)$.

The following Lemma (e.g., shown in Proposition 2 by Yitzhaki (1996)) clarifies the weighted average representation of the linear regression coefficient.[6]

Lemma 1. *[**Weighted Average Representation of Linear Regression**] Let $\mathbb{E}^*(Y|X) = \alpha + \beta X$ denote the best linear predictor of Y given X. Then*

$$\begin{aligned} \beta &= [Var(X)]^{-1} Cov(Y, X) = \int_{\mathcal{S}_X} w(x) g'(x) dx, \\ w(x) &= F_X(x) \left(\mu_X - \mathbb{E}\{X|X \leq x\}\right) \sigma_X^{-2}, \end{aligned}$$

where $\mu_X = \mathbb{E}(X)$, $\sigma_X^2 = Var(X)$, $w(x) \geq 0$, $\int_{\mathcal{S}_X} w(x)dx = 1$, and $\mathcal{S}_X$ is the support of X.

Moreover, Yitzhaki (1996) proves that under normal and uniform distributions, the weight function $w(\cdot)$ in Lemma 1 is maximized at the median of X.[7] This implies that the linear regression approach assigns *smaller* weights to the marginal effects in the lower and higher tails of the income distribution than the middle part of the income distribution. This property is certainly not desirable when assessing mobility with an aversion to inequality, and when the focus is on lower incomes.

Next, we consider the rank-rank regression. Since both U and V are distributed from the standard uniform in general, we use Lemma 1 to obtain the following result for the rank-rank regression in (2).

[6] Note that Yitzhaki (1996)'s results can be extended to the case of a multiple regression.

[7] For the normal distribution, $w(x)$ is equal to its density function. For the uniform distribution with the support $[\underline{x}, \overline{x}]$, $w(x) = 6(\overline{x} - x)(x - \underline{x})(\overline{x} - \underline{x})^{-3}$.

Lemma 2. *Let $\mathbb{E}^*(V|U) = \alpha_r + \beta_r U$ denote the best linear predictor of V given U. Then*

$$\beta_r = [Var(U)]^{-1}Cov(U,V) = \int_0^1 w(u)g_r'(u)du, \tag{3}$$

$$w(u) = 6u - 6u^2, \tag{4}$$

where $w(u) \geq 0$ for $0 \leq u \leq 1$, and $\int_0^1 w(u)du = 1$.

Lemma 2 implies that the weight attains the maximum at the median $x = F_X^{-1}(1/2)$. The closer to the median the x, the greater the weight.

Lemma 2 clarifies the weighted average representation of the derivatives of the rank-rank model, $g_r'(u)$, but not the weighted average representation of the actual correlation or persistence between incomes of different generations, $g'(x)$. Hence, the weighting schemes are not directly comparable between linear level regression and rank-rank regression. By substituting $U = F_X(X)$ into (3) and (4) in Lemma 2, we obtain

$$\beta_r = \int_{\mathcal{S}_X} w(F_X(x))g_r'(F_X(x))f_X(x)dx = \int_{\mathcal{S}_X} 6(F_X(x) - F_X^2(x))g_r'(F_X(x))f_X(x)\frac{g'(x)}{g'(x)}dx,$$

which, as summarized in the following Proposition, clarifies the weighted average representation of the derivative of the true income-transmission process $g(x)$ in (1).

Proposition 1. *[**Weighted Average Representation of Rank-Rank Regression**] Under conditions in Lemma 2, the summary measure β_r in (3) becomes*

$$\beta_r = \int_{\mathcal{S}_X} w_r(x)g'(x)dx, \tag{5}$$

where

$$w_r(x) = 6(F_X(x) - F_X^2(x))g_r'(F_X(x))f_X(x)(g'(x))^{-1}. \tag{6}$$

Proposition 1 exposes an undesirable property of the weighting scheme under the rank-rank regression. In general, the weighting scheme does not integrate to one ($\int w_r(x)dx \neq 1$) since typically $g_r'(F(x)) \neq g'(x)$ for any x. Analytical solutions derived for normal, lognormal, and uniform distributions in Corollary 1 illustrate and confirm this point. Consequently, the rank-rank measures do not meaningfully reflect concerns about mobility in income levels and intensity of movements in incomes. They can be even larger than the maximum of the heterogenous (levels) mobility or smaller than any of subgroup level mobilities.[8]

[8] One exception is for the uniform distributions. As stated in case (*iii*) of Corollary 1, under additional

Corollary 1. ***Weighting Schemes of Rank-Rank Regression for Three Parametric Distributions:*** *We provide the weighting schemes under three specific cases.*
(i) Normal Distribution. Suppose that the conditional distribution of ϵ given $X = x$ is normal with zero mean and variance σ_ϵ^2, and that Y is normally distributed with mean $\mu_y = \mathbb{E}(g(X))$ and variance $\sigma_y^2 = \sigma_\epsilon^2 + Var(g(X))$. Then,

$$w_r(x) = c(x)(F_X(x) - F_X^2(x)), \tag{7}$$

where

$$c(x) = \frac{6}{\sqrt{2\pi(\sigma_\epsilon^2 + \sigma_y^2)}} \times exp\left\{ -\frac{(\sigma_\epsilon^2 + \sigma_y^2)(\sigma_\epsilon^2 \mu_y^2 + g^2(x)\sigma_y^2) - (\mu_y \sigma_\epsilon^2 + g(x)\sigma_y^2)^2}{2\sigma_y^2 \sigma_\epsilon^2 (\sigma_\epsilon^2 + \sigma_y^2)} \right\}. \tag{8}$$

(ii) Lognormal Distribution. Under the same conditions as those under the normal distribution except that the conditional distribution of Y given $X = x$ is lognormal and that Y is lognormally distributed, $w_r(x)$ is identical to that in (7).
(iii) Uniform Distribution. Suppose that Y is uniformly distributed on $[\underline{y}, \overline{y}]$. Then,

$$w_r(x) = 6(F_X(x) - F_X^2(x))(\overline{y} - \underline{y})^{-1}. \tag{9}$$

If X is also uniformly distributed on $[\underline{x}, \overline{x}]$ with $\overline{x} - \underline{x} = \overline{y} - \underline{y}$ and $g(x) = \alpha + x\beta$,[9] the implied rank regression is also linear with $g_r'(u) = \beta_r$ and $\beta_r = \beta$. Moreover, $\int_{\underline{x}}^{\overline{x}} w_r(x)dx = 1$.

These results are novel in their own right, and for the first time, we are able to connect the rank regression findings to the underlying income transmission process for level incomes. In further work, these relations can also be illuminating in terms of the value of "descriptive statistics" for underlying objects in models, such as more complex measures of vertical and/or horizontal mobility.

restrictions for the case of uniform distributions, the implied rank-rank regression is also linear with the same measure of mobility as that in the level regression. Moreover, the underlying weighting scheme for this particular case is proper.

[9] Note that it allows for the lower and upper bounds of X and Y to be different. That is, both the lower and upper bounds for children income can be greater or smaller than that for parents income.

4 Mobility Measures Based on Social Preferences

Subjectivity in the implicit weighting schemes behind the traditional regression-based IGE measures leads us to consider an alternative approach motivated by well founded principles in the literature on inequality, poverty, and cross-sectional mobility. Our goal is to highlight the otherwise implicit role of inequality aversion in all statistical measures, and to embed such consideration in measuring IGEs. Two social welfare preferences are used here: the first is based on the Gini evaluation functions (as in Yitzhaki (1996)), and the second one the Lorenz evaluation functions. Both functions are widely used in the analyses of social welfare and inequality and characterized by one parameter representing the degree of inequality aversion (see Aaberge et al. (2021)). We show how these two measures vary with degrees of inequality aversion and showcases the flexibility of our approach to incorporate a wide range of possible weighting schemes consistent with the underlying evaluation functions.

4.1 Mobility Measures based on Gini Evaluation Functions

The Gini evaluation function is given by

$$P^I_\kappa(u) = 1 - (1-u)^\kappa \text{ for } u \in [0,1], \tag{10}$$

where $\kappa > 1$ is the inequality-aversion parameter; $u = F_X(x)$ is the income position or rank. The higher κ, the more inequality-averse a society. On the one extreme case $\kappa = 1$, society is indifferent to inequality; on the other extreme case $\kappa \to \infty$, society cares most about the welfare of the poor. The derivative of the preference function, $dP^I_\kappa(u)/du = \kappa(1-u)^{\kappa-1}$, reflects the weight placed on a particular income position u in the definition of welfare functions (e.g., Weymark (1981), Yaari (1987, 1988), Aaberge (2000)).

Gini-based Mobility Measure, $s^I(\kappa)$. The first class of mobility measure based on Gini evaluation functions (10) is defined as

$$s^I(\kappa) = \frac{Cov(Y, [1-F_X(X)]^{\kappa-1})}{Cov(X, [1-F_X(X)]^{\kappa-1})}, \tag{11}$$

where the denominator is the extended Gini variability index, and the numerator is the extended Gini covariance. Note that the constant of the derivative of the preference function (κ) is dropped in the measure since it appears in both denominator and numerator. Below we present a systematic analysis of the weighting scheme behind our measure.

Underlying Weights of $\mathbf{s^I(\kappa)}$: $\mathbf{w^I_\kappa(u)}$. Yitzhaki (1996) establishes that $s^I(\kappa)$ can be expressed as a weighted average of the marginal effect $g'(x)$ given by

$$s^I(\kappa) = \int_{\mathcal{S}_X} w^I_\kappa(x) g'(x) dx, \tag{12}$$

where the underlying weight scheme satisfies $w^I_\kappa(x) > 0$, $\int_{\mathcal{S}_X} w^I_\kappa(x)dx = 1$, and

$$w^I_\kappa(x) = \frac{[1 - F_X(x)] - [1 - F_X(x)]^\kappa}{\int_{\mathcal{S}_X} \{[1 - F_X(t)] - [1 - F_X(t)]^\kappa\} dt}. \tag{13}$$

Moreover, $w^I_\kappa(\cdot)$ can be rewritten as a function of $u = F_X(x)$,

$$w^I_\kappa(u) = c^I(\kappa) \left[(1 - u) - (1 - u)^\kappa\right], \tag{14}$$

where $c^I(\kappa) = \left(\int_0^1 \{(1-u) - (1-u)^\kappa\} dF_X^{-1}(u)\right)^{-1} > 0$ is a positive constant, depending on $F_X(\cdot)$ and κ. As shown below, the expression of $w^I_\kappa(\cdot)$ in terms of u is convenient for analyzing its properties.

Properties of $\mathbf{w^I_\kappa(u)}$. The first- and second-order derivatives of $w^I_\kappa(u)$ are given, respectively, by

$$\frac{dw^I_\kappa(u)}{du} = c^I(\kappa) \left[\kappa(1-u)^{\kappa-1} - 1\right] \quad \text{and} \quad \frac{d^2 w^I_\kappa(u)}{du^2} = c^I(\kappa)\kappa(1-\kappa)(1-u)^{\kappa-2} < 0. \tag{15}$$

The first-order condition implies that the maximizer (the turning point) of $w^I_\kappa(u)$ is given by $u^I(\kappa) = 1 - \kappa^{\frac{1}{1-\kappa}}$. The second-order condition implies that $w^I_\kappa(u)$ is strictly concave in u. As the upper left panel of Figure 1 illustrates, the weight increases for lower values of u, reaches a maximum, and then declines. The key properties of the weighting scheme are summarized in Proposition 2.

Proposition 2. *We obtain the following properties:*
(i) $u^I(\kappa)$ *is strictly decreasing in* κ *for* $\kappa > 1$.
(ii) $\lim_{\kappa\to\infty} u^I(\kappa) \to 0$.
(iii) For a sufficiently large κ, $w^I_\kappa(u) \approx c^I(\kappa)(1-u)$ *for* $u \in [u^I(\kappa), 1]$.

Property (i) states that the location of maximum weight (turning point) decreases in the inequality aversion parameter. This property is consistent with the typical policies goals or preferences that are averse to inequality. As a society or policymaker becomes more inequality averse, the individuals from the more disadvantaged families should receive the maximum weight. The pattern is illustrated in the upper right panel of Figure 1. Property (*ii*) states that when the inequality aversion tends to infinity, the largest weight is indeed placed on the poorest individuals. More importantly, the relative weight for the poor is larger than that for the rich when we increase the inequality aversion. This later feature is evident in the upper left panel of Figure 1 and formally stated in property (*iii*). For a sufficiently large inequality aversion parameter κ, the weighting scheme can be approximated by a downward slopping line for almost all the values of u. This result can be of practical importance as well, since this type of weighting schemes is usually consistent with what many empirical researchers have in mind. Moreover, it suggests that the researchers simply need to assign the inequality aversion parameter a relatively large value to obtain such weighting scheme.

4.2 Mobility Measures based on Lorenz Evaluation Functions

Our second class of mobility measures is based on the Lorenz family of evaluation functions, defined as

$$P_{\nu}^{II}(u) = (\nu u - u^{\nu})(\nu - 1)^{-1} \text{ for } u \in [0,1], \nu > 0, \text{ and } \nu \neq 1 \tag{16}$$

where ν is the preference parameter capturing the extent of inequality aversion (see, e.g., Aaberge (2000)). Note that the inequality aversion *increases* with parameter κ for the Gini measures, while it *decreases* with parameter ν for the Lorenz measures.

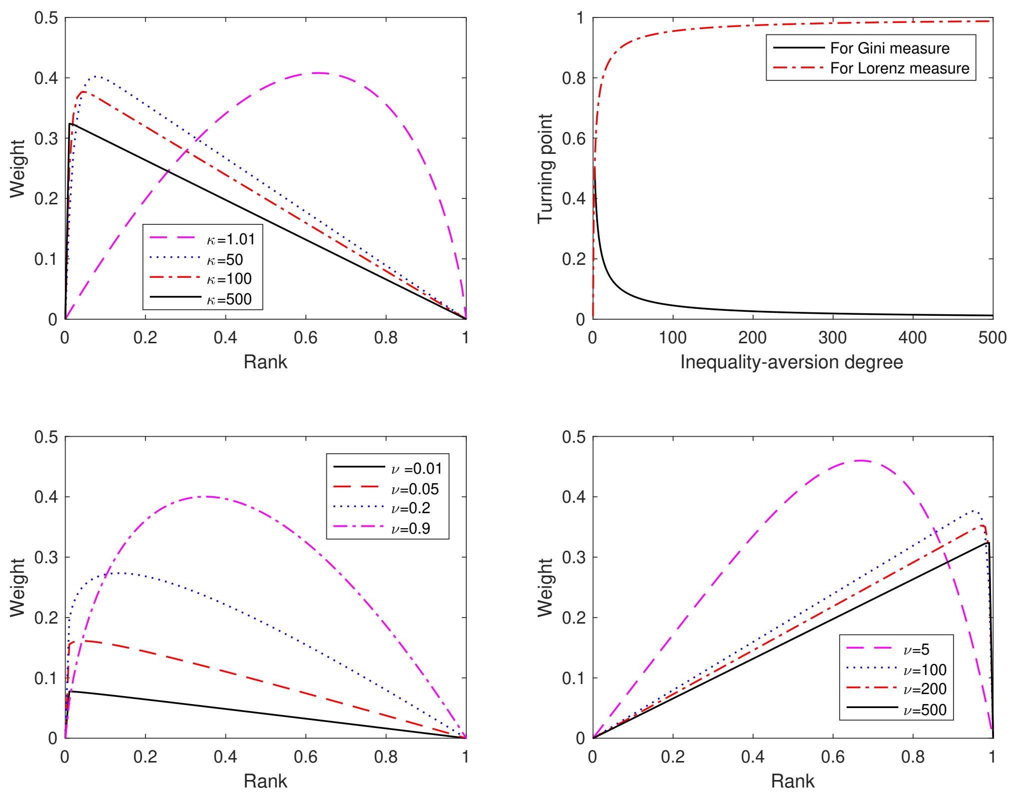
Lorenz-based Mobility Measure, $s^{II}(\nu)$. Similar to the first class of mobility measures, based on the first derivative of the Lorenz evaluation function $dP_{\nu}^{II}(u)/du = \nu(1 - u^{\nu-1})(\nu - 1)^{-1}$, the second class of mobility measures is defined as

$$s^{II}(\nu) = \frac{Cov(Y, 1 - F_X(X)^{\nu-1})}{Cov(X, 1 - F_X(X)^{\nu-1})}, \tag{17}$$

where the constant of the derivative of the preference function $(\nu(\nu - 1)^{-1})$ is dropped in the measure since it appears in both denominator and numerator.

Underlying Weights of $\mathbf{s^{II}}(\nu) : \mathbf{w_{\nu}^{II}(u)}$. Using the arguments similar to (12), the second measure $s^{II}(\nu)$ can also be expressed as a weighted average of the marginal effect $g'(x)$.

Figure 1: Shapes of weight functions and turning points for different inequality-aversions



Note: The lower and upper left panesl are for $w_{\nu}^{II}(u)$ with $0 < \nu < 1$ and for $w_{\kappa}^{I}(u)$ respectively. The lower right panel is for $w_{\nu}^{II}(u)$ with $\nu > 1$. The upper right panel is for turning points $u^{I}(\kappa)$ under Gini measure and $u^{II}(\nu)$ under Lorenz measure. The income X has standard normal distribution.

Lemma 3. *The measure of mobility in (17) can be expressed as*

$$s^{II}(\nu) = \int_{\mathcal{S}_X} w_\nu^{II}(x) g'(x) dx, \tag{18}$$

where $w_\nu^{II}(x) > 0$, $\int_{\mathcal{S}_X} w_\nu^{II}(x) dx = 1$, *and*

$$w_\nu^{II}(x) = \frac{F_X(x) - F_X(x)^\nu}{\int_{\mathcal{S}_X} [F_X(t) - F_X(t)^\nu] dt}. \tag{19}$$

Moreover, $w_\nu^{II}(\cdot)$ can also be rewritten as a function of $u = F_X(x)$,

$$w_\nu^{II}(u) = c^{II}(\nu)(u - u^\nu), \tag{20}$$

where $c^{II}(\nu) = \left(\int_0^1 (u - u^\nu) dF_X^{-1}(u) \right)^{-1}$ is a constant, which is positive for $\nu > 1$ and negative for $0 < \nu < 1$.

Properties of $\mathbf{w_\nu^{II}(u)}$. The first- and second-order derivatives of $w_\nu^{II}(u)$ are given, respectively, by

$$\frac{dw_\nu^{II}(u)}{du} = c^{II}(\nu) \left(1 - \nu u^{\nu-1} \right) \quad \text{and} \quad \frac{d^2 w_\nu^{II}(u)}{du^2} = \nu(1-\nu) c^{II}(\nu) u^{\nu-2} < 0. \tag{21}$$

The first-order condition implies that the maximizer (the turning point) of $w_\nu^{II}(u)$ is given by $u^{II}(\nu) = \nu^{\frac{1}{1-\nu}}$. The second-order condition implies that $w_\nu^{II}(u)$ is strictly concave in u. As the bottom left and right panels of Figure 1 illustrate, the weight increases for lower values of u, reaches a maximum, and then declines. The key properties of the weighting scheme are summarized in Proposition 3.

Proposition 3. *We obtain the following properties:*
(i) $u^{II}(\nu)$ *is strictly increasing in* ν *for* $\nu > 0$.
(ii) $\lim_{\upsilon \to 0} u^{II}(\nu) \to 0$ *and* $\lim_{\nu \to \infty} u^{II}(\nu) \to 1$.
(iii.a) For a sufficiently small $\nu < 1$, $w_\nu^{II}(u) \approx c^{II}(\nu)(u - 1)$ *for* $u \in [u^{II}(\nu), 1]$.
(iii.b) For a sufficiently large $\nu > 1$, $w_\nu^{II}(u) \approx c^{II}(\nu) u$ *for* $u \in [0, u^{II}(\nu)]$.

It is useful to compare Proposition 3 vs. Proposition 2 to understand the similarities and differences between the Gini- and Lorenz-based estimators and how they could complement each other in practice.

Meanings of Inequality Aversion Parameters The major difference between $s^I(k)$

(the Gini family of mobility measures) and $s^{II}(\nu)$ (the Lorenz family of mobility measures) is the weighting function. $w^I_\kappa(\cdot)$ varies with $\kappa > 1$ for the former, whereas $w^{II}_\nu(\cdot)$ varies with $\nu > 0$ for the latter. These two parameters have the opposite meaning. Moreover, the parameter space for the inequality-aversion parameter (κ vs ν) is different for the two classes of evaluation functions, as the space for ν, the parameter for the Lorenz family, covers $(0, 1) \cup (1, \infty)$. As a result, the behavior of the weighting schemes for the Lorenz measures also differ in the two sub-regions of the space.

Similarities – Economic Implications Three properties in Proposition 3 have economic meanings qualitatively similar to their counterparts in Proposition 2. First, Proposition 3 (*i*) is similar to its counterpart in Proposition 2. As illustrated by the two left panels and the upper right panel in Figure 1, when the inequality aversion is higher (ν is smaller), the turning point of the income (that is, the income location of the maximum weight) decreases. Note that the opposite patterns of the curves in the upper right panel is due to the opposite meanings of the two parameters κ and ν.

Proposition 3 (*ii*) again states that at the the limit, the maximum weight (turning point) is placed on the individuals from the poorest families. In addition, Proposition 3 (*iii.a*) is similar to Proposition 2 (*iii*): for a sufficiently large inequality aversion ($\nu \to 0$), the weights *decrease* monotonically with income levels. In general, individuals from poorer families receive more weight weight than those from richer families.

Differences Despite many qualitative similarities of the weighting schemes between the two estimators, the weights may be quite quantitatively different, which can be critical for relevant policy evaluations. Proposition 3 (*iii.b*) is unique. For a sufficiently *small* inequality aversion ($\nu \to \infty$), the weights can *increase* monotonically with income levels, which may correspond to some form of "efficiency" consideration in measuring mobility. This result suggests the flexibility of our measures in accommodating a wide range of possible weighting schemes that emphasize different policy goals and targets. .

As indicated in the upper left and bottom left panels of Figure 1, at the extreme inequality aversion ($\kappa \to \infty$ and $\nu \to 0$), the weighting schemes are both nearly linear and place the maximum weight on the children from the poorest families. However, the weights differ in the magnitudes and in the exact positions where the maximum weight is placed. For example, the weights behind the Lorenz family seem more homogenous. As we will see from the empirical examples below, these seemingly subtle differences can substantially alter our impressions of mobility and the conclusions for between-group comparisons in the presence of nonlinearity.

Finally, although the weighting schemes exhibit similar shapes when policy makers are more averse to inequality (a larger κ for the Gini measure and a smaller ν for the Lorenz measure), they behave very differently under less inequality-aversions. The lower right panel of Figure 1 implies, the weighting scheme under Lorenz measures is still approximately linear and places the maximum weight on the children from the richest families. In contrast, the upper left panel indicates, the weighting scheme under Gini measures seems concave and places the maximum weight on the children from the middle families. Combining the results of the Gini measures (when κ is large) and those of the Lorenz (when $\nu > 1$ is small), one may gain insights into the possible nature of the nonlinearity of mobility under different underlying evaluation functions. Later in our empirical analysis, we will see an example like this.

4.3 Estimation

We provide the estimation of our proposed summary measures of mobility. Suppose that the data consist of an independent sample $\{Y_i, X_i\}_{i=1}^n$ of size n. Given that these two summary measures of mobility have similar expressions, without loss of generality, we focus on the estimation of the summary measure of mobility under the Gini evaluation functions.

Since (11) takes the form of a Wald-IV estimator, we can conveniently estimate $s^I(\kappa)$ using the two steps. Let $Q_i = [1 - F_X(X_i)]^{\kappa-1}$. First, we obtain the estimate of Q_i by $\widehat{Q}_i = [1 - \widehat{F}_X(X_i)]^{\kappa-1}$, where $\widehat{F}_X(X_i) = \frac{1}{n}\sum_{j=1}^n \mathbb{I}(X_j \leq X_i)$, and $\mathbb{I}(\cdot)$ is an indicator function equal to one if the argument is met and zero otherwise. Next, we run an IV regression of Y_i on X_i, with $\widehat{Q}_i$ being the IV for X_i. The IV estimator of the coefficient on X_i is the estimator $\widehat{s}^I(\kappa)$ of $s^I(\kappa)$.

4.4 Estimation in the presence of covariates

How to flexibly account for covariates and perform counterfactual analysis in the context of IGM is not trivial and relatively underdeveloped. Here, we extend our framework to systematically incorporate covariates in estimation, which could prove useful for counterfactual analysis to understand the sources of mobility, as well as the between-group differences.

Let $Z = (Z_1, \cdots, Z_d)'$ be a d-dimensional vector of continuous variables.[10] The true

[10] For discrete covariates such as gender and race, the estimation is identical to that for the (unconditional) measures for each sub-group. This is often the case for the intergenerational mobility literature (e.g.,

income-transmission process in (1) becomes

$$Y = g(X, Z) + \epsilon, \tag{22}$$

where the error term ϵ satisfies $\mathbb{E}(\epsilon|X, Z) = 0$. In the spirit of (11), we define the summary measure of mobility under the Gini evaluation functions, conditional on $Z = z$, as

$$s^I(\kappa, z) = \frac{Cov(Y, [1 - F_{X|Z}(X|Z)]^{\kappa-1}|Z = z)}{Cov(X, [1 - F_{X|Z}(X|Z)]^{\kappa-1}|Z = z)}. \tag{23}$$

Following Yitzhaki (1996), we obtain the counterpart of (12) in the presence of covariates as

$$s^I(\kappa, z) = \int_{\mathcal{S}_{X|Z=z}} w^I_\kappa(x, z) g'(x, z) dx, \tag{24}$$

where $w^I_\kappa(x, z) > 0$, $\int_{\mathcal{S}_{X|Z=z}} w^I_\kappa(x, z) dx = 1$, $\mathcal{S}_{X|Z=z}$ is the conditional support of X given $Z = z$, and

$$w^I_\kappa(x, z) = \frac{[1 - F_{X|Z}(x|z)] - [1 - F_{X|Z}(x|z)]^k}{\int_{\mathcal{S}_{X|Z=z}} \{[1 - F_{X|Z}(t|z)] - [1 - F_{X|Z}(t|z)]^\kappa\} dt}. \tag{25}$$

Estimation Method #1: Fully Nonparametric Model: we propose the estimation of the heterogeneous mobility measures $s^I(\kappa, z)$ and then obtain the average mobility measure across heterogenous covariates. Let $K_h(\cdot)$ be a d-dimensional kernel and $h = (h_1, \cdots, h_d)$ be a sequence of bandwidths which depend on the sample size n.[11] Let $Q^Z_i = [1 - F_{X|Z}(X_i|Z_i)]^{\kappa-1}$ and $\widehat{Q}^Z_i = [1 - \widehat{F}_{X|Z}(X_i|Z_i)]^{\kappa-1}$, where

$$\widehat{F}_{X|Z}(X_i|Z_i) = \frac{n^{-1} \sum_{j=1}^n \mathbb{I}(X_j \leq X_i) K_h(Z_j - Z_i)}{n^{-1} \sum_{j=1}^n K_h(Z_j - Z_i)} \tag{26}$$

is the standard kernel estimator. A natural plug-in estimator of $Cov(Y, Q^Z|Z = z)$ is given by

$$\widehat{Cov}(Y, Q^Z|Z = z) = \frac{n^{-1} \sum_{i=1}^n K_h(Z_i - z)(Y_i - \widehat{\mathbb{E}}(Y_i|Z_i))(Q^Z_i - \widehat{\mathbb{E}}(Q^Z_i))}{n^{-1} \sum_{i=1}^n K_h(Z_i - z)}, \tag{27}$$

see the analysis of black-white differences in intergenerational mobility by Bhattacharya and Mazumder (2011)).

[11] For the Gaussinal kernel, for example, $K_h(z) = (h_1 \times \cdots \times h_d)^{-1} \left(\sqrt{2\pi}\right)^{-d} exp(-\sum_{l=1}^d \frac{(z_l/h_l)^2}{2})$.

where

$$\widehat{\mathbb{E}}(Y_i|Z_i) = \Big(\frac{1}{n}\sum_{j=1}^{n} Y_j K_h(Z_j - Z_i)\Big)\Big(\widehat{f}_Z(Z_i)\Big)^{-1}, \quad (28)$$

$$\widehat{\mathbb{E}}(Q_i^Z) = \Big(\frac{1}{n}\sum_{j=1}^{n} Q_j^Z K_h(Z_j - Z_i)\Big)\Big(\widehat{f}_Z(Z_i)\Big)^{-1}, \quad (29)$$

$$\widehat{f}_Z(Z_i) = \frac{1}{n}\sum_{j=1}^{n} K_h(Z_j - Z_i). \quad (30)$$

The estimator of $Cov(X, Q^Z|Z = z)$ is defined by replacing Y_i with X_i in (27). Based on the estimates $\{\widehat{Q}_i^Z\}_{i=1}^n$, the estimator of $s^I(\kappa, z)$ is given by

$$\widehat{s}^I(\kappa, z) = \frac{\widehat{Cov}(Y, \widehat{Q}^Z|Z = z)}{\widehat{Cov}(X, \widehat{Q}^Z|Z = z)} = \frac{\sum_{i=1}^n K_h(Z_i - z)(Y_i - \widehat{\mathbb{E}}(Y_i|Z_i))(\widehat{Q}_i^Z - \widehat{\mathbb{E}}(\widehat{Q}_i^Z))}{\sum_{i=1}^n K_h(Z_i - z)(X_i - \widehat{\mathbb{E}}(X_i|Z_i))(\widehat{Q}_i^Z - \widehat{\mathbb{E}}(\widehat{Q}_i^Z))}. \quad (31)$$

The asymptotic normality of the kernel estimator $\widehat{s}^I(\kappa, z)$ can be established by following the related literature (e.g., see Li and Racine (2007) and Yin et al. (2010)).

Estimation Method #2: A Semiparametric Approach: In practice, we can adopt a more computationally convenient and efficient way: an alternative semi-parametric estimation of the summary measure. First, to guarantee that the estimate of Q_i^Z lies between zero and one, we run the logistic regression of $\widehat{Q}_i$ on Z_i to obtain the estimate $\widetilde{Q}_i^Z = \Lambda(Z_i'\widehat{\delta})$, where $\widehat{\delta}$ is the estimator of δ in the logistic specification $\mathbb{E}(\widehat{Q}_i|Z_i) = \Lambda(Z_i'\delta)$, $\Lambda(\cdot)$ is the cumulative distribution of a logistic variable, and δ is the vector of coefficients. Second, we run the OLS regressions of Y_i, X_i, $Y_i\widetilde{Q}_i^Z$, and $X_i\widetilde{Q}_i^Z$ on Z_i to obtain the estimates $\widetilde{\mathbb{E}}(Y_i|Z_i)$, $\widetilde{\mathbb{E}}(X_i|Z_i)$, $\widetilde{\mathbb{E}}(Y_i\widetilde{Q}_i^Z|Z_i)$ and $\widetilde{\mathbb{E}}(X_i\widetilde{Q}_i^Z|Z_i)$, respectively.[12] Then, the estimator of $s^I(\kappa, z)$ is given by

$$\widetilde{s}^I(\kappa, z) = \frac{\widetilde{Cov}(Y, \widetilde{Q}^Z|Z = z)}{\widetilde{Cov}(X, \widetilde{Q}^Z|Z = z)} = \frac{\widetilde{\mathbb{E}}(Y_i\widetilde{Q}_i^Z|Z_i) - \widetilde{\mathbb{E}}(Y_i|Z_i)\widetilde{\mathbb{E}}(\widetilde{Q}_i^Z)}{\widetilde{\mathbb{E}}(X_i\widetilde{Q}_i^Z|Z_i) - \widetilde{\mathbb{E}}(X_i|Z_i)\widetilde{\mathbb{E}}(\widetilde{Q}_i^Z)}. \quad (32)$$

With the conditional estimates, one may obtain the average summary measure of

[12] We can also run these regressions on a polynomial function of Z_i that would well approximate the estimated objects (e.g., see the polynomial series estimations in distributional analysis by Firpo et al. (2009)).

mobility in the presence of covariates as follows

$$\overline{s}^I = \int_{\mathcal{S}_Z} s^I(\kappa, z) dF_Z(z), \tag{33}$$

where $\mathcal{S}_Z$ is the support of Z with the distribution function $F_Z(\cdot)$. Using the empirical distribution function $\widehat{F}_Z(z) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}(Z_i \leq z)$, the average mobility measure $\overline{s}^I$ can be estimated by

$$\widehat{\overline{s}^I} = \int_{\mathcal{S}_Z} s^I(\kappa, z) d\widehat{F}_Z(z) = n^{-1}\sum_{i=1}^{n} s^I(\kappa, Z_i). \tag{34}$$

4.5 Use of Conditional Mobility $s^I(\kappa, z)$ in Further Analyses

Once we obtain $s^I(\kappa, z)$, we are able to further consider the source(s) of the between-group differences in our mobility measures and perform counterfactual analysis, if desired. Suppose that we are interested in measuring mobility for groups A and B

$$s_g^I(\kappa) = \int_{\mathcal{S}_{Z^g}} s_g^I(\kappa, z) f_g(z) dz, \quad g = \{A, B\},$$

where $\mathcal{S}_{Z^g}$ is the support of Z for group $g = \{A, B\}$.

One may consider a hypothetical counterfactual state where group A is endowed with the same distribution of observable characteristics among group B, but the income generating process is fixed. The counterfactual mobility measure is defined as

$$s_c^I(\kappa) = \int_{\mathcal{S}_{Z^B}} s_A^I(\kappa, z) f_B(z) dz.$$

The difference between these two groups can then be decomposed into two components, process differences (due to the differences in the income-transmission process) and composition differences, as follows

$$s_A^I(\kappa) - s_B^I(\kappa) = \underbrace{[s_A^I(\kappa) - s_c^I(\kappa)]}_{\text{Composition Differences}} + \underbrace{[s_c^I(\kappa) - s_B^I(\kappa)]}_{\text{Process Differences}}.$$

The counterfactual mobility measure $s_c^I(\kappa)$ can be estimated as

$$\widehat{s}_c^I(\kappa) = n_B^{-1}\sum_{i=1}^{n_B} \widehat{s}_A^I(\kappa, Z_i^B),$$

where $\widehat{s}_A^I(\kappa, \cdot)$ is the estimator of $s_A^I(\kappa, \cdot)$ using the sample of size n_A for group A, and $\{Z_i^B\}_{i=1}^{n_B}$ is the sample of size n_B for group B.

5 Empirical Illustration

5.1 Data

Our application is based on data from the Panel Study of Income Dynamics (PSID), which includes information at the household and individual levels for a nationally representative sample of the population of the United States. The data collection began in 1968 and has since continued to update information on the individuals of the original sample and their descendants. The long panel structure allows us to match children to their parents for intergenerational analysis, as well as to obtain their incomes at a wide range of stages over the life-cycle for both generations. See Mazumder (2016, 2018b) for excellent accounts of the unique advantages of the PSID data for analysis of intergenerational mobility, and over administrative tax data. Because of these advantages, the PSID data are widely used in the literature on estimation of intergenerational mobility. Use of alternative datasets does not impact the illustrative purpose of our analysis, or the central message of the paper.

To facilitate comparison to the literature, especially those studies using the PSID, we follow closely the standard practices in the literature to construct our sample and relevant variables, and therefore provide only limited details here. Following the literature (e.g., Solon (1992); Durlauf et al. (2017)), we include only the Survey Research Center component of the PSID, but exclude the Survey of Economic Opportunity (SEO) component to prevent over-representing the poverty sample. Recent literature has also noted some serious irregularities in the sampling of SEO respondents that can "preclude easy generalization to any well-defined population" (Bloome (2015); An et al. (2020)).

In our analysis, we use (the logarithm of) permanent incomes for both children and parents. Following the literature (e.g., Durlauf et al., 2017), we define the permanent income as the average of annual family incomes, which include the taxable income of all earners in the family, from all sources, and transfer payments. We exclude zero and negative incomes. These income variables are converted to 2015 dollars using the Consumer Price Index.[13] We also follow Mazumder (2018b) to take advantage of the very long panel structure of the PSID and center the average around age 40 (between 30 and 50). The

[13] Source: `https://fred.stlouisfed.org/series/CPALTT01USA661S`

choice of age 40 follows the rule of thumb in the literature that largely overcomes the life cycle bias (Haider and Solon (2006); Mazumder (2018b)). The life-cycle bias is due to the heterogenous life cycle earnings profiles, where individuals with high lifetime income often have relatively low income when younger, and use of the incomes when they are young can then bias the estimates downward (Jenkins, 1987). We also restrict the sample to those individuals with at least three observations of annual incomes (e.g., Durlauf et al., 2017).

These standard practices also mitigate some of the known issues such as the issue of zero incomes that typically arise when using the administrative data due to non-employment. First, the family total income in the PSID includes sources of income such as transfers that are not available in the administrative tax record, and it is still reported "even when it may be too low to be filed for tax purposes" (An et al., 2020). Second, the PSID has a better coverage of life-cycles than the administrative records. Therefore, very few instances of zero incomes exist in the PSID, and the instances of multiple years of zero incomes are even rarer. Discussing these issues with the PSID for estimation of intergenerational mobility, Mazumder (2016) concludes that "the concerns about the sensitivity of results around how to handle years of zero income is effectively a non-issue when using family income." See An et al. (2020) for more details on this issue as well.

5.2 Results

5.2.1 Baseline Results

To facilitate the comparison to the literature, we first estimate the IGMs using the traditional regression-based approaches. Panel A of Table 2) reports the results. The level linear regression using the full sample yields an estimate of about 0.54 (Column (1)), consistent with the previous literature using PSID with an average of multiple years of annual incomes. The rank-rank regression, on the other hand, yields an estimate of about .39, similar to .341 reported in Chetty et al. (2014) using the federal income tax records. The estimated correlation using the rank-rank regression is substantially smaller than when using the level regression, suggesting a higher level of mobility in the U.S.. From our theoretical analysis above, the substantial difference between the two approaches is indicative of the nonlinearity of the income transmission process, and stems from the differences in their respective weighting functions. The weighting schemes for both level and rank-rank regressions are generally unknown, except for a few specific parametric distributions. Moreover, the weights are not necessarily proper weights for the rank-rank regressions.[14]

We now turn to our proposed estimators. In Panel B of Table 2, we first present the results based on the Gini evaluation function for inequality aversion $\kappa = 1.1, 2, 10, 50, 100, 500$. The summary measures of mobility vary drastically with respect to the inequality aversion parameter, and so does our impression of the mobility level in the U.S. For example, the correlation coefficient is between .5905 ($\kappa = 2$) and .2828 ($\kappa = 500$), with the difference being more than 100 percent ($\frac{(.5905-.2828)}{.2828} \times 100 \approx 109$). As a larger κ is associated with larger weights on poorer households, evidently, our results suggest a substantially more mobile society when focusing more on the individuals from the disadvantaged families.

The pattern of the changes with respect to κ is, however, not monotonic. We observe that at the relatively low level of inequality aversion, an increase in inequality aversion (from $\kappa = 1.1$ to 2) leads to a larger coefficient and hence a higher level of immobility. On the other hand, for $\kappa > 2$, when the inequality aversion parameter increases and we place more weights on the individuals from the disadvantaged families, we actually observe that the correlation coefficients between child and parental incomes decrease substantially in magnitudes, suggesting a more mobile society.

The level-regression coefficient is .5371, falling between the coefficient using $\kappa = 10$ (.5641) and that using $\kappa = 50$ (.4870). While we do not have any idea about the weighting schemes behind the rank-rank regression, its coefficient is closer to the mobility measure suggested by our method between using $\kappa = 100$ and $\kappa = 500$ (when placing more weights on the lower tail than the upper tail of the parental income distribution).

Panels C.1. and C.2. of Table 2 presents the second set of the results based on the Lorenz family of evaluation functions. Recall that the inequality aversion parameter, ν, has the opposite meaning of κ for the Gini-based measures. For $\nu \in (0, 1)$ in Panel C.1., we again find that as the inequality aversion increases (ν decreases), the coefficients decrease in magnitudes. This pattern is similar to the results using the Gini-based measures. However, the variation in the estimates is substantially smaller. The coefficients vary from .5696 to .5804. This is not surprising because Figure 1 suggests that even though the shapes of the underlying weighting functions between the Gini and Lorenz family of measures may appear to be similar, they are not the same and the actual weights also differ drastically. For example, in the case of high inequality aversion (when $\nu = .01, \kappa = 500$), they both place the maximum weight on the children from the lowest-income families,

[14] It is possible that the two variables may follow a normal distribution. To examine this possibility, Table 1 also reports the Jarque-Bera test of normality, and we reject the null hypothesis of normality for both child and parental income distributions at the one percent level.

and the weighting function seems almost linear, but the variations in the actual weights are smaller for the Lorenz family than for the Gini family. This result again highlights the importance of clarifying the policy objectives or evaluation functions (and hence the weights) in measuring and summarizing IGMs.

For $\nu > 1$ in Panel C.2. of Table 2, we find that as ν increases and the inequality aversion decreases, the magnitudes of the coefficients decrease. As we place more weights on the children from higher-income individuals, our measures suggest a more mobile society. This result is particularly interesting, especially when we view it with the result based on the Gini estimator. As Figure 1 implies, as $\kappa > 1$ increases from a very small value, the Gini-based measure starts from placing the maximum weight on the middle-income families, and gradually place more weights on the *lower*-income families. In contrast, as $\nu > 1$ increases, the maximum weight associated with the Lorenz-based measure also starts the middle-income families, but gradually place more weights on the *higher*-income families. Hence, the estimates of these two measures complement each other by highlighting the features from different parts of the distribution. Together, our results may suggest that mobility is higher at both ends than at the middle.

In sum, we reach two conclusions. First, our results provide strong evidence that the income transmission process is highly nonlinear, and suggest that both children from disadvantaged and richer families may have a higher level of mobility than those from the "middle class". The latter is only *suggestive* since we pick only a few inequality aversion parameters (hence highlighting only a few selected parts of the distribution as well) and the nonlinear pattern may not be as smooth and straightforward as the ones suggested here. Second and more importantly, due to the strong presence of nonlinearity, the subjective weighting schemes matter, in fact, a lot when forming a general impression of the mobility in a society. Bearing these results in mind, we now further examine how the potential nonlinearity and varying weighting schemes may impact our understanding of geographic differences in IGMs and that of the evolution of IGMs.

5.2.2 Geographic Disparities in Mobility

We first examine geographic differences in mobility. Following the literature, we compare four regions where an individual grew up: the Northeast, the North Central, the South, and the West. The regression-based results are displayed in Table 3. We again find that the rank-rank regression results yield substantially smaller coefficients on the parental log income for all geographic areas, suggesting a more mobile society than the level regressions. The discrepancies between the two approaches also vary drastically across regions. For example, the difference in the implied mobility can be as large as 55 percent for the West ($55 \approx (.4482 - .2898)/.2898 \times 100$).

Both the level and rank-rank regression results indicate significant geographic disparities in IGMs. A common pattern in geographic heterogeneity emerges for both approaches: the South is less mobile than the West, in line with Chetty et al. (2014).[15] Furthermore, the South is the least mobile region, while the West is the most mobile region. However, the general pattern for the entire country can differ drastically across the methods and their corresponding weighting schemes. For example, the rank-rank regression implies a substantially higher disparity in mobility between the South and the West. Relatively speaking, the rank-rank regression suggests that the South is at least 52 percent less mobile than the West ($52.86 \approx (0.4430 - .2898)/.2898 \times 100$), compared to 30 percent suggested by the level regression ($30.63 \approx (0.5855 - 0.4482)/0.4482$). The pattern for the other two regions is not definitive. While the level regression suggests that the North Central is less mobile than the Northeast, the rank-rank regression suggests the opposite.

We now turn to our proposed estimators. The results based on the Gini evaluation functions are presented in Table 4. We start with $k = 1.1$ (the maximum weight is placed roughly at the middle part of the parental income distribution). The result also suggests that the South is the least mobile region, while the West is the most mobile region, but the difference between the South and the rest of the country is also not so big as the traditional approaches suggest. The immobility level in the South is about 29 percent higher than the West ($\approx (0.5956 - 0.4611)/.4611$); the difference is slightly smaller than what is suggested by the level regression ($.3063 \approx (.5855 - .4482)/.4482$) but substantially smaller than what is suggested by the rank-rank regression ($.5286 \approx (.4430 - .2898)/.2898$).

Varying the inequality aversion parameters impacts both the size and patterns of IGM across regions. First, the variation of the estimates with respect to the inequality aversion parameters differs across regions. For example, for the Northeast, when we place more weights on the children from the disadvantaged families, the size of the mobility decreases by 80 percent when comparing the largest value (.1115, $\kappa = 500$) and the smallest value (.5481, $\kappa = 2$). By contrast, the coefficient is only 16 percent smaller for the North Central when comparing the smallest coefficient (.4941 when $\kappa = 500$) with the largest (.5906 when $\kappa = 2$). The stark contrast implies that the income transmission process and the extent of nonlinearity differ significantly across groups.

[15] Specifically, both traditional measures find that the coefficients are smaller in magnitude for the children from the West than those from the South.

Second, the patterns of the changes with respect to κ for the Northeast and North Central are consistent with the full sample (i.e., when we increase the inequality aversion, the coefficient first increases and then decreases), but they are not for the South and the West. For the South, the coefficients *decrease* with respect to the inequality aversion parameter, while for the West, the coefficients *increase*, fluctuating around an increasing trend. In contrast to the rest of the country, the coefficient increases from .4611 to .7825 for the West. When placing more and more weights on the children from the disadvantaged families, we actually find that the West becomes less and less mobile.

Our impression of the *relative* mobility levels is also senstive to the change of the inequality aversion parameter. It starts to change when $\kappa = 10$, and we observe that the West is actually the least mobile region. As we continue to increase the inequality aversion parameter, a more stable relative ranking emerges. In fact, when $\kappa = 500$ and we place more weights on the individuals from the most disadvantaged families, the Northeast is the most mobile region and the West is the least mobile region.

We turn to the results based on the Lorenz family of evaluation functions. The results are presented in Table 5. For $0 < \nu < 1$, we find similar patterns to the results based on the Gini family. We continue to find that when we increase the inequality aversion parameter, the coefficients decrease for the Northeast, the North Central, and the South, but it increases for the West. As mentioned above, the weighting schemes for the Lorenz family, while placing more weights on the lower tail of the distribution, are less steep than the ones for the Gini family. It is not surprising that the coefficients vary with respect to the inequality aversion parameter, but less drastically. However, even a minor change in the coefficients impacts the relative ranking of geographic differences in mobility: when we place more weights on the children from the disadvantaged families (when $\nu = .1, .2, .5,$), the West is the most mobile, while the North Central is the least mobile.

For $\nu > 1$, another interesting pattern arises. As we increase ν and place more weights on the individuals from richer families, the coefficients decrease for the Northeast, the North Central, and the West, which is similar to the full-sample results. For the South, we instead observe the opposite. When we increase ν from 1.1 (the maximum weight roughly at the middle-income families) to 500 (the maximum weight roughly at the richest families), the coefficient increases from .6013 to .7690, suggesting a much less mobile region.

The results here highlight the fact that there exists significant heterogeneity in the income transmission processes both within and across regions, and evidently that, the weighting schemes play an important role in forming both our impression of the mobility for a region as well as the impression of the regional disparities in mobility. Moreover, viewing the results based on the Gini and the Lorenz families together, there is some suggestive evidence that the children from richer families may have a higher level of immobility or "affluence trap" in the South, while children from both disadvantaged and richer families may have a higher level of mobility than those from the "middle class" for the rest of the country.

5.2.3 Dynamics of Mobility

To examine how the mobility evolves across cohorts, we consider four cohorts (those born before 1954, between 1955 and 1961, between 1962 and 1967, after 1968). The regression-based results are displayed in Table 6. We continue to find that the rank-rank regression results suggest a more mobile society than the level regression results. All the coefficients are much smaller when using the rank-rank regressions than when using the level regressions. The discrepancies in the mobility levels implied between the traditional approaches can be as large as 60 percent for the cohorts born between 1962 and 1967 ($60.3 = (.5226 - .326)/.3260 \times 100$).

Both the level and rank-rank regression results imply an increase in the magnitudes of the correlation coefficients and a decrease in mobility over time when comparing the (first) cohort born before 1954 and the (last) cohort born after 1968. Specifically, both methods suggest about 13 percent decrease in mobility (level regression: $\frac{(.5597-.4949)}{.4949} \times \approx 13$; rank-rank regression $\frac{(.4197-.3718)}{.3718} \times 100 \approx 13$). However, the dynamics and the magnitudes of the changes between the first and the last cohorts differ across the methods used. For example, the level regressions suggest an increasing trend between cohorts; we find that relative to the cohort born before 1954, the coefficient is larger for the cohort born during the period 1962-1967. The rank-rank regression suggests the opposite.

We now turn to our proposed estimators. The results based on the Gini evaluation functions are reported in Table 7. Varying the inequality aversion parameter again can drastically revise our view of mobility for a particular cohort, as well as that of the dynamics of the mobility across cohorts. First, the coefficients do not monotonically vary with the inequality aversion parameter (κ), and the patterns differ drastically across cohorts. For the cohorts born before 1954, between 1955-1961, and after 1968, we observe that the coefficients first increase and then decrease when we place more and more weights on the children from the low-income families. They peak at different inequality aversion parameters (for the cohort born before 1954, the largest coefficient is .5747 when $\kappa = 50$, while for the cohort born after 1968, the largest coefficient is .6344 when $\kappa = 2$). On the other hand, for the cohort born between 1962 and 1967, we actually find that the coefficients first decrease and then increase, reaching the maximum when the maximum weights are placed on the children from the lowest-income families with $\kappa = 500$.

Second, our relative ranking of the mobility levels across cohorts also depend crucially on the part of the distribution which a particular weighting scheme emphasizes. For example, when $\kappa = 1.1$, we observe the same ranking as the level regression, where it is more mobile for the cohort born before 1954 than for the cohort born after 1968. Such impression is reversed when $\kappa = 50$ (.5747 (Before 1954) vs .4751 (After 1968)).

Turning to our results based on the Lorenz family in Table 8, we continue to find the importance of the weighting schemes. We observe a far less variation in the coefficients when we vary the inequality aversion parameter for the Lorenz-based measures. More importantly, we also observe a reversed trend, compared to the one implied by Gini-based measures. For example, among those born between 1962 and 1967, when we increase the level of inequality aversion (from $\nu = .8$ to $\nu = .1$), we observe monotonically decreasing coefficients, which suggests a more mobile society. By contrast, for the same cohort, when we increase the level of inequality aversion (from $\kappa = 1.1$ to $\kappa = 500$) for the Gini-based measures, we observe the coefficients fluctuate.

These results imply that we could have a highly nonlinear income transmission process that fluctuates a lot (and hence the derivative and the implied mobility level) at adjacent values and can be drastically different.

5.2.4 Before Conclusions

The results here, both theoretical and empirical, may be uncomfortable for some. The ubiquitous heterogeneity and nonlinearity may imply that any conclusions regarding the mobility can be subjective and sensitive to the varying parameter. That is true. However, some of the *qualitative* conclusions do not have to be. It is important to see what consensus may arise from this kind of analysis. When no uniform conclusions can be reached, our paper points out the need to explicate the commitment to certain policy goals when measuring mobility. For example, many may agree that the measurement of mobility should reflect our care for the poor, and that monitoring the changes or policy effectiveness should place more weights on the children from more disadvantaged families. Only when such qualifying statements are made can our policy discussions be more meaningful and fruitful.

6 Conclusions

In this paper, we consider the fundamental decision-theoretic foundation of the intergenerational mobility measures in the presence of nonlinearity and heterogeneity. We first recast the dominant regression-based measures as a weighted average of intergenerational income elasticities at different parts of the parental income distribution. Our careful analysis of the weighting schemes underlying these traditional measures exposes the undesirable features of these approaches to measuring intergenerational mobility for any society or group. The weights for the rank-rank regression are even radical and not proper weights that could be at odds with most of the policy goals or evaluation functions that we would typically consider reasonable in practice. These results prompt us to provide a unified framework to embed policy goals or evaluation functions in our estimation of IGMs. Two one-parameter families of summary measures of mobility are proposed, and their properties are thoroughly examined. These two sets of mobility measures are flexible and can accommodate a wide range of the weighting schemes that we typically desire. The estimation of these measures is also easy to implement and allows practitioners to focus on the issues at hand. For the sake of completeness, we also extend our framework to further permit inclusion of additional covariates, which can be useful for counterfactual analysis and understanding the sources of the IGMs.

We apply our method to the PSID to estimate the intergenerational mobility in the U.S.. Our central message is loud and clear: our impression of the mobility in a country or group depends crucially on which part of the distribution we would like to highlight in practice. The sensitivity of the mobility measures for a particular group also has a substantial impact on our between-group comparisons of mobility levels. Our perspectives on the geographic differences in mobility and the dynamics of mobility can be challenged, depending on the weighting schemes.

References

Aaberge, R., "Characterizations of Lorenz curves and income distributions," *Social Choice and Welfare* 17 (2000), 639–653.

Aaberge, R., T. Havnes and M. Mogstad, "Ranking Intersecting Distribution Functions," *forthcoming, Journal of Applied Econometrics* (2021).

An, Y., L. Wang and R. Xiao, "A Nonparametric Nonclassical Measurement Error

Approach to Estimating Intergenerational Mobility Elasticities," *Journal of Business & Economic Statistics* (2020), 1–17.

Bhattacharya, D. and B. Mazumder, "A nonparametric analysis of black–white differences in intergenerational income mobility in the United States," *Quantitative Economics* 2 (2011), 335–379.

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

Chetty, R., N. Hendren, M. R. Jones and S. R. Porter, "Race and economic opportunity in the United States: An intergenerational perspective," *The Quarterly Journal of Economics* 135 (2020), 711–783.

Chetty, R., N. Hendren, P. Kline and E. Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics* 129 (2014), 1553–1623.

De Chaisemartin, C. and X. d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review* 110 (2020), 2964–96.

De Chaisemartin, C. and X. dHaultfoeuille, "Fuzzy differences-in-differences," *The Review of Economic Studies* 85 (2018), 999–1028.

Durlauf, S., A. Kourtellos and C.-M. Tan, "Status Trap," *Journal of Business and Economic Statistics* 35 (2017).

Firpo, S., N. M. Fortin and T. Lemieux, "Unconditional quantile regressions," *Econometrica* 77 (2009), 953–973.

Haider, S. and G. Solon, "Life-cycle variation in the association between current and lifetime earnings," *American Economic Review* 96 (2006), 1308–1320.

Hertz, T., *Unequal Chances: Family Background and Economic Success*, chapter Rags, Riches and Race: The Intergenerational Economic Mobility of Black and White Families in the United States (Princeton, NJ.: Princeton University Press, 2005), 338–358.

Jenkins, S., "Snapshots versus movies:?Lifecycle biases? and the estimation of intergenerational earnings inheritance," *European Economic Review* 31 (1987), 1149–1158.

Krueger, A. B., "The rise and consequences of inequality in the United States," *Speech at the Center for American Progress* 12 (2012).

Landersø, R. and J. J. Heckman, "The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US," *The Scandinavian journal of economics* 119 (2017), 178–230.

Li, Q. and J. S. Racine, *Nonparametric econometrics: theory and practice* (Princeton University Press, 2007).

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

Maasoumi, E. and L. Wang, "The gender gap between earnings distributions," *Journal of Political Economy* 127 (2019), 2438–2504.

Mazumder, B., "Estimating the intergenerational elasticity and rank association in the United States: Overcoming the current limitations of tax data," *Research in Labor Economics* 43 (2016), 83–129.

———, "Intergenerational mobility in the United States: What we have learned from the PSID," *The Annals of the American Academy of Political and Social Science* 680 (2018a), 213–234.

———, "Intergenerational mobility in the United States: What we have learned from the PSID," *The ANNALS of the American Academy of Political and Social Science* 680 (2018b), 213–234.

Mogstad, M. and M. Wiswall, "Testing the quantity–quality model of fertility: Estimation using unrestricted family size models," *Quantitative Economics* 7 (2016), 157–192.

Sen, A., *Inequality reexamined* (Oxford University Press, 1992).

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

———, "Cross-country differences in intergenerational earnings mobility," *Journal of Economic Perspectives* 16 (2002), 59–66.

Weymark, J. A., "Generalized Gini inequality indices," *Mathematical Social Sciences* 1 (1981), 409–430.

Yaari, M. E., "The dual theory of choice under risk," *Econometrica: Journal of the Econometric Society* (1987), 95–115.

———, "A controversial proposal concerning inequality measurement," *Journal of Economic Theory* 44 (1988), 381–397.

Yin, J., Z. Geng, R. Li and H. Wang, "Nonparametric covariance model," *Statistica Sinica* 20 (2010), 469.

Yitzhaki, S., "On using linear regressions in welfare economics," *Journal of Business & Economic Statistics* 14 (1996), 478–486.

Table 1: Jarque-Bera test of Normality

Panel A: Child Income

Full Sample (1)	**Northeast** (2)	**North Central** (3)	**South** (4)	**West** (5)
80.54 (0.000)	17.03 (0.000)	30.05 (0.000)	14.16 (0.001)	24.54 (0.000)
	Before 1954	**1954-1961**	**1961-1967**	**Post 1967**
	62.66 (0.000)	14.97 (0.001)	6.51 (0.038)	30.33 (0.000)

Panel B: Father's Income

Full Sample (1)	**Northeast** (2)	**North Central** (3)	**South** (4)	**West** (5)
396.09 (0.000)	270.05 (0.000)	235.03 (0.000)	21.03 (0.000)	30.00 (0.000)
	Before 1954	**1954-1961**	**1961-1967**	**Post 1967**
	34.41 (0.000)	39.70 (0.000)	33.22 (0.000)	267.04 (0.000)

Table 2: Measures of Immobility (Full Sample)

Panel A: Conventional Regression-based results						
	Level Regression	Rank-Rank Regression				
	0.5371***	0.3924***				
	(0.0292)	(0.0204)				
Panel B: Gini Family of Measures ($\kappa > 1$)						
	Low → Higher Inequality Aversion					
	$k = 1.1$	$k = 2$	$k = 10$	$k = 50$	$k = 100$	$k = 500$
	0.5814***	0.5905***	0.5641***	0.4870***	0.4535***	0.2828**
	(0.0311)	(0.0317)	(0.0380)	(0.0547)	(0.0668)	(0.1121)
Panel C.1. Lorenz Family of Measures ($\nu \in (0, 1)$)						
	Low → Higher Inequality Aversion					
	$v = .8$	$v = .5$	$v = .2$	$v = .1$		
	0.5804***	0.5796***	0.5736***	0.5696***		
	(0.0308)	(0.0306)	(0.0309)	(0.0311)		
Panel C.2. Lorenz Family of Measures ($\nu > 1$)						
	High → Low Inequality Aversion					
	$v = 1.1$	$v = 2$	$v = 10$	$v = 50$	$v = 100$	$v = 500$
	0.5786***	0.5680***	0.4965***	0.3648***	0.3135***	0.2700**
	(0.0312)	(0.0325)	(0.0412)	(0.0583)	(0.0699)	(0.1142)
No of obs.	2042					

Table 3: Measures of Immobility: Conventional Regression Approaches (By Region)

	By Region			
	Northeast	**North Central**	**South**	**West**
	(1)	(2)	(3)	(4)
Panel A: Regression Approach				
Log of	0.4650***	0.5193***	0.5855***	0.4482***
Father's Income	(0.0634)	(0.0497)	(0.0544)	(0.0827)
Panel B: Rank-Rank Approach				
Rank of	0.3693***	0.3636***	0.4430***	0.2898***
Father's Income	(0.0462)	(0.0335)	(0.0390)	(0.0531)
Observations	407	776	532	327

Table 4: Measures of Immobility: Gini Family (By Region)

	Parameter	**Northeast**	**North Central**	**South**	**West**
		(1)	(2)	(3)	(4)
Low	$\kappa = 1.1$	0.5306***	0.5782***	0.5956***	0.4611***
		(0.0p5em—692)	(0.0535)	(0.0561)	(0.0870)
	$\kappa = 2$	0.5481***	0.5906***	0.5758***	0.4938***
↓		(0.0711)	(0.0552)	(0.0570)	(0.0875)
	$\kappa = 10$	0.5442***	0.5750***	0.5056***	0.5649***
High		(0.0851)	(0.0686)	(0.0701)	(0.1013)
Inequality	$\kappa = 50$	0.4309***	0.5171***	0.4393***	0.4743***
		(0.1127)	(0.0993)	(0.1075)	(0.1516)
Aversion					
	$\kappa = 100$	0.3408***	0.5643***	0.3850***	0.4871**
		(0.1272)	(0.1212)	(0.1396)	(0.2005)
	$\kappa = 500$	0.1115	0.4941**	0.3892*	0.7825*
		(0.1936)	(0.1992)	(0.2358)	(0.4121)

Table 5: Measures of Immobility: Lorenz Family (By Region)

	Parameter	Northeast	North Central	South	West
		(1)	(2)	(3)	(4)
Panel A: $\nu \in (0,1)$					
High	$\nu = .1$	0.5264***	0.5779***	0.5547***	0.5019***
		(0.0689)	(0.0547)	(0.0567)	(0.0856)
↓	$\nu = .2$	0.5298***	0.5790***	0.5615***	0.4972***
Low		(0.0685)	(0.0540)	(0.0561)	(0.0851)
Inequality Aversion	$\nu = .5$	0.5329***	0.5799***	0.5783***	0.4820***
		(0.0682)	(0.0532)	(0.0555)	(0.0851)
	$\nu = .8$	0.5301***	0.5780***	0.5912***	0.4664***
		(0.0686)	(0.0532)	(0.0558)	(0.0862)
Panel B: $\nu > 1$					
High	$\nu = 1.1$	0.5249***	0.5745***	0.6013***	0.4516***
		(0.0692)	(0.0535)	(0.0564)	(0.0877)
↓	$\nu = 2$	0.5057***	0.5611***	0.6211***	0.4173***
Low		(0.0714)	(0.0550)	(0.0591)	(0.0927)
Inequality Aversion	$\nu = 10$	0.3715***	0.4980***	0.6516***	0.3308***
		(0.0870)	(0.0660)	(0.0798)	(0.1240)
	$\nu = 50$	0.2848**	0.3277***	0.7022***	0.1386
		(0.1212)	(0.0878)	(0.1319)	(0.1930)
	$\nu = 100$	0.3097**	0.2307**	0.7192***	0.0181
		(0.1464)	(0.1043)	(0.1683)	(0.2373)
	$\nu = 500$	0.3027	0.1910	0.7690***	-0.1225
		(0.2270)	(0.1630)	(0.2644)	(0.3894)
Observations		407	776	532	327

Table 6: Measures of Immobility: Conventional Regression Approaches (By Cohort)

	By Birth Cohort			
	Before 1954	**1954-1961**	**1961-1967**	**Post 1967**
	(1)	(2)	(3)	(4)
Panel A: Regression Approach				
Log of Father's Income	0.4949***	0.5303***	0.5226***	0.5597***
	(0.0754)	(0.0602)	(0.0789)	(0.0422)
Panel B: Rank-Rank Approach				
Rank of Father's Income	0.3718***	0.3965***	0.3260***	0.4197***
	(0.0507)	(0.0448)	(0.0496)	(0.0300)
Observations	337	422	365	918

Table 7: Measures of Immobility: Gini Family (By Cohort)

	Parameter	Northeast	North Central	South	West
		(1)	(2)	(3)	(4)
Low	$\kappa = 1.1$	0.5054***	0.5706***	0.5505***	0.6183***
		(0.0798)	(0.0638)	(0.0829)	(0.0452)
↓	$\kappa = 2$	0.5230***	0.5875***	0.5150***	0.6344***
		(0.0808)	(0.0652)	(0.0859)	(0.0459)
	$\kappa = 10$	0.5691***	0.5916***	0.3741***	0.5914***
High		(0.0963)	(0.0784)	(0.1121)	(0.0541)
Inequality	$\kappa = 50$	0.5747***	0.5171***	0.2493	0.4751***
		(0.1437)	(0.1149)	(0.1758)	(0.0757)
Aversion					
	$\kappa = 100$	0.5436***	0.4492***	0.3184	0.4074***
		(0.1822)	(0.1435)	(0.2195)	(0.0911)
	$\kappa = 500$	0.2716	0.3696*	0.9267*	0.2254
		(0.2824)	(0.2012)	(0.4758)	(0.1488)
	Observations	337	422	365	918

Table 8: Measures of Immobility: Lorenz Family (By Cohort)

	Parameter	Northeast	North Central	South	West
		(1)	(2)	(3)	(4)
Panel A: $\nu \in (0,1)$					
High	$\nu = .1$	0.5308***	0.5766***	0.4795***	0.6018***
		(0.0796)	(0.0642)	(0.0866)	(0.0447)
↓	$\nu = .2$	0.5274***	0.5770***	0.4904***	0.6077***
Low		(0.0789)	(0.0637)	(0.0853)	(0.0445)
Inequality Aversion	$\nu = .5$	0.5174***	0.5752***	0.5194***	0.6165***
		(0.0785)	(0.0631)	(0.0831)	(0.0444)
	$\nu = .8$	0.5081***	0.5708***	0.5424***	0.6172***
		(0.0791)	(0.0633)	(0.0825)	(0.0448)
Panel B: $\nu > 1$					
High	$\nu = 1.1$	0.5001***	0.5652***	0.5601***	0.6138***
		(0.0801)	(0.0640)	(0.0827)	(0.0454)
↓	$\nu = 2$	0.4817***	0.5466***	0.5934***	0.5955***
Low		(0.0840)	(0.0665)	(0.0848)	(0.0475)
Inequality Aversion	$\nu = 10$	0.4106***	0.4284***	0.6183***	0.5084***
		(0.1079)	(0.0834)	(0.1034)	(0.0609)
	$\nu = 50$	0.3327**	0.3009**	0.4690***	0.3769***
		(0.1518)	(0.1191)	(0.1432)	(0.0868)
	$\nu = 100$	0.2947	0.2672*	0.4616***	0.2971***
		(0.1807)	(0.1443)	(0.1701)	(0.1042)
	$\nu = 500$	0.2447	0.2151	0.5004*	0.2488
		(0.2546)	(0.1979)	(0.2757)	(0.1749)
Observations		337	422	365	918

Appendix

A. Proof of Corollary 1

Proof. First, we consider the case of normal distributions. Since the conditional distribution of ϵ given $X = x$ is normal with zero mean and variance σ_ϵ^2, the conditional distribution of Y given $X = x$ is normal with mean $\mu_{y|x} = g(x)$ and variance $\sigma_{y|x}^2 = \sigma_\epsilon^2$. Given that Y is normally distributed, the mean and variance of Y are $\mu_y = \mathbb{E}(g(X))$ and $\sigma_y^2 = \sigma_\epsilon^2 + Var(g(X))$, respectively. Note that $F_Y(y) = Pr(Y \leq y) = Pr((Y-\mu_y)/\sigma_y \leq (y-\mu_y)/\sigma_y) = \Phi((y-\mu_y)/\sigma_y)$. Then, it is easy to show that $g_r'(u)$ in (3) becomes

$$\begin{aligned}
g_r'(u) &= \frac{d}{du}\{\mathbb{E}(F_Y(Y)|F_X(X) = F_X(x) = u)\} = \frac{d}{du}\left\{\int_{-\infty}^{\infty}\Phi((y-\mu_y)/\sigma_y)f(y|x)dy\right\} \\
&= \frac{d}{du}\left\{\int_{-\infty}^{\infty}\Phi\left(\frac{y-\mu_y}{\sigma_y}\right)\frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}dy\right\} \\
&= \int_{-\infty}^{\infty}\Phi\left(\frac{y-\mu_y}{\sigma_y}\right)\frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}\frac{y-g(x)}{\sigma_\epsilon^2}\frac{g'(x)}{f_X(x)}dy \\
&= -\frac{g'(x)}{f_X(x)}\int_{-\infty}^{\infty}\Phi\left(\frac{y-\mu_y}{\sigma_y}\right)\frac{d}{dy}\left\{\frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}\right\} \\
&= \frac{g'(x)}{f_X(x)}\frac{1}{\sigma_y}\int_{-\infty}^{\infty}\phi\left(\frac{y-\mu_y}{\sigma_y}\right)\left\{\frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}\right\}dy \\
&= \frac{g'(x)}{f_X(x)}\frac{1}{\sigma_y}\int_{-\infty}^{\infty}\left\{\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}\right\}\left\{\frac{1}{\sigma_\epsilon\sqrt{2\pi}}e^{-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}\right\}dy \\
&= \frac{g'(x)}{f_X(x)}\frac{1}{\sigma_y}\frac{1}{\sigma_\epsilon\sqrt{2\pi}}\int_{-\infty}^{\infty}\left\{\frac{1}{\sqrt{2\pi}}e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}-\frac{(y-g(x))^2}{2\sigma_\epsilon^2}}\right\}dy \\
&= \frac{g'(x)}{f_X(x)}\frac{1}{\sqrt{2\pi(\sigma_\epsilon^2+\sigma_y^2)}} \times exp\left\{-\frac{(\sigma_\epsilon^2+\sigma_y^2)(\sigma_\epsilon^2\mu_y^2+g^2(x)\sigma_y^2)-(\mu_y\sigma_\epsilon^2+g(x)\sigma_y^2)^2}{2\sigma_y^2\sigma_\epsilon^2(\sigma_\epsilon^2+\sigma_y^2)}\right\} \\
&= a(x)g'(x)(f_X(x))^{-1},
\end{aligned} \tag{35}$$

where

$$a(x) = \frac{1}{\sqrt{2\pi(\sigma_\epsilon^2+\sigma_y^2)}} \times exp\left\{-\frac{(\sigma_\epsilon^2+\sigma_y^2)(\sigma_\epsilon^2\mu_y^2+g^2(x)\sigma_y^2)-(\mu_y\sigma_\epsilon^2+g(x)\sigma_y^2)^2}{2\sigma_y^2\sigma_\epsilon^2(\sigma_\epsilon^2+\sigma_y^2)}\right\}. \tag{36}$$

Therefore, $w_r(x)$ in (6) becomes

$$w_r(x) = 6(F_X(x) - F_X^2(x))g_r'(F_X(x))f_X(x)(g'(x))^{-1} = c(x)(F_X(x) - F_X^2(x)), \tag{37}$$

where $c(x) = 6a(x)$.

Second, for the case of lognormal distribution, using arguments similar to the case of normal distributions, it can be shown that $w_r(x)$ remains the same as that in (37). Specifically,

$$\begin{aligned}
g_r'(u) &= \frac{d}{du}\{\mathbb{E}(F_Y(Y)|F_X(X) = F_X(x) = u)\} = \frac{d}{du}\left\{\int_0^\infty \Phi\left(\frac{log(y) - \mu_y}{\sigma_y}\right) f(y|x)dy\right\} \\
&= \frac{d}{du}\left\{\int_0^\infty \Phi\left(\frac{log(y) - \mu_y}{\sigma_y}\right) \frac{1}{y\sigma_\epsilon\sqrt{2\pi}} exp\left\{-\frac{(ln(y) - g(x))^2}{2\sigma_\epsilon^2}\right\} dy\right\} \\
&= \int_0^\infty \Phi\left(\frac{log(y) - \mu_y}{\sigma_y}\right) \frac{d}{du}\left\{\frac{1}{y\sigma_\epsilon\sqrt{2\pi}} exp\left(-\frac{(ln(y) - g(x))^2}{2\sigma_\epsilon^2}\right)\right\} dy \\
&= \int_0^\infty \Phi\left(\frac{log(y) - \mu_y}{\sigma_y}\right) \frac{1}{y\sigma_\epsilon\sqrt{2\pi}}\left\{exp\left(-\frac{(ln(y) - g(x))^2}{2\sigma_\epsilon^2}\right)\frac{ln(y) - g(x)}{\sigma_\epsilon^2}\frac{g'(x)}{f(x)}\right\} dy \\
&= \int_0^\infty \Phi\left(\frac{log(y) - \mu_y}{\sigma_y}\right) \left\{\frac{1}{\sigma_\epsilon\sqrt{2\pi}} exp\left(-\frac{(ln(y) - g(x))^2}{2\sigma_\epsilon^2}\right)\frac{ln(y) - g(x)}{\sigma_\epsilon^2}\frac{g'(x)}{f(x)}\right\} dln(y) \\
&= \int_{-\infty}^\infty \Phi\left(\frac{t - \mu_y}{\sigma_y}\right) \left\{\frac{1}{\sigma_\epsilon\sqrt{2\pi}} exp\left(-\frac{(t - g(x))^2}{2\sigma_\epsilon^2}\right)\frac{t - g(x)}{\sigma_\epsilon^2}\frac{g'(x)}{f(x)}\right\} dt \\
&= a(x)g'(x)(f_X(x))^{-1},
\end{aligned}$$

where the last equality, which is the same as (35), is obtained by using the same arguments as those for the case of the normal distribution. Therefore, $w_r(x)$ in (6) is identical to that in (37).

Third, we consider the case of uniform distributions. Note that $\mathbb{E}(\epsilon|X) = 0$ implies $\mathbb{E}(\epsilon|F_X(X)) = 0$. Then,

$$g_r(U) = \mathbb{E}[V|U] = \mathbb{E}[F_Y(g(X) + \epsilon)|F_X(X)] = \frac{g(X) + \mathbb{E}[\epsilon|F_X(X)] - \underline{y}}{\overline{y} - \underline{y}} = \frac{g(F_X^{-1}(U)) - \underline{y}}{\overline{y} - \underline{y}}.$$

Therefore, $w_r(x)$ in (6) becomes

$$w_r(x) = 6(F_X(x) - F_X^2(x))(\overline{y} - \underline{y})^{-1}. \tag{38}$$

In addition, we provide a result connecting summary measures between level and rank regressions under additional assumptions. Let $\Delta_y = \overline{y} - \underline{y}$. If X is also uniformly distributed on $[\underline{x}, \overline{x}]$ with $\Delta_x = \overline{x} - \underline{x}$ and $g(x) = \alpha + x\beta$, we obtain

$$g_r(u) = \frac{g(u\Delta_x + \underline{x}) - \underline{y}}{\Delta_y} = \frac{\alpha + \beta(u\Delta_x + \underline{x}) - \underline{y}}{\Delta_y} = \frac{\alpha + \beta\underline{x} - \underline{y}}{\Delta_y} + \frac{\beta\Delta_x}{\Delta_y}u = \alpha_r + \beta_r u,$$

where

$$\alpha_r = (\alpha + \beta \underline{x} - \underline{y})/\Delta_y \quad \text{and} \quad \beta_r = \beta\Delta_x/\Delta_y.$$

Suppose $\Delta_x = \Delta_y$, we obtain $\beta_r = \beta$. Moreover, $w_r(x)$ in (38) becomes $w_r(x) = 6(x - \underline{x})(\overline{x} - x)\Delta_x^{-3}$, and it can be shown that $\int_{\underline{x}}^{\overline{x}} w_r(x)dx = 1$. □

B. Proof of Proposition 2

Proof. (i) $u^I(\kappa)$ is strictly decreasing in κ for $\kappa > 1$ because

$$\frac{du^I(\kappa)}{d\kappa} = -\frac{\kappa^{\frac{1}{1-\kappa}}(1 - \kappa + \kappa \log \kappa)}{\kappa(1-\kappa)^2} < 0 \text{ for } \kappa > 1. \tag{39}$$

(ii) Using L'Hopital's rule leads to

$$\lim_{\kappa\to\infty} \kappa^{\frac{1}{1-\kappa}} = \lim_{\kappa\to\infty} \exp\left(\log\left(\kappa^{\frac{1}{1-\kappa}}\right)\right) = \exp\left(\lim_{\kappa\to\infty} \log\left(\kappa^{\frac{1}{1-\kappa}}\right)\right) = \exp(0) = 1. \tag{40}$$

Therefore, $\lim_{\kappa\to\infty} u^I(\kappa) = \lim_{\kappa\to\infty}\left(1 - \kappa^{\frac{1}{1-\kappa}}\right) = 0$.

(iii) Recall the first derivative in (15) is

$$\frac{dw_\kappa^I(u)}{du} = c^I(\kappa)\kappa(1-u)^{\kappa-1} - c^I(\kappa), \tag{41}$$

where

$$c^I(\kappa) = \left(\int_0^1 \{(1-u) - (1-u)^\kappa\}\, dF_X^{-1}(u)\right)^{-1}.$$

First, the first term in (41) converges to zero as $\kappa \to \infty$ because $\kappa(1-u)^{\kappa-1}$ converges to zero for any $u \in (0, 1]$, and $c^I(\kappa)$ converges to a non-zero constant. This implies that for a sufficiently large value of κ, the approximating first derivative in (41) is a constant. Second, the rate of convergence for $\kappa(1-u)^{\kappa-1}$ seems much faster than that for the κ-related component $\int_0^1 (1-u)^\kappa dF_X^{-1}(u)$ in $c^I(\kappa)$. Therefore, for a sufficiently large value of κ, the approximate slope of $w_\kappa^I(u)$ is $-c^I(\kappa)$. As the upper left panel of Figure 1 indicates, when κ is larger, all of the curves become closer to lines, with different slopes corresponding to different values of κ. Last, combining the approximate slope $-c^I(\kappa)$ with the additional condition $w_\kappa^I(1) = 0$, we obtain that for a sufficiently large κ, the approximating expression for $w_\kappa^I(u)$ with $u \in [u^I(\kappa), 1]$ is $c^I(\kappa)(1-u)$. □

C. Proof of Proposition 3

Proof. (i) $u^{II}(\nu)$ is strictly increasing in $\nu > 0$ because it can be shown that

$$\frac{du^{II}(\nu)}{d\nu} = \frac{\nu^{\frac{1}{1-\nu}}(1-\nu+\nu log\ \nu)}{\nu(1-\nu)^2} > 0. \tag{42}$$

(ii) It is clear that $u^{II}(\nu) = \nu^{\frac{1}{1-\nu}} \to 0$ as $\nu \to 0$. Using the same proof as in (40), we obtain $u^{II}(\nu) \to 1$ as $\nu \to \infty$.

(iii) Recall the first derivative in (21) is

$$\frac{dw_\nu^{II}(u)}{du} = c^{II}(\nu) - c^{II}(\nu)\nu u^{\nu-1}, \tag{43}$$

where

$$c^{II}(\nu) = \left(\int_0^1 (u-u^\nu)dF_X^{-1}(u)\right)^{-1}.$$

First, the second term in (43) converges to zero as $\nu \to 0$ because $\nu u^{\nu-1}$ converges to zero for any $u \in (0,1]$, and $c^{II}(\nu)$ converges to a non-zero constant. This implies that for a sufficiently small ν, the approximating first derivative in (43) is a constant. Second, the rate of convergence for $\nu u^{\nu-1}$ seems much faster than that for the ν-related component $\int_0^1 u^\nu dF_X^{-1}(u)$ in $c^{II}(\nu)$, in the sense that $\nu u^{\nu-1}$ is negligible relative to $c^{II}(\nu)$. Therefore, for a sufficiently small ν, the approximate slope of $w_\nu^{II}(u)$ is $c^{II}(\nu)$. As the lower left panel of Figure 1 implies, when ν is smaller, all of the curves become closer to lines, with different slopes corresponding to different values of ν. Last, combining the approximate slope $c^{II}(\nu)$ with the additional condition $w_\nu^{II}(1) = 0$, we obtain that for a sufficiently small ν, the approximating expression for $w_\nu^{II}(u)$ with $u \in [u^{II}(\nu), 1]$ is $c^{II}(\nu)(u-1)$.

(iv) First, the second term in (43) converges to zero as $\nu \to \infty$ because $\nu u^{\nu-1}$ converges to zero for any $u \in [0,1)$, and $c^{II}(\nu)$ converges to a non-zero constant. This implies that for a sufficiently large ν, the approximating first derivative in (43) is a constant. Second, similar to the proof of property (iii), when ν is larger, $\nu u^{\nu-1}$ is negligible relative to $c^{II}(\nu)$. Therefore, for a sufficiently large ν, the approximate slope of $w_\nu^{II}(u)$ is $c^{II}(\nu)$. As the lower right panel of Figure 1 shows, when ν is larger, all of the curves become closer to lines, with different slopes corresponding to different values of ν. Last, combining the approximate slope $c^{II}(\nu)$ with the additional condition $w_\nu^{II}(0) = 0$, we obtain that for a sufficiently large ν, the approximating expression for $w_\nu^{II}(u)$ with $u \in [0, u^{II}(\nu)]$ is $c^{II}(\nu)u$. □