WPS8123

Policy Research Working Paper 8123

Approximating Income Distribution Dynamics Using Aggregate Data

Aart Kraay
Roy Van der Weide

WORLD BANK GROUP
Development Research Group
Macroeconomics and Growth Team
June 2017

Policy Research Working Paper 8123

Abstract

This paper proposes a methodology to approximate individual income distribution dynamics using only time series data on aggregate moments of the income distribution. Under the assumption that individual incomes follow a lognormal autoregressive process, this paper shows that the evolution over time of the mean and standard deviation of log income across individuals provides sufficient information to place upper and lower bounds on the degree of mobility in the income distribution. The paper demonstrates that these bounds are reasonably informative, using the U.S. Panel Study of Income Dynamics where the panel structure of the data allows us to compare measures of mobility directly estimated from the micro data with approximations based only on aggregate data. Bounds on mobility are estimated for a large cross-section of countries, using data on aggregate moments of the income distribution available in the World Wealth and Income Database and the World Bank's PovcalNet database. The estimated bounds on mobility imply that conventional anonymous growth rates of the bottom 40 percent (top 10 percent) that do not account for mobility substantially understate (overstate) the expected growth performance of those initially in the bottom 40 percent (top 10 percent).

This paper is a product of the Macroeconomics and Growth Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at akraay@worldbank.org and rvanderweide@worldbank.org.

Approximating Income Distribution Dynamics Using Aggregate Data

Aart Kraay (World Bank)

Roy Van der Weide (World Bank)

Keywords: mobility, inequality, top incomes, poverty

JEL Codes: D3, I3

1818 H Street NW, Washington DC 20433, akraay@worldbank.org, rvanderweide@worldbank.org. We are grateful to Luis Servén for helpful comments. The views expressed here are the authors' and do not reflect those of the World Bank, its Executive Directors, or the countries they represent.

1. Introduction

Understanding the extent of mobility – defined here as changes in individuals' relative incomes – is crucial to interpreting movements over time in average incomes in different parts of the income distribution. For example, the policy implications, and even the political acceptability, of a given change in average income in the top 1 percent of the income distribution depends crucially on whether the identity of those in the top income group is stable over time, or instead whether some of the initially rich fall out of the top group and are replaced with those who were initially poorer. This distinction also matters at the lower end of the income distribution. Consider for example the World Bank's declared goal of promoting "shared prosperity", defined as growth in the bottom 40 percent of a country's income distribution.[1] This emphasis on the bottom 40 percent sits well with the World Bank's focus on the conditions of the poorest in a country. However, to the extent that the role of such a goal is to evaluate the efficacy of interventions targeted towards the initially-poor members of the bottom 40 percent, it is necessary to know whether membership in this group is stable over time, or whether the initially-poor beneficiaries of these interventions become richer and move out of the bottom 40 percent and are replaced by those who started out richer and became relatively poorer.

The ability to document and analyze income distribution dynamics at the individual or household level is severely constrained by the scarcity of panel data sets that track individual or household incomes over time.[2] In advanced economies such as the United States, long and high-quality panel data on household incomes is available in the Panel Study on Income Dynamics (PSID). However, such high-quality long panel data sets in developing countries are rare. Instead, it is much more common for survey data to be available in the form of repeated cross-sections, with different samples of individuals appearing in each round of the survey. In cases where researchers have access to record-level survey data, it is possible to deploy a variety of pseudo-panel estimation techniques to obtain approximations to income distribution dynamics at the individual level.[3] In many cases, however, access to record-level data from

[1] See for example Basu (2013).

[2] The different datasets we work with in this paper have different units of observation at the micro level, including individuals, households, and tax units. They also differ in whether they measure consumption or income. For terminological convenience we will refer to *income* distribution dynamics at the *individual* level wherever it is possible to do so without confusion.

[3] See for example Dang, Lanjouw, Luoto and McKenzie (2014) for a recent application of such techniques to transitions into and out of poverty. Another approach is to move away from measures of income mobility and instead focus on measures of occupational and/or educational mobility, based on comparison of occupations and/or education levels of parents and children, using individual surveys in which information on the occupations and/or education of multiple generations in the same family can be obtained at a single point in time. See for example Long and Ferrie (2013) for a comparison of the US and Great Britain, and Sinha (2017) for evidence from India.

the survey is restricted, making it difficult to deploy pseudo-panel techniques that rely on micro data from repeated cross sections. In yet other cases, particular for older datasets such as historical income tax records, the micro data may simply no longer exist, and therefore cannot be analyzed to measure mobility directly. In these cases, researchers may only have access to summary statistics on the distribution of income such as the mean and the Gini coefficient, or alternatively grouped data on the proportion of individuals in different income brackets, but not the record-level data from the entire micro dataset itself.

This paper proposes and implements a method for approximating individual-level income distribution dynamics in these situations where only aggregate summary statistics on the distribution of income are available. Specifically, we consider a situation in which the researcher only has access to time series data on the mean of the income distribution, as well as some measure of dispersion of income across individuals, such as the Gini coefficient or the income share of a particular subgroup. We assume that the underlying surveys are sufficiently comparable that the observed aggregate moments characterize the same population across the successive surveys.

Our key identifying assumption is that individual incomes follow an autoregressive lognormal process with individual fixed effects. Under this data generating process, the joint distribution of individual incomes in any two periods is fully characterized by the mean and standard deviation of log incomes in the two periods (which can be inferred from the observed aggregate moments of the income distribution), and a mobility parameter characterizing the covariance of individual incomes between the two periods (that we cannot observe directly in the aggregate data). This mobility parameter depends on two underlying parameters of the data generating process: the autoregressive coefficient on log income, and the cross-sectional variance of the individual effects. We show that the limited information embodied in the evolution over time of the observed aggregate moments is sufficient to allow us to estimate the autoregressive coefficient on log income as well as to place bounds on the variance of the individual effects. This in turn permits us to place bounds on the extent of mobility in the income distribution, even though we do not observe income dynamics at the individual level.

In order to assess the usefulness of these bounds, we use record-level data from the PSID that permits us to directly measure mobility using micro-level panel data. We then discard all of the information in the record-level data, and use only the two time-series on the mean and standard deviation of log income to retrieve bounds on mobility based on aggregate data. We show that these bounds contain the point estimates obtained using the record-level data, and moreover are reasonably tight. Encouraged by these findings, we apply our methodology to two large cross-country datasets. The first

of these is the World Wealth and Income Database (WID), constructed by Thomas Piketty and his collaborators[4], which reports summary statistics on average income and top income shares compiled from tabulated data on tax records. From this database, we retrieve time-series data on mean income and the bottom 90% income share for 19 countries with reasonably complete annual time series data during the post-World War II period. For the median country in this sample, we have 60 years of annual data on aggregate moments on which to base our estimates of mobility. The second data source is the compendium of household survey data for a large set of mostly developing countries in the World Bank's PovcalNet database. This dataset provides summary statistics on the distribution of income (or consumption, depending on the survey) for over 1400 household surveys in a large cross-section of developing countries, in some cases extending back to the 1980s. We retrieve data on mean income (or consumption) and the Gini coefficient for these surveys, and restrict attention to the set of 28 countries for which we have at least 10 surveys. The median country in this sample has 16 household surveys spanning a period of 17 years.

We apply our methodology to find bounds on mobility for the countries in these two datasets. Our estimates confirm that among the high income countries in the WID database, the US ranks as the country with the lowest income persistence, while countries with high income persistence include all of the Scandinavian countries. Interestingly, when we correlate our estimates with GDP per capita, we find that lower income persistence is associated with higher income. This observation applies both to the WID and the PovcalNet estimates. As we discuss in more detail below, this observation is not an obvious consequence of our assumptions, and suggests that countries with greater persistence must have other characteristics conducive to more rapid growth.

We use these estimated bounds on mobility to generate bounds on the difference between "anonymous" growth rates of group average incomes (that do not require membership in the group to be stable over time) with corresponding "non-anonymous" growth rates (that track the performance of the same initial group of individuals over time). When we compare estimates of anonymous income growth rates among the bottom 40% (the World Bank's measure of Shared Prosperity) to their non-anonymous counterparts, the difference is substantial: on average, non-anonymous growth rates exceed anonymous ones by four percentage points in the WID and three percentage points in PovcalNet in annual terms. This means that by tracking shared prosperity anonymously, policy makers could inadvertently overlook the

[4] The full team behind the WID database includes Facundo Alvaredo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman.

successes with which originally poor households have been able to increase their incomes. This ordering is reversed when we examine growth of average incomes at the top of the income distribution – anonymous top incomes grow faster than non-anonymous top incomes do. Also here the magnitude of the differences we observe is considerable, at over 10 percentage points in the WID and 8 percentage points in PovcalNet. This suggest that anonymous top income growth rates, while a good indicator of changes in inequality in society, are not well-suited for estimating the expected income success of those who occupy the top of the income distribution at any given point in time.

Our paper provides an alternative approach to the large literature on estimating income distribution dynamics using pseudo-panel data techniques that track the evolution of cohorts over time.[5] This literature circumvents the need for true panel data that tracks individuals over time, but it does assume that the researcher has access to microdata from repeated series of cross-sectional household surveys. In contrast, our approach does not require access to any record level data, but instead shows how the parameters governing individual income distribution dynamics can be retrieved from time series data on aggregate moments only.[6] The obvious advantage of this is that by reducing the data requirements to a minimum, it significantly expands the number of countries where this approach can be applied. It also makes it possible to obtain estimates for historic periods for which aggregate moments have been preserved, but record level data will be hard if not impossible to come by. The price we pay for working with time series rather than pseudo-panel data is that we have to work with a much smaller number of observations. We make the most of the data that we have by relying on finite sample estimation methods (Andrews, 1993), and optimizing the trade-off between bias and precision. Finally, while we are by no means the first to notice that anonymous and non-anonymous growth rates can in principle diverge widely in the presence of mobility, we are – to our knowledge – the first to be able to provide estimates of the gap between the two for a large cross-section of countries.[7]

The rest of this paper proceeds as follows. In Section 2 we state our main assumptions regarding the lognormal data generating process for individual incomes, and show how mobility is fully

[5] See for example Deaton (1985), Moffitt (1993), Collado (1997), Verbeek and Vella (2005), Antman and McKenzie (2007), Inoue (2008), and Dang, Lanjouw, Luoto and McKenzie (2014).

[6] In this respect, our work is similar in spirit to Caselli and Ventura (2000), study a series of growth models with heterogenous agents and show that under fairly general heterogeneity, (1) the aggregate economy behaves as if aggregate variables represent the decisions of a representative agent; and (2) the evolution of the aggregate variables characterizing the decisions of the representative are informative about the individual-level dynamics of income, consumption, and wealth.

[7] See Jenkins and Van Kerm (2006), Grimm (2007), Van Kerm (2009), and Bourguignon (2011) for discussions of the difference between anonymous growth incidence curves and their non-anonymous counterparts.

characterized by the observed aggregate moments of the income distribution, together with a single key unobserved mobility parameter. In Section 3, we show how to obtain bounds on this mobility parameter using only information embedded in the evolution over time of the aggregate moments of the income distribution. Section 4 validates our methodology in the PSID data, and Section 5 provides estimates of mobility for a large cross-section of countries based only on the aggregate summary statistics on the distribution of income available in the WID and PovcalNet databases. In Section 6 we apply the lognormality assumption and the estimates of mobility from the previous sections to document the differences between anonymous and non-anonymous growth rates of group average incomes. Section 7 concludes.

2. Lognormal Income Distribution Dynamics

Throughout the paper, we rely on the following assumption regarding the process generating individual incomes:

Assumption A1: The logarithm of income of individual i at time t, y_{it}, is generated by the following autoregressive process:

$$y_{it} = \rho y_{it-1} + \delta_t + \lambda_i + \varepsilon_{it}, \quad (1)$$

where ρ is an autoregressive parameter satisfying $0 \leq \rho < 1$; δ_t is a common factor; the innovations λ_i and ε_{it} are independent and normally distributed with zero mean and variances σ_λ^2 and $\sigma_{\varepsilon t}^2$; and initial income is $y_{i0} = \frac{\lambda_i}{1-\rho} + \delta_0 + \varepsilon_{i0}$.

Although very simple, this data generating process gives rise to non-trivial income distribution dynamics through the interplay of two forces. On the one hand, realizations of the idiosyncratic shock ε_{it} generate changes in individuals' relative incomes over time. On the other hand, there are two sources of persistence over time in relative incomes: the autoregressive term ρy_{it-1}, and the individual effect λ_i. Overall income distribution dynamics will reflect the balance of these sources of changes and persistence in relative incomes. Note also that we allow the common factor, δ_t, and the variance of the idiosyncratic errors, $\sigma_{\varepsilon t}^2$, to vary over time, but we require the autoregressive parameter ρ and the variance of the individual effect σ_λ^2 to be stable over time.

Since the innovations in log income, as well as initial log income, are assumed to be normally distributed, log incomes are normally distributed in every period, as summarized in the following proposition:

Proposition 1: Assumption A1 implies that log individual incomes y_{it} and y_{it-1} are jointly normally distributed:

$$\begin{pmatrix} y_{it} \\ y_{it-1} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix}, \begin{pmatrix} \sigma_t^2 & \beta_{t-1}\sigma_{t-1}^2 \\ \beta_{t-1}\sigma_{t-1}^2 & \sigma_{t-1}^2 \end{pmatrix} \right), \tag{2}$$

where μ_t and σ_t^2 denote the cross-sectional mean and variance of log income at time t, and $0 \leq \beta_{t-1} \equiv \rho + \frac{\sigma_\lambda^2}{\sigma_{t-1}^2}\frac{1}{1-\rho} \leq 1$.

Proof: See Appendix A

Proposition 1 states that log incomes in any two consecutive periods are normally distributed.[8] The proof in the appendix generalizes this to the case of any two non-consecutive periods, which will be useful when applying our methodology to irregularly-spaced survey data. In addition to establishing lognormality of income at all points in time, Proposition 1 introduces a key composite parameter β_{t-1} which summarizes the comovement of individual incomes over time. This composite parameter can be interpreted as the OLS estimator of the slope of a regression of individual income on lagged individual income, i.e. $\beta_{t-1} = \frac{COV(y_{it}, y_{it-1})}{\sigma_{t-1}^2}$. This regression coefficient captures two distinct sources of persistence in individual incomes. First, higher values of the autoregressive parameter ρ naturally imply greater persistence in individual incomes. Second, higher values of $\frac{\sigma_\lambda^2}{\sigma_{t-1}^2}$ imply that a greater share of the dispersion across individuals in y_{it-1} is generated by dispersion in the individual effect, λ_i. This implies greater persistence, since by definition individuals receive the same individual effect in each period. Note that since β_{t-1} summarizes persistence in individual incomes, it is closely negatively related with mobility. In fact, one of the most common empirical summary statistics of mobility in panel datasets where

[8] See Lopez and Serven (2006) for cross-country evidence that the lognormal distribution matches well the reported data on quintile shares for a large compilation of household surveys across countries and over time. Battistin, Blundell and Lewbel (2009) focus on US microdata and show that the distribution of income is close to, but not exactly, lognormal, and that the distributions of permanent income and consumption are very close to lognormal. See also Cowell and Flachaire (2015), Section 6.3.1.2.

individual incomes are tracked over time is $1-\beta_{t-1}$. For this reason, we will refer to the composite parameter β_{t-1} as a "mobility parameter", as it is the workhorse empirical measure of relative mobility.

Let $\Phi(.)$ denote the cumulative normal distribution function, and define the quantile function $y_t(p) = \mu_t + \sigma_t\Phi^{-1}(p)$, which returns the log income level associated with each percentile p of the income distribution. Define the random variable $y_{t|t-1}(p)$ as log income at time t of an individual who was at the p^{th} percentile of the income distribution at time $t-1$. The probability distribution of $y_{t|t-1}(p)$ fully characterizes the mobility prospects of an individual who starts out at the p^{th} percentile of the income distribution, and is summarized in Proposition 2.

Proposition 2: *Given Assumption A1, $y_{t|t-1}(p)$ is distributed normally with the following mean and variance:*

$$\mu_{t|t-1}(p) \equiv E\left[y_{t|t-1}(p)\right] = \mu_t + \beta_{t-1}\sigma_{t-1}\Phi^{-1}(p) \tag{3}$$

$$\sigma^2_{t|t-1} \equiv E\left[\left(y_{t|t-1}(p) - \mu_{t|t-1}(p)\right)^2\right] = \sigma_t^2 - \beta_{t-1}^2\sigma_{t-1}^2, \tag{4}$$

Proof: *See Appendix A*.

Proposition 2, which follows immediately from applying the properties of conditional mean and variance of the bivariate normal distribution to Equation (2), shows a key feature of the lognormal data generating process in Assumption A1 -- the distribution of $y_{t|t-1}(p)$ depends only on aggregate moments (the mean and variance of log income), and a single mobility parameter β_{t-1}. In the following section, we show that it is possible to retrieve empirical bounds on β_{t-1} using only information in the evolution over time of the aggregate moments of the income distribution. This in turn implies that we can empirically characterize income mobility using only data on aggregate moments of the income distribution, i.e. using only the time series data on μ_t and σ_t^2.

Proposition 2 also helps to clarify the relative mobility interpretation of β_{t-1}. To see this, subtract $y_{t-1}(p) = \mu_{t-1} + \sigma_{t-1}\Phi^{-1}(p)$ from both sides of Equation (3) to obtain:

$$\mu_{t|t-1}(p) - y_{t-1}(p) = (\mu_t - \mu_{t-1}) + (\beta_{t-1} - 1)\sigma_{t-1}\Phi^{-1}(p) \tag{5}$$

This expression decomposes the expected change in log income of an individual starting at the p^{th} percentile of the income distribution at time $t-1$ into two components. The first term $\mu_t - \mu_{t-1}$

corresponds to the change in the mean of log income, which by definition contributes equally to everyone's income growth, and thus leaves relative incomes unchanged. This term can be thought of as capturing absolute mobility. The second term $(\beta_{t-1}-1)\sigma_{t-1}\Phi^{-1}(p)$ corresponds to the expected change in relative income, and thus captures relative mobility. In the benchmark case of $\beta_{t-1}=1$, the expected change in relative income is zero, i.e. there is no relative mobility in expectation. The lower is β_{t-1}, the greater is relative mobility since the expected change in relative incomes becomes larger in absolute value. In particular, when $\beta_{t-1}<1$, individuals starting out in the bottom half of the income distribution at time $t-1$ will expect to see faster-than-average income growth (since when $p<0.5$ we have $\Phi^{-1}(p)<0$ and $(\beta_{t-1}-1)\sigma_{t-1}\Phi^{-1}(p)>0$). That is, the initially poor (in expectation) get richer in relative terms. Conversely, an individual starting in the top half of the income distribution with $p>0.5$ can expect to have income growth below the average. That is, the initially rich (in expectation) get poorer in relative terms.

The mobility parameter β_{t-1} also governs uncertainty about changes in relative incomes. From Equation (4) it is clear that for a given initial dispersion in incomes σ_{t-1}^2, higher values of β_{t-1} imply less uncertainty about changes in relative incomes. The intuition for this is straightforward. Given an initial dispersion in incomes, higher values of β_{t-1} are due to higher values of the autoregressive parameter ρ, and higher values of the variance of the individual effect, σ_λ^2, both of which are sources of greater persistence in individual incomes. This in turn means a smaller role for idiosyncratic shocks to income, ε_{it}, and therefore less uncertainty about relative incomes.

3. Estimating Upper and Lower Bounds on Mobility Using Aggregate Data

A key feature of the data generating process in Assumption A1 is that it implies a simple autoregressive processes for the evolution over time of the aggregate moments of the income distribution, as summarized in the following proposition:

Proposition 3: *Assumption A1 implies that the mean and variance of log income follow:*

(6) $$\mu_t=\rho\mu_{t-1}+\delta_t$$

(7) $$\sigma_t^2=\rho^2\sigma_{t-1}^2+\left(\frac{1+\rho}{1-\rho}\right)\sigma_\lambda^2+\sigma_{\varepsilon t}^2$$

Proof: *See Appendix A*

Proposition 3 shows how the evolution over time of aggregate moments of the income distribution reflects the parameters of the underlying data generating process. In this section we show that, given time series data on the aggregate moments μ_t and σ_t^2, we can use Proposition 3 to recover estimates of the autoregressive parameter, ρ, as well as bounds on the variance of the individual effect, σ_λ^2. This in turn means that we can recover bounds on the key mobility parameter β_{t-1}, and thereby obtain approximate income distribution dynamics using only aggregate data.

We consider situations in which time series data on mean income and a summary statistic on dispersion, such as the Gini coefficient or a group income share, are available for a series of surveys of the same population. We assume that the unobserved movements over time in the common component of income, δ_t, can be approximated with a linear function of time plus an i.i.d. zero-mean error term, i.e. $\delta_t = \delta_0 + \delta_1 t + u_t$. Similarly, we assume that we can approximate the unobserved movements over time in the variance of the idiosyncratic shock to income, $\sigma_{\varepsilon t}^2$ with a linear function of time plus an i.i.d. zero-mean error term. Since the variance of the individual effect, σ_λ^2, is constant over time, this means that we can write the second and third terms in Equation (7) as $\left(\frac{1+\rho}{1-\rho}\right)\sigma_\lambda^2 + \sigma_{\varepsilon t}^2 = \eta_0 + \eta_1 t + v_t$.

Inserting these approximations into Equations (6) and (7) results in the following system of two equations:

(8) $$\mu_t = \rho\mu_{t-1} + \delta_0 + \delta_1 t + u_t$$

(9) $$\sigma_t^2 = \rho^2\sigma_{t-1}^2 + \eta_0 + \eta_1 t + v_t$$

Given time series data on the mean and variance of log income, we can obtain an estimate of ρ by simply regressing μ_t on its lag and a time trend (i.e. from Equation (8)), or we can obtain an estimate of ρ^2 by regressing σ_t^2 on its lag and a time trend (i.e. from Equation (9)). In some of our empirical applications, notably the PSID and the PovcalNet data, the available time series are quite short, rarely longer than 20 years, and frequently shorter. This has two implications for our estimation strategy. First, the short time series raises concerns about small-sample bias in the estimation of ρ. Specifically, Andrews (1993) shows that the OLS estimator of the autoregressive coefficient in a linear AR(1) process around a deterministic

trend is biased downwards in finite samples, and that this bias can be substantial. Andrews (1993) proposes a bias-corrected estimator that addresses this problem, though at the cost that the bias-corrected estimator has higher variance than the OLS estimator. We balance this tradeoff between bias and precision by taking a linear combination of the OLS estimator and the bias-corrected estimator that minimizes mean squared error. We apply this procedure to obtain an estimator of ρ from Equation (8), and an estimator of ρ^2 from Equation (9). Second, the scarcity of time series data on aggregate moments suggests that it is important to combine information from the dynamics of μ_t and σ_t^2 into a single estimate of ρ. We do this by taking a mean squared error-minimizing linear combination of the estimators based on Equation (8) and Equation (9), to arrive at a single estimate of ρ reflecting the dynamics of both the mean and the variance of log income. Further details on the estimation strategy are in Appendix B.

Given this estimate of ρ, which we denote $\hat{\rho}$, we next obtain bounds on the mobility parameter $\beta_{t-1} = \rho + \frac{\sigma_\lambda^2}{\sigma_{t-1}^2}\frac{1}{1-\rho}$. Note that β_{t-1} is an increasing function of the variance of the individual effect σ_λ^2. Thus, a lower bound on β_{t-1} can be obtained by setting $\sigma_\lambda^2 = 0$, i.e. $\beta_{t-1}^{LOWER} = \hat{\rho}$. This lower bound corresponds to the benchmark of the highest degree of mobility consistent with our estimate of $\hat{\rho}$ based on aggregate data, since we have turned off any additional persistence coming from the individual effect.

We obtain an upper bound β_{t-1}, i.e. a lower bound on mobility, by finding a corresponding upper bound on the variance of the individual effect. To do this, note that the variance of the idiosyncratic component of the error term in Equation (7) must be weakly positive, i.e. $\sigma_{\varepsilon t}^2 \geq 0$. Using Equation (7) and the estimate of ρ, this implies that $\sigma_\lambda^2 \leq \left(\frac{1-\hat{\rho}}{1+\hat{\rho}}\right)(\sigma_t^2 - \hat{\rho}^2\sigma_{t-1}^2)$. Given our assumption that ρ and σ_λ^2 are stable over the estimation sample, this upper bound must hold for every period $t \in T$, where T represents the time periods that comprise the estimation sample. This means that the tightest possible upper bound for the variance of the individual effect is $\sigma_\lambda^2 \leq \min_{t\in T}\left(\left(\frac{1-\hat{\rho}}{1+\hat{\rho}}\right)(\sigma_t^2 - \hat{\rho}^2\sigma_{t-1}^2)\right) \equiv \overline{\sigma_\lambda^2}$. Inserting this into the expression for β_{t-1}, and recalling that our data generating process implies that $\beta_{t-1} \leq 1$, we have $\beta_{t-1}^{UPPER} = \min\left(1, \hat{\rho} + \frac{\overline{\sigma_\lambda^2}}{\sigma_{t-1}^2}\frac{1}{1-\hat{\rho}}\right)$.

Note that the bounds β_{t-1}^{LOWER} and β_{t-1}^{UPPER} depend only on the aggregate moments of the income distribution, σ_t^2 and σ_{t-1}^2, as well as the estimate $\hat{\rho}$, which as discussed above can also be obtained using only aggregate moments. This means that we can obtain bounds on the mobility parameter β_{t-1} and in turn approximate individual-level income distribution dynamics using only aggregate data. Whether these

bounds are useful, in the sense of delineating a reasonably narrow range of values for β_{t-1}, is an empirical question to which we turn in the remainder of the paper.

4. Comparing Actual and Approximate Mobility in the PSID

In this section, we verify that our methodology for placing bounds on the mobility parameter β_{t-1} provides reasonably informative bounds on true mobility, using data from the US Panel Study of Income Dynamics (PSID). We first estimate bounds on β_{t-1} by applying the approach described in the previous sections to the time-series of the cross-sectional mean and standard deviation of log income in the PSID. We then take advantage of the panel structure of the PSID micro data to estimate the mobility parameter β_{t-1} directly, and compare these micro estimates with the bounds obtained using only aggregate data.

We work with annual rounds of the PSID between 1977 and 1997.[9] The unit of observation in the PSID is the household, and we take nominal family income per capita deflated by the national consumer price index as our measure of real income per capita. Since measures of inequality are sensitive to extreme observations, we discard a small number of household-year observations corresponding to implausibly low and implausibly high per capita incomes.[10] We then compute the mean and variance of log per capita income, using the sampling weights provided in the PSID, and apply our methodology for estimating ρ and obtaining bounds on β_{t-1} to the resulting two time-series of aggregate moments.

Table 1 summarizes our results. Panel A reports the estimates and standard errors of the autoregressive parameter ρ. We first report estimates based on the dynamics of the mean of log income (Equation (8)) and the variance of log income (Equation (9)). The three columns correspond to the OLS estimates, the small sample bias-corrected OLS estimates, and the MSE-minimizing linear combination of the two. Comparing the first two columns suggests that the bias correction is important – the bias corrected estimates of ρ are substantially larger than the OLS estimates, increasing from 0.65 to 1 (based in the dynamics of μ_t), and increasing from 0.76 to 0.96 (based on the dynamics of σ_t^2). However, the

[9] After 1997, the PSID switches to biannual frequency. We also considered working with a biannual version of the PSID from 1977 to the present. However, over this longer time series, there is clear evidence of a structural break in the time series for the standard deviation of log income. Accommodating this trend break in a series of only 17 biannual observations led to estimates of ρ based on Equations (8) and (9) that were highly imprecise. For this reason, we work with the shorter 1977-1997 time period with annual data. In the following subsection, we allow for structural breaks when estimating Equation (8) and (9) using longer time series on aggregate moments available in the World Wealth and Income Database.

[10] Specifically, we drop households log per capita income below $400 or above $440,000, corresponding to the bottom 0.8 percent and top 0.04 percent of the household-year observations in the raw data.

comparison also reveals that the bias-corrected estimates are much less precise than the OLS estimates. Recognizing this bias versus precision tradeoff, we calculate a MSE-minimizing combination of the OLS and bias corrected-OLS estimates, reported in the third column. The last column reports the MSE-minimizing weight on the OLS estimator, which is 1.07 for the first equation, and 0.80 for the second.[11] This results in MSE-minimizing estimates closer to the OLS estimates, at 0.63 and 0.80 in the two equations. These are our preferred estimates of ρ based on the observed dynamics of the mean and variance of log income.

Despite the short time series, these estimates are reasonably precise, with standard errors of 0.07 and 0.17 respectively.[12] However, since these combined estimates minimize MSE, they remain downward-biased. Comparing the MSE-minimizing estimates in the third column with the bias-corrected OLS estimates in the second column suggests this downward bias is non-trivial, at 0.37 for the estimate based on the dynamics of the mean of log income, and at 0.20 for the estimate based on the dynamics of the variance of log income. To further improve the precision of our estimates, we combine the two single-equation MSE-minimizing estimates into a single estimate by taking an MSE-minimizing linear combination of the two, which is reported in the last row of Table A. This results in our final estimate, which is $\hat{\rho} = 0.81$ with a standard error of 0.18.

In the bottom panel of Table 1 we report our upper-bound estimate of the variance of the individual effect given our preferred estimate of $\hat{\rho} = 0.81$, which is $\overline{\sigma_\lambda^2} = 0.021$. To put this estimate in perspective, it is useful to consider a benchmark version of Equation (7) in which the variance of the idiosyncratic shock to income is constant, i.e. $\sigma_{\varepsilon t}^2 = \sigma_\varepsilon^2$. In this case, the cross-sectional variance of log income converges to a steady-state value $\sigma^{*2} = \frac{\sigma_\lambda^2}{(1-\rho)^2} + \frac{\sigma_\varepsilon^2}{(1-\rho^2)}$, and $\frac{\sigma_\lambda^2}{(1-\rho)^2}/\sigma^{*2}$ can be interpreted as the share of steady-state inequality due to the variation in the individual effect. In our sample, the average over time of the variance of log income is 0.73, and using this as an estimate of σ^{*2} together with our preferred estimate of ρ results in an upper bound of 83 percent of steady state inequality is due to the variance of the individual effect.

[11] As discussed in Appendix B, the MSE-minimizing weights need not be between zero and one.

[12] A somewhat surprising feature of the estimates in the first row of Table 2 is that the standard error of the preferred estimator (which denotes the MSE-minimizing combination of the OLS and BC estimators) is notably smaller than the standard errors of both the OLS and the BC estimator. To interpret this, it is important to recall that we are minimizing MSE, and so the standard error of the combined estimator is not a good summary of the desirability of the estimator since it does not reflect the bias that also is present.

In the next two columns of the bottom panel of Table 1, we report the lower bound and the average over time of the upper bound on the mobility parameter, β_{t-1}, which are 0.81 and 0.97, respectively. In Figure 1, we display the time-evolution of our estimated bounds on mobility based on the aggregate moments. Recall that the lower bound is $\hat{\rho}$, which is constant over time, while the upper bound is equal to $\beta_{t-1}^{UPPER} = \min\left(1, \hat{\rho} + \frac{\overline{\sigma_\lambda^2}}{\sigma_{t-1}^2}\frac{1}{1-\hat{\rho}}\right)$ and varies over time with the observed data on σ_{t-1}^2. Purely for visual reference, the dashed line indicates the midpoint of the range between the lower and upper bound estimates.

The key question of interest is how these bounds on the mobility parameter β_{t-1} compare with actual mobility as measured at the household level. To answer this question, recall that β_{t-1} is simply the slope coefficient from an OLS regression of household-level log per capita income on its lagged value. Given the panel structure of the PSID, we can immediately retrieve a time series of estimates of β_{t-1} by regressing income on lagged income at the household level in successive rounds of the PSID. In our baseline specification, we estimate this series of OLS regressions in each period using all households with the requisite per capita income data in the current and previous period. These baseline estimates are superimposed on the macro estimates based on aggregate data in the top panel of Figure 1. These baseline micro estimates of β_{t-1} fall within the bounds estimated using the macro data for all but the last four years of the sample. Over the entire time period, the micro estimates of β_{t-1} average to 0.83, falling within the range of the average lower and upper bounds of 0.80 and 0.97. However, as is apparent from Figure 1, these baseline micro estimates fall closer to the bottom of the range based on the macro estimates.

There are two features of the PSID microdata that suggest relevant variants on these baseline micro estimates of mobility. The first is that the PSID is a rotating panel, and only around a quarter of the household-year observations correspond to households that appear in all 21 rounds between 1977 and 1997, while the median household is observed for 16 of the 21 PSID rounds. In our baseline specification, the number of households in each cross-sectional regression ranges from 5,570 to 7,587. This raises the possibility that at least some of the fluctuations over time in the micro estimates of β_{t-1} reflect changes in the composition of households in the PSID from year to year. To investigate the possible role of changing sample composition, we generate an alternative set of micro estimates of β_{t-1} by estimating the same series of cross-sectional regressions of household per capita income on lagged income, but restricting attention to the much-reduced sample of 1,637 households that appear in all 21 rounds of the PSID.

The second issue concerns the interpretation of year-to-year fluctuations in household per capita income. These fluctuations in part reflect changes in household size, as well as changes in the number of income earners in the household, which may not be well-captured by our simple lognormal data generating process. While it is possible to directly observe household size, it is not possible to cleanly identify fluctuations in household income per capita due to changes in the number of income earners in the household.[13] Beyond these concerns, there are also perennial thorny questions concerning the effects of measurement error in family income on the estimates of β_{t-1}, which could lead to different biases depending on its correlation with income and over time.[14] One crude way of partially addressing both these concerns simultaneously is to filter our estimation sample for influential observations. We do this using a Cook's distance criterion, and drop all household-year observations in our pooled baseline regression sample corresponding to the top one percent of observations on the Cook's distance statistic.

The bottom panel of Figure 1 superimposes these two alternative micro estimates of β_{t-1} on the same grey shaded region corresponding to the bounds on mobility based on aggregate data. Both variants result in slightly higher micro estimates of β_{t-1} , with both averaging to 0.86, as opposed to 0.83 in the baseline. These alternative micro estimates of β_{t-1} now fall closer to the center of the range based on the macro data, and particularly so in the middle decade of our 20-year sample. Overall, the picture that emerges from Figure 1 is that the bounds on mobility we can estimate using only data on the evolution of aggregate moments of the income distribution are reasonably narrow, and for the most part include estimates of mobility estimated from the micro panel data.

5. Cross-Country Estimates of Mobility Using Only Aggregate Data

Encouraged by the results of the previous section, which show that our approach delivers reasonable bounds on mobility in the PSID where mobility can be directly observed in the record-level panel data, in this section we obtain bounds on the mobility parameter β_{t-1} in two multi-country datasets where we have data only on aggregate moments of the income distribution. We work with two such datasets. The first is the World Wealth and Income Database (WID), which assembles estimates of top

[13] This is because the PSID collects information on family income and income of the household head, but does not collect data separately on incomes of other household members.

[14] The limited available information based on linking survey incomes with administrative data in the US suggests that the overall effect of measurement error on mobility measures is small (see Jantti and Jenkins (2015), Section 10.4.1), although these authors caution that this evidence is best interpreted as a reflection of how little is known about this issue.

income shares based on tabulations of income tax records for a large number of countries, based on the work of Anthony Atkinson, Thomas Piketty, and their collaborators.[15] We retrieve time series data on mean income and the income share of the bottom 90 or 99 percent of income earners (depending on data availability), for a set of 19 countries with long annual time series data on these aggregate moments in the post World War II period.[16] We convert these two summary statistics into the mean and variance of log income using our maintained assumption of lognormality.[17] The WID predominantly contains data from advanced economies, and our set of 19 countries from this source includes only three developing countries: China, India, and Mauritius.

Our task of estimating ρ based on the evolution over time of these aggregate moments is complicated by the fact that the time period covered for many of these countries is long, and likely spans some structural breaks. Most of the countries in our sample have data beginning shortly after World War II, and visual inspection of trends in the mean and variance of log income suggest a trend break in these series around the 1970s for many countries (see Figure 2). Three countries in our sample (United States, France, and Germany) also have fairly long time series data prior to World War II, where again it seems plausible that income dynamics may have been different relative to the post-war period. We therefore divide the data into sub-periods, and allow for different time trends by sub-period. For the three countries with pre-World War II data, we consider the available data up to 1939 as one distinct period. For the post-World War II period, we allow the data to select a single structural break in the time trends. Figure 2 illustrates this process of identifying trend breaks for different time periods for the three countries in our sample with pre-World War II data. The left and right panels of the figure report the time series of the mean and standard deviation of log income, with the trend lines for different sub-periods superimposed.

We impose the restriction that ρ is the same over the entire time-series available for each country, but we allow the intercept and time trend in Equations (8) and (9) to differ across the sub-periods for each

[15] See for example Atkinson and Piketty (2007, 2010), Atkinson et al. (2011), Banerjee and Piketty (2005), and Roine and Waldenstrom (2008).

[16] Our default is to use the bottom 90 percent share when available. We use data on bottom 99 percent for 5 countries (United Kingdom, India, Japan, Mauritius, and Singapore), for which data on bottom 90 percent is either limited or not available in the WID database.

[17] While the top incomes data in the WID is based on tabulated tax records, mean income is an estimate of income of all individuals including those who do not file tax returns, often based on national accounts measures of household income (see e.g. Atkinson et al. (2011) for details). The WID report data for income shares higher than the top 10% that we use here, and the very top income shares are based on fitting a Pareto distribution to the highest observed income groups. We use the top 10% share since it is least likely to reflect the Pareto imputation of the top tail of the income distribution, and therefore is more likely to be consistent with our lognormality assumption.

country.[18] We compute OLS and bias-corrected OLS estimates of ρ based on the evolution over time of the mean of log income, and based on the evolution over time of the variance of log income, for each country. We also compute the MSE-minimizing weighted average of the two for each equation separately, and the MSE-minimizing weighted average of these across the two equations to arrive at our preferred estimate of ρ for each country.

Our second application of this approach draws on the PovcalNet database maintained by the World Bank. This database is a large compendium of household survey data for developing countries, and is the basis for the World Bank's global poverty estimates. The database, as accessed in September 2016, contains records corresponding to 1,411 household surveys covering virtually the entire developing world, and for some countries extends back to the 1980s. For each survey, the PovcalNet database reports the mean of either per capita income or per capita consumption, depending on the welfare measure used in the survey, as well a number of summary measures of poverty and inequality. We extract time series data on mean income and the Gini coefficient for 28 countries with at least 10 household surveys. Two countries in our sample, China and Indonesia, have separate surveys for rural and urban populations, resulting in a total of 30 time-series of household surveys to which to apply our methodology.[19] Again relying on the assumption of lognormality, we convert these series into series for the mean and variance of log income, and then we implement the procedure described in Section 3 to obtain estimates of ρ and upper and lower bounds on mobility. Because the available time series in PovcalNet is much shorter than in the WID, we do not allow for trend breaks when estimating Equations (8) and (9). A visual inspection of the time-series of the mean and variance of log income for these countries shows no obvious signs of structural breaks in the data. As with the WID dataset, we generate OLS and bias-corrected OLS estimates of ρ, as well as an MSE-minimizing combination of the two, for each of the 30 surveys.

We report the results of these two empirical applications of our methodology in Tables 2 and 3. In the first column of both tables we present our preferred estimate of ρ for each country. The underlying OLS and bias-corrected OLS estimates from the equations for the evolution of the mean and variance of log income, and the corresponding MSE-minimizing weighted averages, are reported in Appendix Tables

[18] To minimize the influence of a small number of observations corresponding to large swings in the mean and variance of log income, we eliminate from our estimation sample observations for which Cook's Distance exceeds 0.2 and/or the Studentized residuals exceed 3.5.

[19] Most of the countries we selected from the PovcalNet database have annual household surveys, but a few have regularly surveys once every two or three years. We annualize our estimates of ρ and β_{t-1} for these countries to make them comparable to those based on annual data. Expressions for the irregularly-spaced versions of our main results are detailed in Appendix A.

B1 and B2, together with details on the estimation sample and the MSE-minimizing weights. The next two columns of Tables 2 and 3 report our upper-bound estimate of the variance of the individual effect, $\overline{\sigma_{\lambda}^{2}}$ and its contribution to steady-state inequality, and the final two columns report the lower bound and the average over time of the upper bound on the mobility parameter, β_{t-1}^{LOWER} and β_{t-1}^{UPPER}.

Figure 3 provides a useful visual summary of our estimates. In the top panel, we plot our preferred estimate of ρ (on the vertical axis) against log real GDP per capita (on the horizontal axis). The red circles correspond to our estimates for developing countries, while the blue squares correspond to advanced economies. The bottom panel plots the over-time average of the mid-point of our range for β_{t-1}, i.e. $\left(\beta_{t-1}^{UPPER} - \beta_{t-1}^{LOWER}\right)/2$ (on the vertical axis) against log real GDP per capita (on the horizontal axis). Our estimates of persistence ρ tend to be higher in the WID data than those based on the PovcalNet data – the mean estimate across countries of ρ is 0.86 in the WID sample, while it is only 0.63 within the PovcalNet sample. It is difficult to say whether this reflects an actual tendency for persistence to be lower in the developing countries covered in PovcalNet when compared with the largely OECD countries in the WID sample. An alternative interpretation is that our OLS estimates of ρ, as well as the MSE-minimizing combination of the OLS and bias-corrected OLS estimates, have greater small-sample downward bias in the PovcalNet sample where the available time series is much shorter than in the WID sample. However, within each group, there is a tendency for persistence to be lower in richer countries, and this correlation is significant at the 10 percent level within the WID sample of countries. The bottom panel of Figure 3, which plots the mobility parameter β_{t-1} against log real GDP per capita, looks broadly similar to the top panel, with the exception that the negative relationship with per capita income is less pronounced in the PovcalNet sample for β_{t-1} as it is for ρ in the top panel. Recall that β_{t-1} is a function of ρ and the share of the variance of the individual effect in overall inequality, $\frac{\sigma_{\lambda}^{2}}{\sigma_{t-1}^{2}}$. The similarity between the top and bottom panels of Figure 3 suggests that cross-country differences in $\frac{\sigma_{\lambda}^{2}}{\sigma_{t-1}^{2}}$ is much smaller than cross-country differences in ρ.

Figure 4 plots the evolution over time of the bounds on β_{t-1} for the United States derived from the WID data. The bounds are displayed as a grey-shaded region. The lower bound of this range is the point estimate of $\hat{\rho} = 0.66$, while the upper bound varies over time with σ_{t-1}^{2}. As in Figure 1, we also plot the mid-point of the range as a dashed line. Finally, we superimpose on this graph the mobility estimates based on the PSID micro data, for the 1997-1997 period. Although the bounds on mobility displayed in Figure 4 are based on very different data than the PSID (tabulated tax records versus a

household panel survey), it is interesting to note that the micro estimates of mobility fall mostly within the bounds based on macro data (although more towards the upper half of the range). It is also useful to interpret the movements over time in the upper bound on the mobility parameter. Recall that this upper bound is $\beta_{t-1}^{UPPER} = \hat{\rho} + \frac{\overline{\sigma_\lambda^2}}{\sigma_{t-1}^2}\frac{1}{1-\hat{\rho}}$ and moves inversely with the variance of log income, σ_{t-1}^2. For example, the decline in β_{t-1}^{UPPER} since 1980, i.e. the increase in our upper-bound estimate of mobility, is driven entirely by the increase in overall inequality during this period, i.e. the increase in σ_{t-1}^2. The rationale for this is straightforward. Given our identifying assumption that ρ and σ_λ^2 are constant over time, we interpret the increase in overall inequality as reflecting an increase over time in the variance of the idiosyncratic shock, $\sigma_{\varepsilon t}^2$. Since these shocks are independent over time, they have only transitory effects on income, and as a result mobility is higher.

6. Application: Anonymous and Non-Anonymous Growth in a Cross-Section of Countries

Popular discussions of trends in inequality frequently refer to income growth rates of "the rich" (defined as the top X% of the income distribution) and income growth rates of "the poor" (defined as the bottom Y% of the income distribution). One prominent example is the World Bank's declared goal of promoting "shared prosperity", defined as growth in the bottom 40 percent of the income distribution. Similarly, a body of work by Thomas Piketty and his collaborators has drawn widespread attention to trends in "top incomes", defined as average incomes in the upper percentiles of the income distribution.

Measuring and interpreting the income growth rates of such population subgroups is complicated by the basic data problem that motivates this paper: true panel data tracking individuals over time is rarely available. Absent such panel data, group-average growth rates typically are calculated based on repeated cross sections. For example, the growth rate of "the poor" would be calculated by comparing average incomes in the bottom Y% at two points in time. Since growth rates calculated in this way do not track individuals over time, they are referred to as "anonymous" growth rates. In contrast, "non-anonymous" growth rates track the same individuals in an initial reference group over time, and can be very different from their "anonymous" counterparts when there is mobility in the income distribution.[20]

[20] We are by no means the first to notice this distinction -- see Jenkins and Van Kerm (2006), Grimm (2007), Van Kerm (2009), and Bourguignon (2011) for discussions of the difference between anonymous growth incidence curves and their non-anonymous counterparts. The novelty in this section of our paper is that we are able to compute estimates of the difference between anonymous and non-anonymous growth rates for a large sample of countries, using our estimates of mobility based only on aggregate moments of the income distribution.

However, given the scarcity of true panel data, these non-anonymous growth rates are rarely observed directly.

This distinction can have significant policy implications as well, precisely because it reflects the underlying degree of mobility in the income distribution. For example, rapid growth in top incomes might be more politically acceptable if it were accompanied by significant mobility of individuals into and out of the top income bracket. Similarly, if the policy objective is to track the effectiveness of an intervention aimed at raising incomes in the poorest Y% of the population, conclusions could be quite different using the anonymous and non-anonymous growth rates, since the latter track the performance of the initial beneficiaries of the intervention while the former does not.

In this section we obtain analytical expressions for the difference between anonymous and non-anonymous growth rates implied by the lognormal data generating process described in Assumption A1. These expressions depend only on aggregate moments and the mobility parameter β_{t-1}. We then use our bounds on β_{t-1} from the previous section to obtain bounds on the difference between anonymous and non-anonymous growth rates in a cross-section of countries, even though the true panel data required to track individuals over time is not available for most of these countries.

We begin by defining the anonymous and non-anonymous versions of the growth incidence curve. The anonymous growth incidence curve, as defined in Ravallion and Chen (2003), returns the proportional change in income at every percentile p of the income distribution:

$$g_t^A(p) \equiv y_t(p) - y_{t-1}(p) = \mu_t - \mu_{t-1} + (\sigma_t - \sigma_{t-1})\Phi^{-1}(p) \tag{10}$$

This growth incidence curve is termed "anonymous" since the individual at the p^{th} percentile at time t will generally be different from the individual at the same percentile in the previous period $t-1$. Note that the curve is increasing in p if inequality increases, i.e. if $\sigma_t > \sigma_{t-1}$, and is decreasing in p if inequality falls. Because the anonymous growth incidence curve does not track individuals over time, a given curve could be consistent with low mobility (if the same individuals occupy the same percentiles of the income distribution at two points in time) or high mobility (if the identity of individuals at a given percentile changes over time).

In contrast, the non-anonymous growth incidence curve tracks the expected growth rate of individuals at each percentile p of the initial income distribution, i.e.

(11) $$g_t^{NA}(p) = E\left[y_{t|t-1}(p)\right] - y_{t-1}(p) \; = \mu_t - \mu_{t-1} + (\beta_{t-1} - 1)\sigma_{t-1}\Phi^{-1}(p)$$

This expression is the same as Equation (5), as is the intuition. The second term measures the extent of mobility in the income distribution. In the benchmark case of no mobility (in expectation), i.e. $\beta_{t-1} = 1$, the non-anonymous growth incidence curve is flat, and expected income growth is the same at every point in the initial income distribution. When the mobility parameter $\beta_{t-1} < 1$, individuals starting out below the mean of log income expect higher-than-average growth, while those starting out above the mean expect lower-than-average growth.

The difference between the two growth incidence curves is:

(12) $$g_t^A(p) - g_t^{NA}(p) = y_t(p) - \mu_{t|t-1}(p) = (\sigma_t - \beta_{t-1}\sigma_{t-1})\Phi^{-1}(p)$$

It is straightforward to verify that our data generating process implies $\sigma_t - \beta_{t-1}\sigma_{t-1} \geq 0$, and so the anonymous growth incidence curve will fall below the non-anonymous growth incidence curve for the bottom half of the income distribution (since $\Phi^{-1}(p) < 0$ when $p < 0.5$).[21] That is, the anonymous growth incidence curve understates the actual growth rate of any individual starting out in the bottom half of the income distribution. Symmetrically, the anonymous growth incidence curve overstates the actual growth rate of any individual starting out in the top half of the income distribution. Note that the difference between the two growth incidence curves is monotonic in β_{t-1} at every percentile p. Therefore, the upper and lower bounds for β_{t-1} obtained in the previous section will map into upper and lower bounds on the difference between the two growth incidence curves.

Figure 5 illustrates the difference between anonymous and non-anonymous growth incidence curves using the PSID microdata and our estimates of mobility based on the aggregate moments of this dataset from Section 4. The left and right panels of the graph show the 10-year average annual growth incidence curves over the periods 1977-1987 and 1987-1997. In each panel, the dashed line shows the anonymous growth incidence curve. In both periods, the anonymous growth incidence curve is weakly upward-sloping, reflecting the increase in overall inequality over this time. The two solid lines show the non-anonymous growth incidence curves based on the actual micro data, calculated in two different ways. The thin line is a kernel-weighted local polynomial smoothed average of the actual growth rates of all

[21] This is because the positive semi-definiteness of the covariance matrix in Equation (2) requires $\sigma_t^2 \geq \beta_{t-1}^2\sigma_{t-1}^2$.

individuals at each percentile of the initial income distribution. The bold line is the lognormal version of this curve, i.e. Equation (11), but using the estimate of β_{t-1} based on the PSID micro data. Interestingly, this lognormal version of the growth incidence curve tracks the non-parametric smoothed estimate of the growth incidence curve fairly closely, suggesting that our assumption of lognormality provides a reasonable approximation to the true income distribution dynamics in this particular setting. As expected, the non-anonymous growth incidence curves indicate much higher (lower) growth for those initially in the bottom (top) half of the income distribution. Finally, we plot the lower and upper bounds on the growth incidence curve based on our bounds on β_{t-1} calculated using only aggregate data. These bounds include the two versions of the actual non-anonymous growth incidence curve based on micro data almost everywhere. It is also useful to note that they exclude the anonymous growth incidence curve everywhere except for very close to the middle of the income distribution where the anonymous and non-anonymous growth incidence curves are very close to each other.

We next consider growth rates of group average incomes, such as the bottom 40 or the top 10 percent of the income distribution. The anonymous and non-anonymous growth rates of group average incomes are simply weighted averages of the corresponding anonymous and non-anonymous growth incidence curves, as summarized in the following proposition:

<u>*Proposition 4:*</u> *Let* $0 \leq p < q \leq 1$ *denote percentiles of the income distribution. The anonymous and non-anonymous growth rate of group average incomes of those between the* p^{th} *and* q^{th} *percentiles of the income distribution are:*

$$g_t^i(p,q) \approx \int_p^q g_t^i(s) w_{t-1}(s) ds\,, \qquad i = A, NA \tag{13}$$

where $w_{t-1}(s)$ *is the share of percentile s in total income of those between the* p^{th} *and* q^{th} *percentiles of the income distribution at time* $t-1$. $g_t^i(p,q)$ *may also be represented as:* $g_t^i(p,q) = g_t^i(z(p,q,\sigma_{t-1}))$, *where* $z(p,q,\sigma)$ *satisfies:* $p \leq z(p,q,\sigma) \leq q$.

<u>*Proof:*</u> *See Appendix A.*

It follows immediately from Proposition 4 and Equations (10) and (11) that the difference between the anonymous and non-anonymous growth incidence curves for group average incomes is:

$$g_t^A(p,q) - g_t^{NA}(p,q) \approx (\sigma_t - \beta_{t-1}\sigma_{t-1}) \int_p^q \Phi^{-1}(s) w_{t-1}(s) ds \tag{14}$$

That is, the difference between the two growth rates of group average incomes is also a weighted average of the difference between the anonymous and non-anonymous growth incidence curve. As before, the difference between the anonymous and non-anonymous growth incidence curves depends on $\sigma_t - \beta_{t-1}\sigma_{t-1} \geq 0$. Whether the anonymous growth rate exceeds or is smaller than the non-anonymous rate now depends on the sign of the integral $\int_p^q \Phi^{-1}(s)w_{t-1}(s)ds$. While a closed-form solution for this integral exists and is given in the proof of Proposition 4 in Appendix A, this term cannot be signed in general. However, for the particular case where $0 \leq p < q \leq 0.5$, it is clear that $\int_p^q \Phi^{-1}(s)w_{t-1}(s)ds \leq 0$ since $\Phi^{-1}(s) \leq 0$ for $0 \leq s \leq 0.5$. This means that for any subgroup in the bottom half of the income distribution, the anonymous group average growth rate is lower than the corresponding non-anonymous growth rate. Naturally, this argument is symmetric: when $0.5 \leq p < q \leq 1$, the anonymous growth rate of group average incomes for any subgroup in the top half of the income distribution exceeds the corresponding non-anonymous growth rate.

We can quantify the difference between anonymous and non-anonymous growth rates using our estimates of the mobility parameter β_{t-1} based on aggregate data from the WID and PovcalNet databases. We begin by calculating the average annual anonymous growth rate of the bottom 40 and top 10 percent over the most-recent 10-year period available for each country. These anonymous growth rates would be the ones typically reported when true panel data are not available. We then compute the lower and upper bounds on the corresponding annual average non-anonymous growth rates, using the bounds on β_{t-1} reported in Tables 2 and 3. The results of this calculation are summarized in Figure 6. The top and bottom panels report results from the WID and PovcalNet samples, respectively. In each panel, we plot the anonymous growth rate on the horizontal axis, the non-anonymous growth rate on the vertical axis, and the diagonal line corresponds to the 45-degree line. The points above the 45-degree line correspond to growth rates of the bottom 40 percent, reflecting our result that non-anonymous growth rates exceed their anonymous counterparts for group averages within the bottom half of the income distribution. Conversely, the points below the 45-degree line correspond to growth rates of the top 10 percent, for which the non-anonymous growth rate is smaller than the anonymous growth rate. The triangles (for the bottom 40 percent) and squares (for the top 10 percent) indicate the midpoint of our range for the non-anonymous growth rate, and the vertical lines display the range between the lower- and upper-bound estimates of the non-anonymous growth rates. The anonymous growth rates and the bounds on the anonymous growth rates are reported by country in Tables 4 and 5.

The most striking feature of Figure 6 and Table 4 is the magnitude of the gap between the non-anonymous and anonymous growth rates. For example, in the WID database, the mean difference between the midpoint of our estimated range for the non-anonymous growth rate, and the corresponding anonymous growth rate of the bottom 40 percent is 4.1 percent per year, while for the top 10 percent it is -10.7 percent per year. These differences are very large when compared with the standard deviation of anonymous growth rates across countries, which is only 2.8 percent per year for the bottom 40 percent, and 2.1 percent per year for the top 10 percent. Similar gaps are observed for the PovcalNet sample, where the mean difference between the non-anonymous and anonymous growth rate of the bottom 40 percent is 3.1 percent per year, and for the top 10 percent the gap is 8.1 percent per year. Equivalently, given the extent of mobility that we have inferred from the aggregate data, we can conclude that the anonymous growth rate of the bottom 40 percent understates the actual growth rate of those initially in the bottom 40 percent by 3.1 percent per year on average. And conversely, anonymous growth in average incomes of the top 10 percent overstates the average growth rate of those initially in the top 10 percent by 8.1 percent per year on average.

7. Conclusions

This paper demonstrates the feasibility of approximating individual-level income distribution dynamics when the researcher only has access to time-series data on aggregate moments of the income distribution such as the mean and the Gini coefficient or top income shares. Our key identifying assumption is that individual incomes follow an autoregressive lognormal process with individual fixed effects. We show that the limited information embodied in the time-series of these aggregate moments is sufficient to place bounds on the extent of mobility in the income distribution, even though we do not observe income dynamics at the individual level. An empirical application using data from the PSID confirms that these bounds generally contain the point estimates that are obtained using the record-level panel data, and moreover are reasonably tight. Encouraged by these findings, we apply our methodology to two large cross-country datasets, namely the WID (including mostly high income countries) and the World Bank's PovcalNet (including mostly developing countries). Some of the cross-country patterns we observe in estimates of mobility seem quite plausible given our priors. For example, among the high income countries, the Scandinavian countries and much of Europe show relatively high levels of income persistence, while the United States, Singapore and Taiwan rank among the countries with low levels of income persistence. When comparing estimates between the WID and PovcalNet, our estimates suggest

that income persistence is lower in the developing world. However, this comparison is difficult to interpret given the differences between the two databases. Specifically, a possible interpretation is that our OLS estimates of ρ, as well as the MSE-minimizing combination of the OLS and bias-corrected OLS estimates, have greater small-sample downward bias in the PovcalNet sample where the available time series is much shorter than in the WID sample.

When we correlate our estimates of income persistence with per capita GDP across countries, we find a negative correlation within the WID and PovcalNet samples – higher persistence is associated with lower incomes. The reason for this is not obvious, and it is unlikely to be mechanical. Note that if countries were identical in all parameters but the income persistence parameter, then lower persistence would be associated with lower average income (since lower persistence implies both a lower mean and lower cross-sectional variance of log income in the steady state, and therefore lower mean income given our lognormality assumption). This seems to suggest that countries with lower income persistence differ from countries with higher persistence in other ways that benefit their income growth.

Our approach to approximating individual-level distribution dynamics with aggregate data makes it possible for the first time to compare anonymous and non-anonymous growth rates in a wide cross section of countries. We show that the anonymous growth rate of the bottom 40 percent (corresponding to the World Bank's definition of "shared prosperity") understates the growth rate of those initially in the bottom 40 percent by 3 to 4 percentage points per year. This implies that the anonymous growth rate is a poor approximation to the success of those initially in the bottom of the income distribution, and is not a good tool for tracking the effects of policy interventions that were targeted to the initially-poor. The exact opposite holds true when tracking growth at the top end of the income distribution. Growth of top incomes will be larger when measured anonymously. Our estimates suggest that for the top 10% the difference is as large as 10 percentage points in annual terms. Thus, tracking anonymous top income shares substantially overstates the expected income growth rates realized by the average individual in the top end of the distribution.

Approximating income distribution dynamics, under the assumption that individual incomes follow a lognormal autoregressive process, can also be applied to an inter-generational context. The advantage of working with analytical approximations in this case is that it provides a means of rationalizing some of the empirical regularities that have been reported in recent studies on inter-generational mobility. One prominent example of such a regularity is that higher levels of inter-generational mobility are more likely to be observed in countries with lower levels of income inequality (see e.g. Corak, 2013;

Chetty et al., 2014). In an ongoing companion study, we verify analytically how growth and inequality jointly determine mobility. This in turn allows us to verify how much growth it would take to offset a rise in inequality, and use this framework to interpret the results of a recent study by Chetty et al. (2017) which reports a dramatic decline in absolute mobility in the United States over the last 60 years. In a separate study we also use our lognormal framework and empirical bounds on mobility to investigate the welfare consequences of cross-country differences in mobility, using the welfare metrics suggested by Atkinson and Bourguignon (1982) and Gottschalk and Spolaore (2002).

References

Andrews, Donald (1993). "Exactly median-unbiased estimation of first order autoregressive/unit root models". Econometrica. 61: 139-165

Antman, Francisca, and David McKenzie (2007). "Earnings mobility and measurement error: A pseudo-panel approach", Economic Development and Cultural Change. 56: 125-161

Atkinson, Anthony and Francois Bourguignon (1982). "The Comparison of Multidimensional Distributions of Economic Status". Review of Economic Studies. 49:183-201.

Atkinson, Anthony, Thomas Piketty, ed. (2007). "Top incomes over the Twentieth Century: A contrast between continental European and English speaking countries". Oxford and New York: Oxford University Press.

Atkinson, Anthony, Thomas Piketty, ed. (2010). "Top incomes: A global perspective". Oxford and New York: Oxford University Press

Atkinson, Anthony, Thomas Piketty, and Emmanuel Saez (2011). "Top incomes in the Long Run of History". Journal of Economic Literature. 49:3-71

Attanasio, Orazio, Erik Hurst, and Luigi Pistaferri (2012). "The evolution of income, consumption, and leisure inequality in the US, 1980-2010". NBER Working Paper Series, 17982

Auten, Gerald, and Geoffrey Gee (2009). "Income Mobility in the United States: New Evidence from Income Tax Data". National Tax Journal. 62:301-328

Auten, Gerald, Geoffrey Gee, and Nicholas Turner (2013). "Income inequality, mobility, and turnover at the top in the US, 1987-2010". American Economic Review: Papers & proceedings. 103: 168-172

Banerjee, Abhijit, and Thomas Piketty (2005). "Top Indian incomes, 1922-2000". World Bank Economic Review. 19: 1-20

Basu, Kaushik (2013). "Shared Prosperity and the Mitigation of Poverty". World Bank Policy Research Department Working Paper No. 6700.

Battistin, Erich, Richard Blundell, and Arthur Lewbel (2009). "Why Is Consumption More Lognormal Than Income? Gibrat's Law Revisited". Journal of Political Economy. 117(6):1140-1154.

Bourguignon, Francois (2010). "Non-Anonymous Growth Incidence Curves, Income Mobility, and Social Welfare Dominance". Journal of Economic Inequality. 9:605-627.

Caselli, Francesco and Jaume Ventura (2000). "A Representative Consumer Theory of Distribution". American Economic Review. 90(4):909-926.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez (2014). "Where is the land of opportunity? The geography of intergenerational mobility in the United States". Quarterly Journal of Economics. 129: 1553-1623

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang (2017). "The fading American dream: Trends in absolute income mobility since 1940". Science. 356: 398-406

Collado, Dolores (1997). "Estimating dynamic models from time series of independent cross-sections". Journal of Econometrics. 82: 37-62

Corak, Miles (2013). “Income inequality, equality of opportunity, and inter-generational mobility”. Journal of Economic Perspectives. 27: 79-102

Cowell, Frank and Emmanuel Flachaire (2015). “Statistical Methods for Distributional Analysis”, in Francois Bourguignon and Anthony Atkinson, eds. The Handbook of Income Distribution. Elsevier.

Dang, Hai-Anh, Peter Lanjouw, Jill Luoto, and David McKenzie (2014). “Using repeated cross-sections to explore movements in and out of poverty”. Journal of Development Economics. 107:112-128

Deaton, Angus (1985). “Panel data from time series of cross-section”. Journal of Econometrics. 30: 109-126.

Gottschalk, Peter and Enrico Spolaore (2002). “On the Evaluation of Economic Mobility”. Review of Economic Studies. 69(1):191-208.

Gottschalk, Peter, and Robert Moffitt (2009). “The Rising Instability of U.S. Earnings”. Journal of Economic Perspectives. 23: 3-24

Grimm, Michael (2007). “Removing the Anonymity Axiom in Assessing Pro-Poor Growth”. Journal of Economic Inequality. 5:179-197.

Inoue, Atsushi (2008). “Efficient estimation and inference in linear pseudo-panel data models”. Journal of Econometrics. 142: 449-466

Jantti, Markus and Stephen Jenkins (2015). “Income Mobility”, in Francois Bourguignon and Anthony Atkinson, eds. The Handbook of Income Distribution. Elsevier.

Jenkins, Stephen, and Philippe Van Kerm (2006). “Trends in income inequality, pro-poor income growth, and income mobility”. Oxford Economic Papers. 58:531-548

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song (2010). “Earnings inequality and mobility in the United States: Evidence from Social Security data since 1937”. Quarterly Journal of Economics. 125:91-128

Leigh, Andrew (2007). “How closely do top income shares track other measures of inequality?”. Economic Journal. 117:619-633

Long, Jason and Joseph Ferrie (2013). “Intergenerational Occupation Mobility in Great Britain and the United States Since 1850”. American Economic Review. 103(4): 1109-1137.

Lopez, Humberto and Luis Serven (2006). “A Normal Relationship? Poverty, Growth, and Inequality”. World Bank Policy Research Department Working Paper No. 3814.

Moffitt, Robert (1993). “Identification and estimation of dynamic models with a time series of repeated cross-sections”. Journal of Econometrics. 59:99-123

Ravallion, Martin, and Shaohua Chen (2003). “Measuring pro-poor growth”. Economics Letters. 78:93-99

Roine, Jesper, and Daniel Waldenstrom (2008). “The evolution of top incomes in an egalitarian society: Sweden, 1903-2004”, Journal of Public Economics, 92: 366-387

Sinha, Rishabh (2017). “Closer, But No Cigar: Intergenerational Mobility Across Caste Groups in India”. Manuscript, World Bank.

Van Kerm, Philippe (2009). “Income mobility profiles”. Economics Letters. 102: 93-95

Verbeek, Marno, and Francis Vella (2005). “Estimating dynamic models from repeated cross-sections”. Journal of Econometrics. 127: 83-102

Table 1: Estimates of Mobility from the PSID

Panel A: Estimates of ρ Based on Aggregate Data

	Estimate of ρ:			MSE-Minimizing
	OLS	**BC-OLS**	**Min-MSE**	**Weight**
Estimates based on:				
Equation for Mean	0.652	1.000	0.628	1.070
of Log Income	0.196	1.964	0.073	
Equation for Variance	0.760	0.956	0.802	0.802
of Log Income	0.136	0.282	0.170	
MSE-Minimizing			0.812	-0.058
Combination of Both			0.180	

Panel B: Estimates of Mobility Parameter β_{t-1}

Upper Bound on Variance of Individual Effect		Mobility Parameter β Average 1977-1997				
Estimate	*Average Share of Variance of log y*	*Lower Bound*	*Upper Bound*	*Micro Estimates Baseline*	*Balanced*	*Cook's D*
0.021	0.826	0.812	0.968	0.826	0.857	0.863

Notes: The first two rows of panel A reports estimates of ρ based on the dynamics of the mean of log income (Equation (8)) and the variance of log income (Equation (9)). OLS refers to ordinary least squares and BC-OLS refers to small-sample bias-corrected estimates using the procedure in Andrews (1993). Min-MSE refers to the mean-squared-error-minimizing linear combination of the OLS and BC-OLS estimates described in Appendix B. The final column reports the MSE-minimizing weight on the OLS estimator. The third row of panel A the MSE-minimizing linear combination of the estimates in Panel A, together with the MSE-minimizing weight on the estimator based on the dynamics of the mean of log income. Standard errors are reported below point estimates. The first two columns of Panel B report the upper bound on the variance of the individual effect (σ_λ^2), and its share in the the steady-state variance of log income. The remaining columns of Panel B report the lower bound on β_{t-1} (i.e. the preferred estimate of ρ from Panel A), the average over time of the upper bound on β_{t-1}, and the average over time of three alternative estimates of β_{t-1} based on the micro data. The "Baseline" estimates use data for all households in each year. The "Balanced Panel" estimates use only the subset of households available in all 21 years. The "Cook's-D" estimates drop households in the top 1 percent of the distribution of Cook's Distance statistic (a measure of influence) in the baseline OLS estimates.

Table 2: Estimates of Mobility from the World Wealth and Income Database

	Preferred Estimate of ρ		Upper Bound on Variance of Individual Effect		Mobility Parameter	
	Estimate	*Standard Error*	*Estimate*	*Average Share of Variance of log y*	*Lower Bound*	*Average Upper Bound*
Australia	0.756	0.090	0.019	0.685	0.756	0.928
Canada	0.808	0.075	0.025	0.704	0.808	0.944
China	0.978	0.038	0.000	0.000	0.978	0.978
Denmark	0.889	0.054	0.003	0.558	0.889	0.955
France	0.824	0.054	0.011	0.432	0.824	0.909
Germany	0.769	0.062	0.031	0.712	0.769	0.940
Great Britain	0.837	0.100	0.013	0.457	0.837	0.919
India	0.768	0.079	0.015	0.280	0.768	0.841
Ireland	0.784	0.040	0.019	0.585	0.784	0.915
Italy	0.950	0.067	0.000	0.165	0.950	0.959
Japan	0.996	0.013	0.000	0.000	0.996	0.996
Mauritius	0.880	0.051	0.000	0.000	0.880	0.880
New Zealand	0.900	0.079	0.001	0.175	0.900	0.918
Norway	0.893	0.087	0.000	0.000	0.893	0.893
Singapore	0.780	0.060	0.040	0.643	0.780	0.924
Spain	0.960	0.076	0.000	0.000	0.960	0.960
Sweden	0.929	0.055	0.000	0.000	0.929	0.929
Taiwan	0.701	0.127	0.025	0.484	0.701	0.860
United States	0.655	0.057	0.068	0.565	0.655	0.869

Notes: The first two columns report our preferred estimate of ρ and its standard error. This estimate is an MSE-minimizing weighted average of the estimates based on the evolution over time of the mean and variance of log income. The underlying estimates and MSE-minimizing weights are reported in Appendix Table B1. The third and fourth columns report the upper bound on the variance of the individual effect (σ_λ^2), and its share in the the steady-state variance of log income. The remaining columns report the lower bound on β_{t-1} (i.e. the preferred estimate of ρ from Column 1), and the average over time of the upper bound on β_{t-1}.

Table 3: Estimates of Mobility from the PovcalNet Database

	Preferred Estimate of ρ		Upper Bound on Variance of Individual Effect		Mobility Parameter	
	Estimate	*Standard Error*	*Estimate*	*Average Share of Variance of log y*	*Lower Bound*	*Average Upper Bound*
Argentina	0.581	0.178	0.107	0.733	0.581	0.896
Armenia	0.673	0.138	0.026	0.668	0.673	0.900
Belarus	0.405	0.278	0.077	0.812	0.405	0.895
Bolivia	0.171	0.316	0.521	0.681	0.171	0.764
Brazil	0.576	0.093	0.184	0.823	0.576	0.929
Chile	0.907	0.127	0.007	0.758	0.907	0.983
China (Rural)	0.680	0.152	0.035	0.769	0.680	0.917
China (Urban)	0.999	0.004	0.000	0.000	0.999	0.999
Colombia	0.485	0.218	0.278	0.859	0.485	0.931
Costa Rica	0.661	0.127	0.079	0.829	0.661	0.944
Ecuador	0.467	0.323	0.206	0.758	0.467	0.880
El Salvador	0.455	0.240	0.179	0.719	0.455	0.866
Georgia	0.745	0.080	0.033	0.878	0.745	0.969
Honduras	0.698	0.127	0.067	0.618	0.698	0.886
Indonesia (Rural)	0.531	0.124	0.040	0.704	0.531	0.876
Indonesia (Urban)	0.725	0.063	0.018	0.517	0.725	0.872
Kazakhstan	0.604	0.116	0.034	0.718	0.604	0.898
Kyrgyzstan	0.609	0.167	0.029	0.491	0.609	0.827
Moldova	0.524	0.263	0.059	0.631	0.524	0.839
Mexico	0.626	0.225	0.102	0.833	0.626	0.939
Panama	0.490	0.220	0.254	0.891	0.490	0.949
Peru	0.504	0.325	0.172	0.740	0.504	0.882
Philippines	0.969	0.066	0.000	0.471	0.969	0.984
Poland	0.932	0.040	0.001	0.638	0.932	0.976
Paraguay	0.508	0.231	0.186	0.722	0.508	0.875
Romania	0.891	0.167	0.003	0.927	0.891	0.992
Russia	0.563	0.244	0.084	0.790	0.563	0.911
Thailand	0.676	0.181	0.055	0.830	0.676	0.949
Turkey	0.677	0.077	0.050	0.857	0.677	0.955
Ukraine	0.884	0.143	0.001	0.424	0.884	0.934

Notes: The first two columns report our preferred estimate of ρ and its standard error. This estimate is an MSE-minimizing weighted average of the estimates based on the evolution over time of the mean and variance of log income. The underlying estimates and MSE-minimizing weights are reported in Appendix Table B2. The third and fourth columns report the upper bound on the variance of the individual effect (σ_λ^2), and its share in the the steady-state variance of log income. The remaining columns report the lower bound on β_{t-1} (i.e. the preferred estimate of ρ from Column 1), and the average over time of the upper bound on β_{t-1}.

Table 4: Anonymous and Non-Anonymous Growth Rates of the Bottom 40 and Top 10 Percent from the World Income and Wealth Database

	Average Annual Growth During Most Recent 10 Years of:							
	Bottom 40%				*Top 10%*			
	Anonymous	Non-Anonymous		Difference	Anonymous	Non-Anonymous		Difference
		Lower	Upper	(Using β-Mid)		Lower	Upper	(Using β-Mid
Australia	-0.001	0.059	0.026	0.043	0.021	-0.124	-0.045	-0.105
Canada	0.004	0.072	0.031	0.047	0.003	-0.185	-0.072	-0.132
China	0.098	0.108	0.108	0.009	0.079	0.053	0.053	-0.026
Denmark	0.005	0.042	0.018	0.026	0.015	-0.074	-0.016	-0.060
France	-0.001	0.049	0.017	0.034	-0.005	-0.128	-0.049	-0.084
Germany	-0.003	0.067	0.024	0.048	0.017	-0.166	-0.053	-0.127
Great Britain	0.014	0.083	0.060	0.057	0.010	-0.191	-0.124	-0.168
India	0.013	0.081	0.061	0.058	0.026	-0.154	-0.099	-0.153
Ireland	-0.023	0.041	0.009	0.048	-0.014	-0.180	-0.097	-0.124
Italy	-0.010	0.018	0.015	0.027	0.001	-0.071	-0.062	-0.067
Japan	-0.012	-0.003	-0.003	0.009	0.010	-0.014	-0.014	-0.024
Mauritius	-0.012	0.041	0.041	0.053	0.039	-0.088	-0.088	-0.127
New Zealand	0.016	0.053	0.047	0.034	0.005	-0.089	-0.073	-0.086
Norway	0.024	0.063	0.063	0.039	0.036	-0.057	-0.057	-0.093
Singapore	0.050	0.127	0.086	0.056	0.032	-0.201	-0.079	-0.172
Spain	0.008	0.026	0.026	0.018	-0.005	-0.052	-0.052	-0.046
Sweden	0.004	0.039	0.039	0.035	0.026	-0.058	-0.058	-0.085
Taiwan	0.000	0.071	0.044	0.057	0.040	-0.135	-0.069	-0.142
United States	-0.022	0.068	0.034	0.073	0.011	-0.250	-0.152	-0.212
Mean	0.008	0.058	0.039	0.041	0.018	-0.114	-0.064	-0.107
Std. Dev.	0.028	0.030	0.027	0.017	0.021	0.074	0.043	0.051

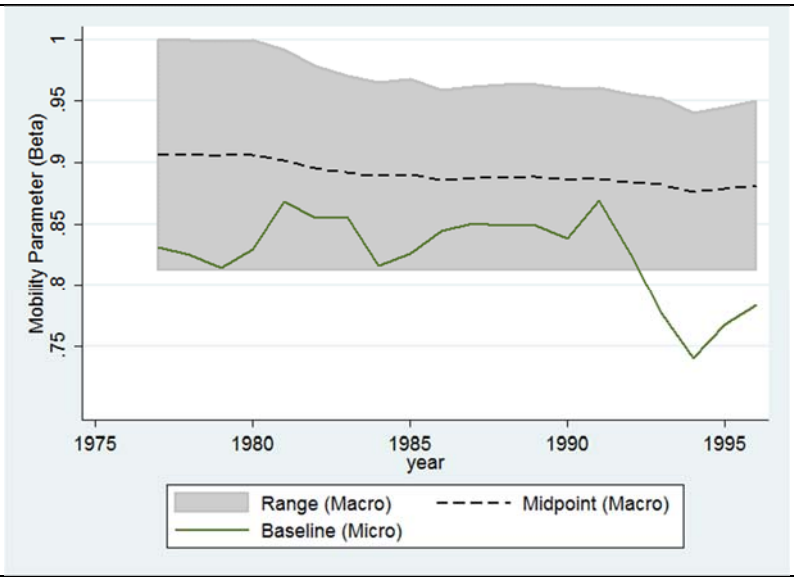
Notes: This table reports anonymous growth rates and bounds on non-anonymous growth rates for the bottom 40 percent (left panel) and top 10 percent (right panel). Growth rates are calculated as annual averages over the most recently-available ten years of data for each country. For details on the time period covered by the available data see Appendix Table B1. Within each panel, the first column reports the anonymous growth rate, and the second and third columns report the lower and upper bounds on the non-anonymous growth rate. The fourth column in each panel reports the difference between the midpoint of the bounds on the non-anonymous growth rate and the anonymous growth rate.

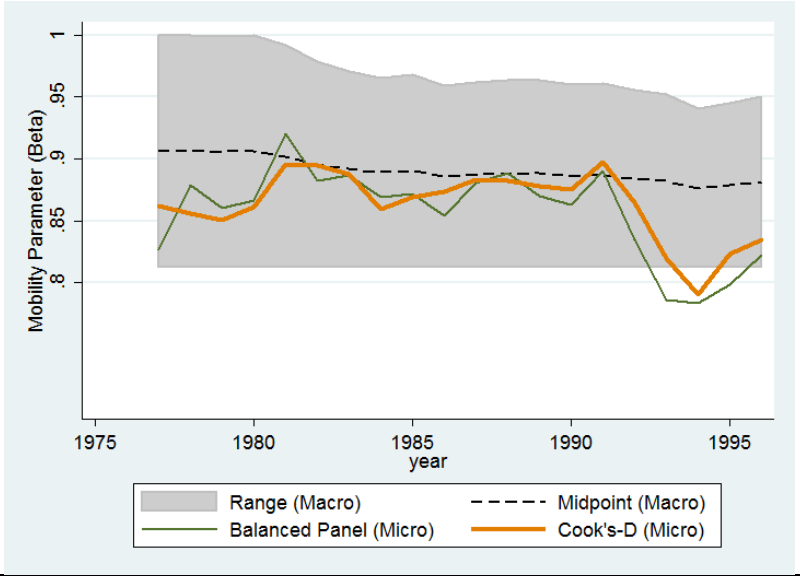
Table 5: Anonymous and Non-Anonymous Growth Rates of the Bottom 40 and Top 10 Percent from the PovcalNet Database

	Average Annual Growth During Most Recent 10 Years of:							
	Bottom 40%				*Top 10%*			
	Anonymous	Non-Anonymous		Difference	Anonymous	Non-Anonymous		Difference
		Lower	Upper	(Using β-Mid)		Lower	Upper	(Using β-Mid
Argentina	0.089	0.146	0.103	0.036	0.022	-0.135	-0.018	-0.098
Armenia	0.041	0.087	0.055	0.029	0.034	-0.072	0.003	-0.068
Belarus	0.109	0.147	0.115	0.022	0.088	0.000	0.075	-0.050
Boliva	0.082	0.147	0.099	0.041	0.021	-0.163	-0.028	-0.117
Brazil	0.066	0.138	0.075	0.040	0.033	-0.173	0.007	-0.116
Chile	0.049	0.090	0.054	0.023	0.022	-0.090	0.008	-0.064
China (Rural)	0.066	0.122	0.085	0.038	0.074	-0.063	0.027	-0.091
China (Urban)	0.063	0.066	0.066	0.003	0.073	0.065	0.065	-0.008
Colombia	0.046	0.121	0.049	0.039	0.040	-0.166	0.033	-0.107
Costa Rica	0.042	0.110	0.058	0.042	0.037	-0.144	-0.004	-0.112
Ecuador	0.068	0.132	0.084	0.040	0.021	-0.159	-0.022	-0.111
El Salvador	0.045	0.105	0.059	0.037	0.002	-0.158	-0.036	-0.100
Georgia	0.022	0.077	0.026	0.030	0.025	-0.110	0.014	-0.073
Honduras	0.027	0.098	0.055	0.049	-0.006	-0.210	-0.085	-0.142
Indonesia (Rural)	0.026	0.075	0.041	0.032	0.060	-0.048	0.026	-0.071
Indonesia (Urban)	0.013	0.075	0.046	0.048	0.062	-0.085	-0.017	-0.113
Kazakhstan	0.086	0.124	0.095	0.024	0.051	-0.038	0.028	-0.055
Kyrgyzstan	0.068	0.109	0.081	0.026	0.053	-0.038	0.025	-0.060
Moldova	0.076	0.117	0.085	0.025	0.041	-0.056	0.020	-0.059
Mexico	0.036	0.103	0.047	0.039	0.028	-0.150	0.000	-0.103
Panama	0.061	0.131	0.068	0.039	0.031	-0.169	0.010	-0.110
Peru	0.073	0.134	0.085	0.036	0.019	-0.149	-0.014	-0.100
Philippines	0.007	0.022	0.014	0.012	0.001	-0.038	-0.018	-0.029
Poland	0.027	0.050	0.034	0.015	0.019	-0.035	0.002	-0.035
Paraguay	0.064	0.130	0.079	0.040	0.019	-0.165	-0.022	-0.113
Romania	0.066	0.098	0.068	0.017	0.066	-0.005	0.062	-0.037
Russia	0.064	0.125	0.075	0.036	0.088	-0.059	0.063	-0.086
Thailand	0.053	0.107	0.057	0.029	0.038	-0.099	0.028	-0.073
Turkey	0.048	0.105	0.058	0.033	0.041	-0.099	0.017	-0.082
Ukraine	0.067	0.091	0.080	0.018	0.046	-0.008	0.018	-0.041
Mean	0.055	0.106	0.067	0.031	0.038	-0.094	0.009	-0.081
Std. Dev.	0.024	0.030	0.023	0.011	0.024	0.066	0.034	0.032

Notes: This table reports anonymous growth rates and bounds on non-anonymous growth rates for the bottom 40 percent (left panel) and top 10 percent (right panel). Growth rates are calculated as annual averages over the most recently-available ten years of data for each country. For details on the time period covered by the available data see Appendix Table B2. Within each panel, the first column reports the anonymous growth rate, and the second and third columns report the lower and upper bounds on the non-anonymous growth rate. The fourth column in each panel reports the difference between the midpoint of the bounds on the non-anonymous growth rate and the anonymous growth rate.

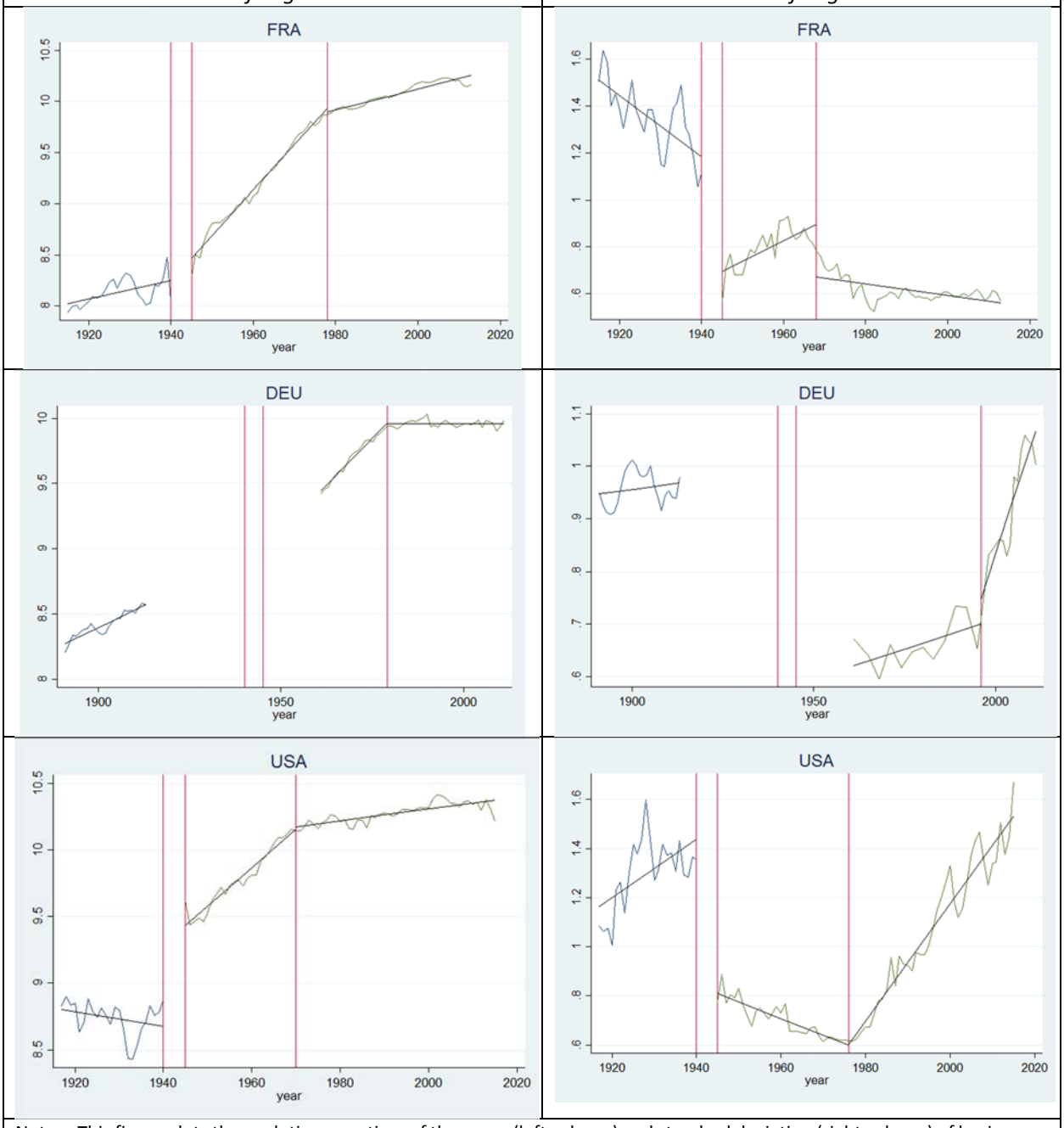
Figure 1: Estimates of Mobility from the PSID Based on Micro and Aggregate Data





Notes: This figure reports estimates of the mobility parameter β_{t-1} based on aggregate moments and true panel micro data, using annual rounds of the Panel Study of Income Dynamics (PSID) over the period 1977-1997. In both panels, the grey-shaded area is the range between the lower and upper-bound estimates of β_{t-1} based on the evolution over time of the cross-sectional mean and variance of log income, and the dashed line marks the mid-point of the grey-shaded range. The solid lines report three alternative estimates of mobility based on micro panel data. All three are based on a series of cross-sectional OLS regressions of log income on lagged log income at the household level. The "Baseline" estimates use data for all households in each year. The "Balanced Panel" estimates use only the subset of households available in all 21 years. The "Cook's-D" estimates drop households in the top 1 percent of the distribution of Cook's Distance statistic (a measure of influence) in the baseline OLS estimates.

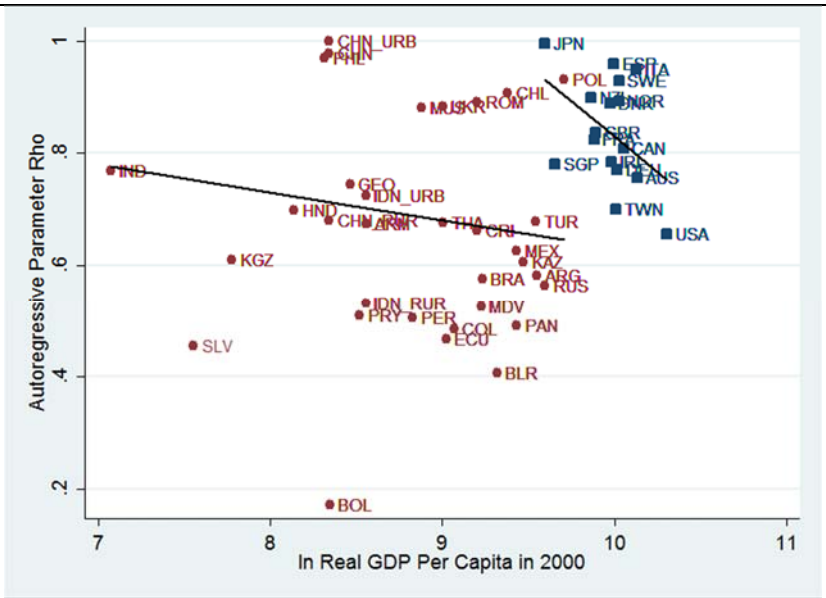
Figure 2: Evolution of Mean and Variance of Log Income, Selected WID Countries



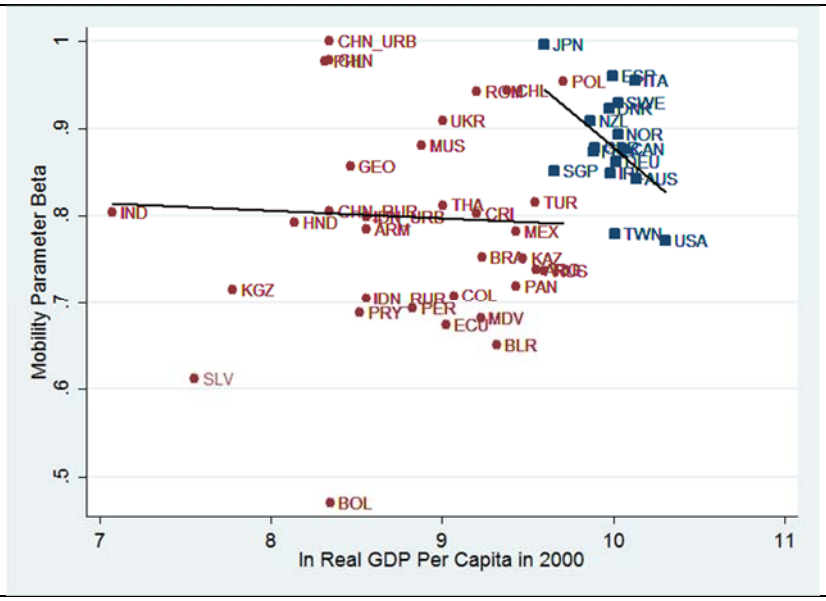
Notes: This figure plots the evolution over time of the mean (left column) and standard deviation (right column) of log income for selected countries in our WID sample. The vertical red lines delineate the subperiods corresponding to trend breaks, and the fitted lines show the estimated linear trends by subperiod. The first period for each country corresponds to all available pre-World War II data. The year of the post-war trend break is selected country-by-country and separately for the mean and variance of log income, based on the supremum of the Wald statistics for the null hypothesis that the pre- and post-break trends are the same.

Figure 3: Cross-Country Differences in Estimated Mobility

Autoregressive Parameter ρ

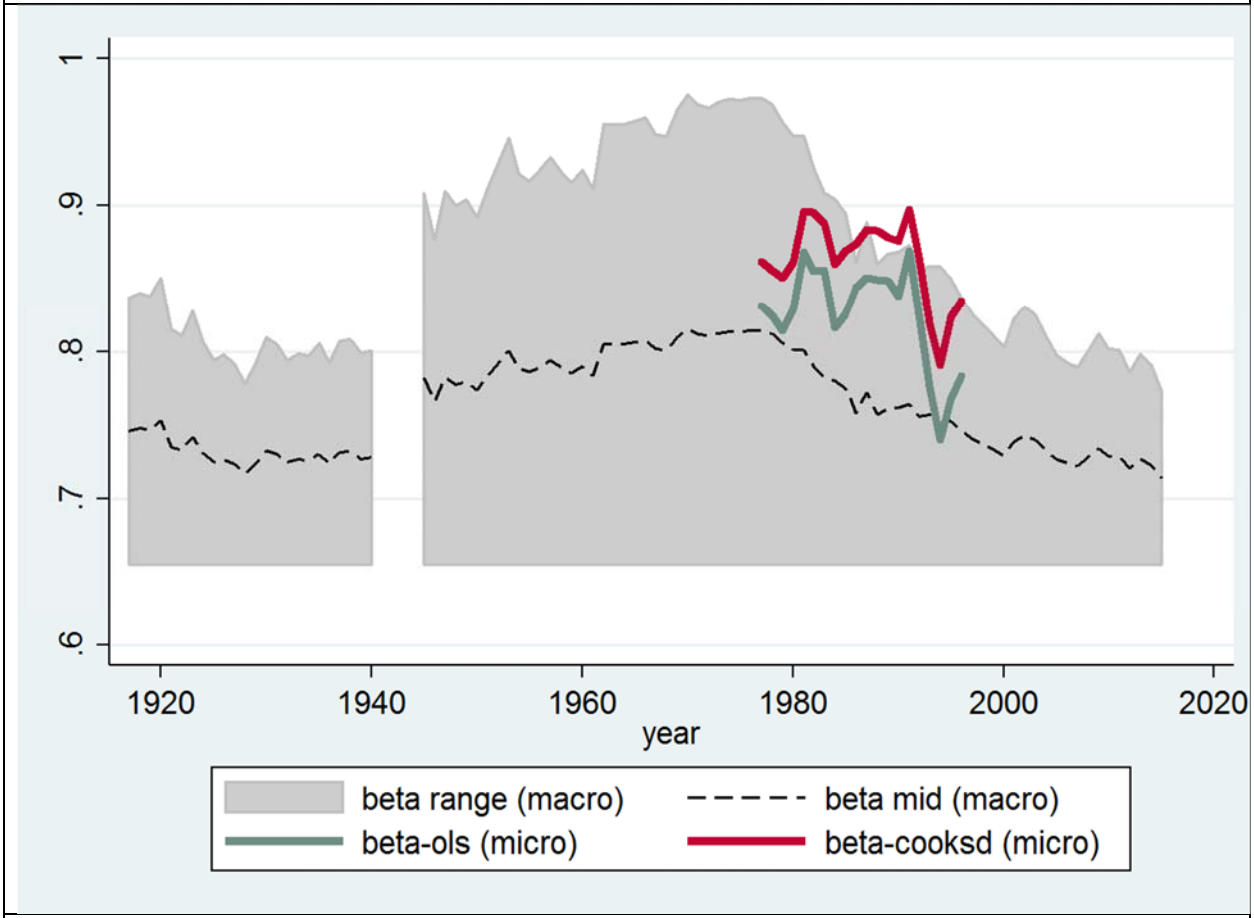


Mid-Point of Range for β_{t-1}



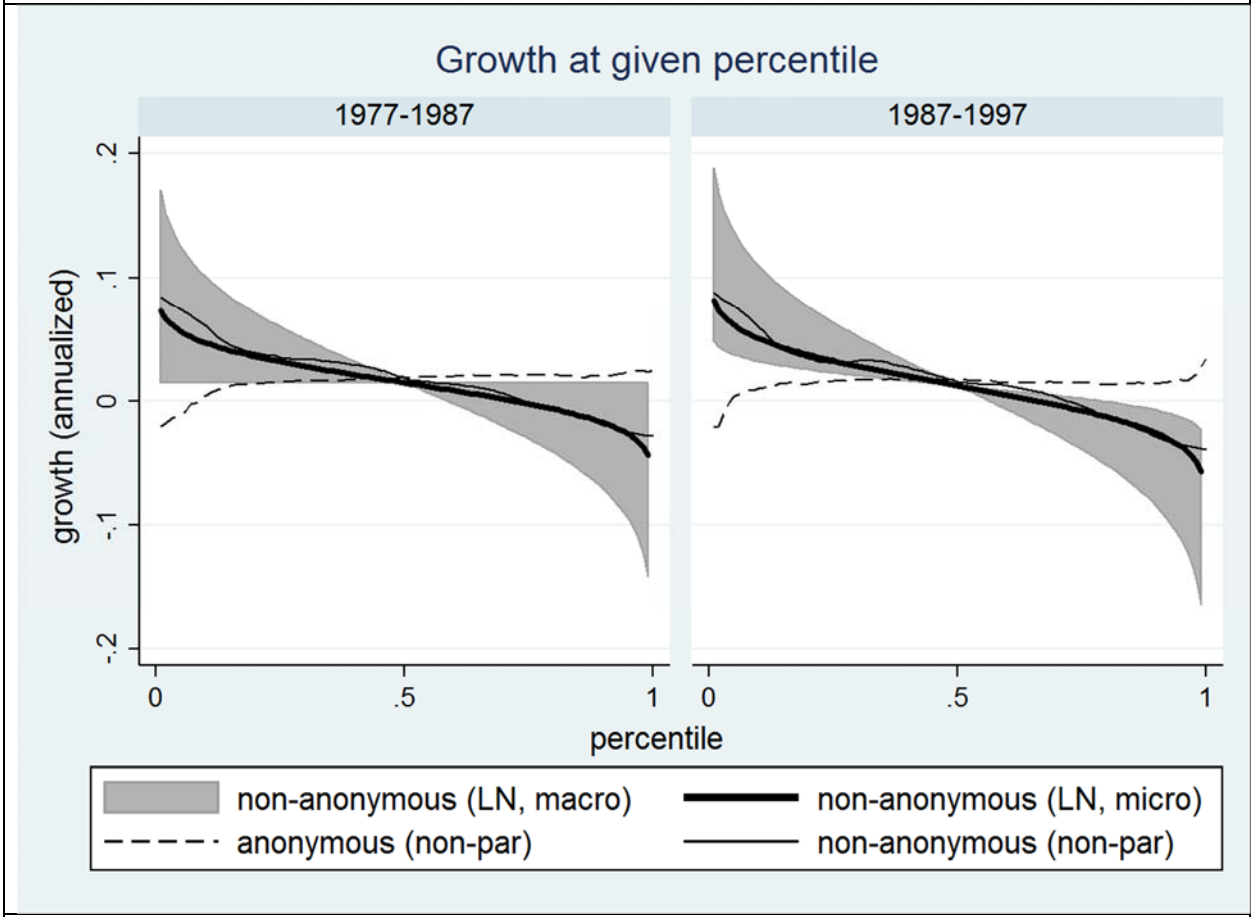
Notes: This figure plots the point estimate of ρ (top panel) and the average over time of the mid-point between the upper and lower bounds on β_{t-1} (bottom panel) on the vertical axis, against the logarithm of real GDP per capita in 2000 on the horizontal axis. In each panel, the red circles correspond to developing countries, and consist of the countries in our PovcalNet sample, as well as China, Mauritius, and India from the WID sample. The blue squres correspond to advanced economies in the WID sample.

Figure 4: Estimated Mobility Trends in the United States



Notes: This figure reports the evolution over time of the estimated bounds on the mobility parameter β_{t-1} for the United States. The grey-shaded region contains the range between the lower and upper bounds on β_{t-1}, and the dashed line indicates the midpoint of the range. The two shorter superimposed time series are the estimates of actual mobility based on the PSID microdata, as described in Section 4.

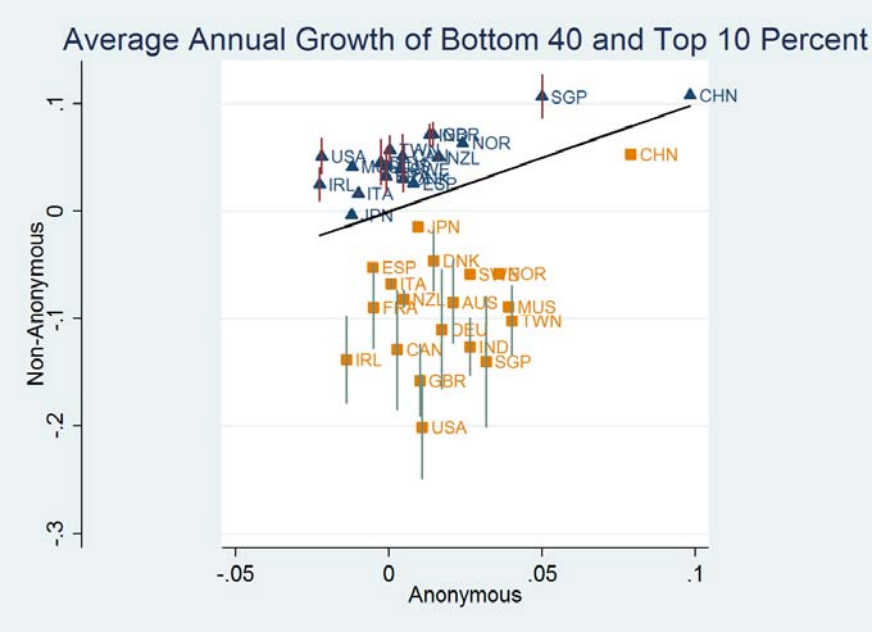
Figure 5: Anonymous and Non-Anonymous Growth Incidence Curves, PSID



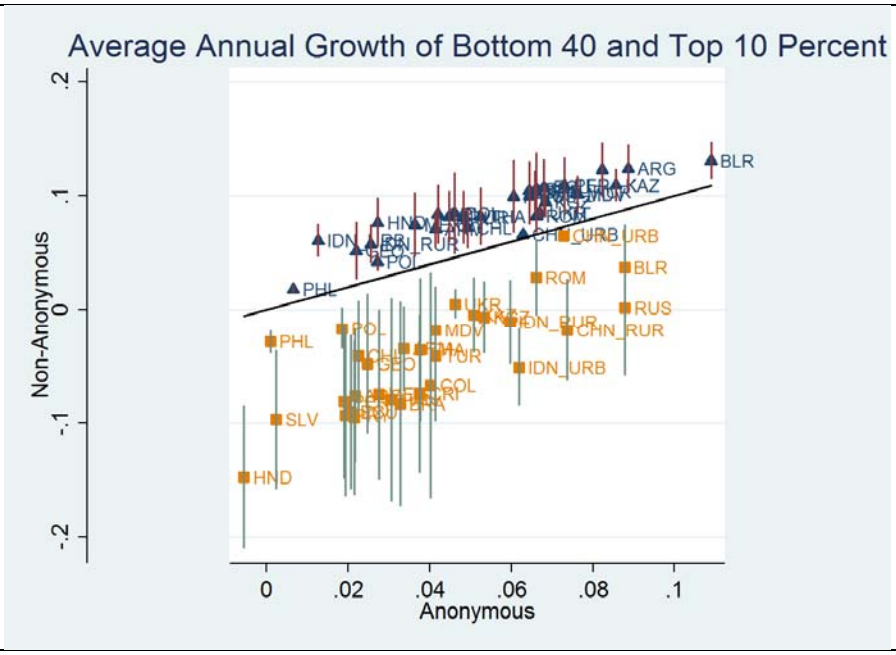
Notes: This figure plots growth incidence curves for the US, based on the PSID microdata and using the bounds on mobility based on aggregate data. The horizontal axis corresponds to percentiles of the income distribution and the vertical axis plots the average annual growth rate over the period 1987-1997 at each percentile. The dashed line reports the anonymous growth incidence curve. The two solid lines report two versions of the actual non-anonymous growth incidence curve based on the PSID micro data. The thin line is a local polynomial smoothed average of the growth rates of households at each percentile of the initial income distribution. The bold line is a lognormal approximation to the non-anonymous growth incidence curve, i.e. Equation (11), using the estimate of β_{t-k} obtained using the PSID microdata. The shaded region corresponds to the region between the lower and upper bounds on the non-anonymous growth incidence curve based only on aggregate data.

Figure 6: Anonymous and Non-Anonymous Growth Rates of Bottom 40 and Top 20 Percent

Estimates Based on WID Data



Estimates Based on PovcalNet Data



Notes: This figure plots average annual anonymous growth rates (on the horizontal axis) against average annual non-anonymous growth rates (on the vertical axis). The solid diagonal line represents the 45-degree line. The top panel reports results from the WID sample, while the bottom panel reports results from the PovcalNet sample. The blue triangles above (below the 45-degree line represent the mid-point estimate of the non-anonymous growth rate of the bottom 40 percent (top 10 percent), over the most recent 10-year period available for each country. The vertical lines represent the range from the lower to the upper-bound estimates of the non-anonymous growth rate.

Appendix A: Proofs

In the main text we assumed for notational convenience that survey data are available in consecutive periods t and $t-1$. In reality however, survey data often are available at irregular frequencies that differ over time and across countries, and in the empirical part of the paper we work with such irregularly-spaced data. In this appendix we provide proofs for the case of irregularly-spaced data, i.e. for two surveys available in periods t and $t-k$. The propositions as stated in the main text obtain for the special case of $k=1$.

Preliminaries

Iterating the data generating process in Equation (1) backwards for k periods results in:

$$y_{it} = \rho^k y_{it-k} + \left(\frac{1-\rho^k}{1-\rho}\right)\lambda_i + \tilde{\delta}_t + \tilde{\varepsilon}_{it} \tag{15}$$

where $\tilde{\delta}_t \equiv \sum_{s=0}^{k-1} \rho^s \delta_{t-s}$ and $\tilde{\varepsilon}_{it} \equiv \sum_{s=0}^{k-1} \rho^s \varepsilon_{it-s}$. For the case $t = t-k$ and $k = t-k$, and using the assumption on initial income in Assumption A1, Equation (15) becomes:

$$\begin{aligned} y_{it-k} &= \rho^{t-k}\left(\frac{\lambda_i}{1-\rho} + \delta_0 + \varepsilon_{i0}\right) + \left(\frac{1-\rho^{t-k}}{1-\rho}\right)\lambda_i + \tilde{\delta}_{t-k} + \tilde{\varepsilon}_{it-k} \\ &= \frac{\lambda_i}{1-\rho} + \rho^{t-k}(\delta_0 + \varepsilon_{i0}) + \tilde{\delta}_{t-k} + \tilde{\varepsilon}_{it-k} \end{aligned} \tag{16}$$

Proof of Proposition 1:

Given Equation (16) which states that y_{it-k} is a linear combination of $\varepsilon_{i0}, \dots, \varepsilon_{it-k}$ and λ_i, and Assumption A1 that these shocks are jointly normally distributed, it follows that y_{it-k} is normally distributed for all $t-k \geq 0$. Equation (16) also implies that $COV[y_{it-k}, \lambda_i] = \frac{\sigma_\lambda^2}{1-\rho}$. To complete the proof of Proposition 1 we need to find $COV[y_{it}, y_{it-k}]$. Using Equation (15) we have:

$$\begin{aligned} COV[y_{it}, y_{it-k}] &= \rho^k \sigma_{t-k}^2 + \left(\frac{1-\rho^k}{1-\rho}\right) COV[y_{it-k}, \lambda_i] \\ &= \left(\rho^k + \frac{1-\rho^k}{(1-\rho)^2}\frac{\sigma_\lambda^2}{\sigma_{t-k}^2}\right)\sigma_{t-k}^2 \\ &= \beta_{t-k,k}\sigma_{t-k}^2 \end{aligned} \tag{17}$$

Setting $k = 1$ and adopting the more compact notational convention that $\beta_{t-1,1} = \beta_{t-1}$ gives $\beta_{t-1} = \rho + \frac{\sigma_\lambda^2}{1-\rho}\frac{1}{\sigma_{t-1}^2}$ as in the main text.

Proof of Proposition 2

From the irregularly-spaced version of Proposition 1 we have:

$$\begin{pmatrix} y_{it} \\ y_{it-k} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_t \\ \mu_{t-k} \end{pmatrix}, \begin{pmatrix} \sigma_t^2 & \beta_{t-k,k}\sigma_{t-k}^2 \\ \beta_{t-k,k}\sigma_{t-k}^2 & \sigma_{t-k}^2 \end{pmatrix}\right) \tag{18}$$

Using the standard properties of the conditional mean and variance of the bivariate normal distribution, as well as the definition of the quantile function $y_{t-k}(p) = \mu_{t-k} + \sigma_{t-k}\Phi^{-1}(p)$ we have:

$$\mu_{t|t-k}(p) \equiv E\left[y_{t|t-k}(p)\right] = E[y_{it}|y_{it-k} = y_{t-k}(p)] = \mu_t + \beta_{t-k,k}\sigma_{t-k}\Phi^{-1}(p) \tag{19}$$

$$\sigma_{t|t-k}^2 \equiv E\left[\left(y_{t|t-k}(p) - \mu_{t|t-k}(p)\right)^2\right] = V[y_{it}|y_{it-k} = y_{t-k}(p)] = \sigma_t^2 - \beta_{t-k,k}^2\sigma_{t-k}^2 \tag{20}$$

Setting $k = 1$ retrieves the result in the main text.

Proof of Proposition 3

Taking unconditional expectations of both sides of Equation (15) gives the following irregularly-spaced analog of Equation (6).

$$\mu_t = \rho^k\mu_{t-k} + \tilde{\delta}_t \tag{21}$$

Taking unconditional variances of both sides of Equation (15) gives the following irregularly-spaced analog of Equation (7):

$$\begin{aligned} \sigma_t^2 &= \rho^{2k}\sigma_{t-k}^2 + \left(\frac{1-\rho^k}{1-\rho}\right)^2\sigma_\lambda^2 + 2\rho^k\left(\frac{1-\rho^k}{1-\rho}\right)COV(y_{it-k}, \lambda_i) + \sum_{s=0}^{k-1}\rho^{2s}\sigma_{\varepsilon t-s}^2 \\ &= \rho^{2k}\sigma_{t-k}^2 + \left(\frac{1-\rho^k}{1-\rho}\right)^2\sigma_\lambda^2 + 2\rho^k\left(\frac{1-\rho^k}{1-\rho}\right)\frac{\sigma_\lambda^2}{1-\rho} + \sum_{s=0}^{k-1}\rho^{2s}\sigma_{\varepsilon t-s}^2 \\ &= \rho^{2k}\sigma_{t-k}^2 + \frac{1-\rho^{2k}}{(1-\rho)^2}\sigma_\lambda^2 + \tilde{\sigma}_{\varepsilon t}^2 \end{aligned} \tag{22}$$

where $\tilde{\sigma}_{\varepsilon t}^2 \equiv \sum_{s=0}^{k-1} \rho^s \sigma_{\varepsilon t-s}^2$. Setting $k = 1$ retrieves the result in the main text.

Proof of Proposition 4

To prove Proposition 4 we first show that the anonymous growth rate of group average incomes is a weighted average of the anonymous growth incidence curve:

$$
\begin{aligned}
g_{t,t-k}^{A}(p,q) &\equiv \frac{Y_t(p,q) - Y_{t-k}(p,q)}{Y_{t-1}(p,q)} \\
&= \frac{1}{q-p}\int_p^q \frac{Y_t(s) - Y_{t-k}(s)}{Y_{t-k}(s)} \frac{Y_{t-k}(s)}{Y_{t-k}(p,q)} ds \\
&\approx \int_p^q \big(y_t(s) - y_{t-k}(s)\big) w_{t-k}(s) ds \\
&= \int_p^q g_{t,t-k}^{A}(s) w_{t-k}(s) ds
\end{aligned} \tag{23}
$$

where $w_{t-k}(s) \equiv \frac{1}{q-p}\frac{Y_{t-k}(s)}{Y_{t-k}(p,q)}$ is the share of percentile s in the total income of the group at time $t-k$. Similarly the non-anonymous growth rate of group average incomes is the same weighted average of the non-anonymous growth incidence curve:

$$
\begin{aligned}
g_{t,t-k}^{NA}(p,q) &\equiv \frac{E\big[Y_{t|t-k}(p,q)\big] - Y_{t-k}(p,q)}{Y_{t-k}(p,q)} \\
&= \frac{1}{q-p}\int_p^q E\left[\frac{e^{y_{t|t-k}(s)} - e^{y_{t-k}(s)}}{e^{y_{t-k}(s)}}\right] \frac{e^{y_{t-k}(s)}}{Y_{t-k}(p,q)} ds \\
&\approx \int_p^q \Big(E\big[y_{t|t-k}(s)\big] - y_{t-k}(s)\Big) w_{t-k}(s) ds \\
&= \int_p^q g_{t,t-k}^{NA}(s) w_{t-k}(s) ds
\end{aligned} \tag{24}
$$

The expression for the difference between the anonymous and non-anonymous group average growth rates in Equation (14) follows from subtracting Equations (23) and (24).

To complete the proof we need to evaluate the integral

$$
\int_p^q \Phi^{-1}(s) w_{t-k}(s) ds = \frac{1}{Y_{t-k}(p,q)} \frac{1}{q-p} \int_p^q \Phi^{-1}(s) e^{\mu_{t-k} + \sigma_{t-k}\Phi^{-1}(s)} ds \tag{25}
$$

To evaluate this remaining integral, recall that $Y_{t-k}(p,q) \equiv \frac{1}{q-p}\int_p^q e^{y_{t-k}(s)} ds = \frac{1}{q-p}\int_p^q e^{\mu_{t-k} + \sigma_{t-k}\Phi^{-1}(s)} ds$. Differentiating $Y_{t-k}(p,q)$ with respect to σ_{t-k} yields:

$$
\frac{\partial Y_{t-k}(p,q)}{\partial \sigma_{t-k}} = \frac{1}{q-p}\int_p^q \Phi^{-1}(s) e^{\mu_{t-k} + \sigma_{t-k}\Phi^{-1}(s)} ds \tag{26}
$$

This means that we can find the integral we need simply by differentiating group average income with respect to σ_{t-k}. Combining Equations (25) and (26) and using the property of the truncated lognormal

distribution that $Y_{t-k}(p,q) = e^{\mu_{t-k}+\frac{\sigma_{t-k}^2}{2}}\left(\Phi(\Phi^{-1}(q)-\sigma_{t-k}) - \Phi(\Phi^{-1}(p)-\sigma_{t-k})\right)/(q-p)$, we obtain:

$$\int_p^q \Phi^{-1}(s)w_{t-k}(s)ds = \frac{\partial Y_{t-k}(p,q)}{\partial \sigma_{t-k}}\frac{1}{Y_{t-k}(p,q)}$$

$$= \frac{\partial}{\partial \sigma_{t-k}}\left(e^{\mu_{t-k}+\frac{\sigma_{t-k}^2}{2}}\left(\Phi(\Phi^{-1}(q)-\sigma_{t-k}) - \Phi(\Phi^{-1}(p)-\sigma_{t-k})\right)\right)\frac{1}{Y_{t-k}(p,q)} \quad (27)$$

$$= \left(\sigma_{t-k} - \frac{\phi(\Phi^{-1}(q)-\sigma_{t-k}) - \phi(\Phi^{-1}(p)-\sigma_{t-k})}{\Phi(\Phi^{-1}(q)-\sigma_{t-k}) - \Phi(\Phi^{-1}(p)-\sigma_{t-k})}\right)$$

For the particular case of average incomes in the bottom q percent this expression simplifies to

$$\int_0^q \Phi^{-1}(s)w_{t-k}(s)ds = \left(\sigma_{t-k} - \frac{\phi(\Phi^{-1}(q)-\sigma_{t-k})}{\Phi(\Phi^{-1}(q)-\sigma_{t-k})}\right) \quad (28)$$

while for the particular case of average incomes in the top p percent this expression simplifies to

$$\int_p^1 \Phi^{-1}(s)w_{t-k}(s)ds = \left(\sigma_{t-k} + \frac{\phi(\Phi^{-1}(p)-\sigma_{t-k})}{1-\Phi(\Phi^{-1}(p)-\sigma_{t-k})}\right). \quad (29)$$

Finally, the fact that $g_t^i(p,q)$ can also be represented as: $g_t^i(p,q) = g_t^i(z(p,q,\sigma_{t-1}))$, where $z(p,q,\sigma)$ satisfies: $p \le z(p,q,\sigma) \le q$, follows directly from the fact that $g_t^i(z)$ is a monotonic function of z.

Appendix B: Details of OLS, Bias-Corrected, and MSE-Minimizing Estimates of ρ

This appendix described our approach to estimating the autoregressive coefficient in log individual incomes using aggregate data. Let $\hat{\rho}_{\mu,OLS}$ denote the OLS estimator of ρ based on Equation (8), and let $\widehat{\rho^2}_{\sigma,OLS}$ denote the OLS estimator of ρ^2 based on Equation (9). Given that the available time series is short, these estimators will exhibit downwards finite-sample bias. We therefore also generate corresponding bias-corrected estimators $\hat{\rho}_{\mu,BC}$ and $\widehat{\rho^2}_{\sigma,BC}$ using the procedure suggested in Andrews (1993). At the core of this procedure is the fact that the distribution of the OLS estimator is exclusively a function of the true autoregressive parameter and the sample size, and thus is independent of the parameters that describe the distribution of the error term and the time-trend. We refer the interested reader to Andrews (1993) for a proof. We take advantage of this result, as suggested by Andrews (1993), by computing the median bias of the OLS estimator as a function of ρ for each sample size separately (using numerical simulations), and then inverting this function to obtain a median-unbiased estimator for ρ.

Comparing the OLS and bias-corrected estimators highlights a tradeoff between bias and variance: while the OLS estimator is substantially downward biased, the bias-corrected estimator is much less precisely estimated. We address this tradeoff by defining the following two weighted averages of the OLS and bias-corrected estimators:

$$\hat{\rho}_{\mu} = \omega_{\mu}\hat{\rho}_{\mu,OLS} + (1-\omega_{\mu})\hat{\rho}_{\mu,BC} \quad (30)$$

(31) $$\widehat{\rho^2}_{\sigma} = \omega_{\sigma}\widehat{\rho^2}_{\sigma,OLS} + (1-\omega_{\sigma})\widehat{\rho^2}_{\sigma,BC}$$

with the weights ω_{μ} and ω_{σ} chosen to minimize the mean squared error of the corresponding linear combination of estimators. Finally, we combine the two estimators in Equation (30) into a single preferred estimator of ρ as follows:

(32) $$\hat{\rho} = \omega\hat{\rho}_{\mu} + (1-\omega)\left(\widehat{\rho^2}_{\sigma}\right)^{\frac{1}{2}}$$

with a weight ω again chosen to minimize mean squared error.

To derive the specific expressions for the MSE-minimizing weights ω_{μ}, ω_{σ}, and ω, it is convenient to work with a generic case of two possibly correlated estimators, $\hat{\rho}_1$ and $\hat{\rho}_2$, with biases b_1 and b_2; variances V_1 and V_2; and covariance C_{12}. With this notation, the MSE of a weighted average of the two estimators is:

(33) $$\begin{aligned} MSE[x\hat{\rho}_1 + (1-x)\hat{\rho}_2] &\equiv Bias[x\hat{\rho}_1 + (1-x)\hat{\rho}_2]^2 + V[x\hat{\rho}_1 + (1-x)\hat{\rho}_2] \\ &= (xb_1 + (1-x)b_2)^2 + x^2V_1 + 2x(1-x)C_{12} + (1-x)^2V_2 \end{aligned}$$

Setting the derivative of this expression with respect to x equal to zero, followed by some straightforward algebra, results in this expression for the MSE-minimizing weight:

(34) $$x^* = \frac{b_2(b_2-b_1) + V_2 - C_{12}}{(b_1-b_2)^2 + V_1 - 2C_{12} + V_2}$$

Using this weight, the MSE-minimizing weighted average of the two estimators is $x^*\hat{\rho}_1 + (1-x^*)\hat{\rho}_2$, and has bias $x^*b_1 + (1-x^*)b_2$ and variance $x^{*2}V_1 + 2x^*(1-x^*)C_{12} + (1-x^*)^2V_2$. Note that the weight need not lie between 0 and 1. In particular when the two estimators are highly correlated, a weight larger than 1 on one of the two estimators, and a corresponding weight less than 0 on the other, will exploit this correlation to reduce the variance of the combined estimator.

To apply this generic expression to Equation (30) we assume that the bias-corrected estimator eliminates the bias in the OLS estimator, so that $b_2 = 0$ and $b_1 = \hat{\rho}_{\mu,OLS} - \hat{\rho}_{\mu,BC}$. Let $V_1 = V\left[\hat{\rho}_{\mu,OLS}\right]$ denote the variance of the OLS estimator. Since the bias-corrected estimator is a function of the OLS estimator, we can linearize to obtain $V_2 = V\left[\hat{\rho}_{\mu,BC}\right] \approx \nabla_{\mu}^2 V\left[\hat{\rho}_{\mu,OLS}\right]$ and $C_{12} = COV\left[\hat{\rho}_{\mu,OLS}, \hat{\rho}_{\mu,BC}\right] \approx \nabla_{\mu} V\left[\hat{\rho}_{\mu,OLS}\right]$, where ∇_{μ} is the gradient of the bias corrected estimator as a function of the OLS estimator evaluated at the OLS estimate, that we compute numerically. Inserting these into Equation (34) results in this MSE-minimizing weight:

(35) $$\omega_{\mu}^{*} = \frac{V\left[\hat{\rho}_{\mu,OLS}\right]\nabla_{\mu}(\nabla_{\mu} - 1)}{\left(\hat{\rho}_{\mu,OLS} - \hat{\rho}_{\mu,BC}\right)^{2} + V\left[\hat{\rho}_{\mu,OLS}\right]\left(\nabla_{\mu} - 1\right)^{2}}$$

Applying the same reasoning to Equation (31) gives this corresponding MSE-minimizing weight:

(36) $$\omega_{\sigma}^{*} = \frac{V[\widehat{\rho^{2}}_{\sigma,OLS}]\nabla_{\sigma}(\nabla_{\sigma} - 1)}{\left(\widehat{\rho^{2}}_{\sigma,OLS} - \widehat{\rho^{2}}_{\sigma,BC}\right)^{2} + V\left[\widehat{\rho^{2}}_{\sigma,OLS}\right](\nabla_{\sigma} - 1)^{2}}$$

Finally, to obtain the MSE-minimizing weight in Equation (32), note that two MSE-minimizing estimators from Equations (30) and (31) are both biased with $b_1 = \hat{\rho}_{\mu} - \hat{\rho}_{\mu,BC}$ and $b_2 = \left(\widehat{\rho^{2}}_{\sigma}\right)^{\frac{1}{2}} - \left(\widehat{\rho^{2}}_{\sigma,BC}\right)^{\frac{1}{2}}$. In addition we have $V_1 = V\left[\hat{\rho}_{\mu}\right]$ and $V_2 = V\left[\left(\widehat{\rho^{2}}_{\sigma}\right)^{\frac{1}{2}}\right] \approx \frac{1}{2}\left(\widehat{\rho^{2}}_{\sigma}\right)^{-\frac{1}{2}} V\left[\left(\widehat{\rho^{2}}_{\sigma}\right)\right]$ and we assume $C_{12} = 0$. This results in

(37) $$\omega^{*} = \frac{\left(\left(\widehat{\rho^{2}}_{\sigma}\right)^{\frac{1}{2}} - \left(\widehat{\rho^{2}}_{\sigma,BC}\right)^{\frac{1}{2}}\right)\left(\left(\widehat{\rho^{2}}_{\sigma}\right)^{\frac{1}{2}} - \left(\widehat{\rho^{2}}_{\sigma,BC}\right)^{\frac{1}{2}} - \left(\hat{\rho}_{\mu} - \hat{\rho}_{\mu,BC}\right)\right) + \frac{1}{2}\left(\widehat{\rho^{2}}_{\sigma}\right)^{-\frac{1}{2}} V\left[\left(\widehat{\rho^{2}}_{\sigma}\right)\right]}{\left(\left(\widehat{\rho^{2}}_{\sigma}\right)^{\frac{1}{2}} - \left(\widehat{\rho^{2}}_{\sigma,BC}\right)^{\frac{1}{2}} - \left(\hat{\rho}_{\mu} - \hat{\rho}_{\mu,BC}\right)\right)^{2} + V\left[\hat{\rho}_{\mu}\right] + \frac{1}{2}\left(\widehat{\rho^{2}}_{\sigma}\right)^{-\frac{1}{2}} V\left[\left(\widehat{\rho^{2}}_{\sigma}\right)\right]}$$

Appendix Table B1: Estimation Details for WID Sample

	Estimation Sample				Estimate of ρ Based on Equation for:								Preferred Estimate of ρ	
	Time Period		*Break Point for:*		*Mean of log Income*				*Variance of log income*					Weight on
	First	Last	Mean	Var-	OLS	BC-OLS	Min-MSE	Weight	OLS	BC-OLS	Min-MSE	Weight	Estimate	Estimate Based
	Year	Year		iance				on OLS				on OLS		Mean of log Incom
Australia	1958	2013	1982	1974	0.723	1.000	0.719	1.015	0.667	0.772	0.759	0.131	0.756	0.066
					0.100	1.000	0.086		0.096	0.097	0.096		0.090	
Canada	1950	2010	1974	1992	0.694	0.896	0.844	0.259	0.714	0.811	0.797	0.151	0.808	0.230
					0.094	0.166	0.148		0.083	0.088	0.088		0.075	
China	1978	2015	1989	2002	0.928	1.000	0.921	1.089	0.807	1.000	0.805	1.012	0.978	1.493
					0.056	0.563	0.011		0.077	0.618	0.068		0.038	
Denmark	1950	2010	1966	1974	0.717	0.934	0.870	0.296	0.885	1.000	0.895	0.920	0.889	0.239
					0.093	0.183	0.156		0.030	0.263	0.051		0.054	
France	1915	2013	1978	1968	0.693	0.868	0.844	0.137	0.741	0.825	0.818	0.092	0.824	0.245
					0.072	0.112	0.106		0.060	0.063	0.063		0.054	
Germany	1891	2011	1979	1996	0.489	0.658	0.640	0.107	0.888	1.000	0.885	1.025	0.769	0.471
					0.103	0.127	0.125		0.046	0.404	0.035		0.062	
Great Britain	1950	2012	1968	1996	0.582	0.737	0.713	0.154	1.006	1.000	1.000	0.000	0.837	0.568
					0.109	0.137	0.132		0.015	0.154	0.154		0.100	
India	1950	1999	1983	1983	0.343	0.492	0.461	0.212	0.738	0.868	0.849	0.155	0.768	0.208
					0.148	0.176	0.170		0.077	0.091	0.089		0.079	
Ireland	1975	2009	1997	1999	0.796	1.000	0.784	1.059	0.787	1.000	0.783	1.016	0.784	0.726
					0.102	1.022	0.048		0.088	0.691	0.075		0.040	
Italy	1974	2009	2000	1984	0.724	1.000	0.740	0.944	0.888	1.000	0.879	1.071	0.950	-0.505
					0.073	0.729	0.109		0.068	0.607	0.025		0.067	
Japan	1945	2010	1974	1974	0.837	1.000	0.831	1.040	0.964	1.000	0.961	1.100	0.996	-0.270
					0.069	0.690	0.044		0.041	0.394	0.004		0.013	
Mauritius	1952	2011	1974	1976	0.869	1.000	0.858	1.079	0.876	1.000	0.878	0.983	0.880	-0.077
					0.085	0.851	0.024		0.041	0.357	0.047		0.051	
New Zealand	1950	2013	1967	1990	0.809	1.000	0.803	1.035	0.823	0.933	0.906	0.261	0.900	0.061
					0.078	0.782	0.054		0.063	0.090	0.084		0.079	
Norway	1950	2011	1986	1993	0.788	1.000	0.782	1.028	0.822	0.932	0.899	0.310	0.893	0.050
					0.083	0.826	0.062		0.071	0.100	0.091		0.087	
Singapore	1950	2012	1997	1997	0.687	0.874	0.850	0.131	0.668	0.759	0.751	0.097	0.780	0.297
					0.070	0.113	0.107		0.070	0.072	0.072		0.060	
Spain	1981	2012	1992	2005	0.692	1.000	0.687	1.015	0.871	1.000	0.864	1.049	0.960	-0.543
					0.111	1.111	0.096		0.063	0.550	0.036		0.076	
Sweden	1945	2013	1990	1975	0.790	1.000	0.802	0.944	0.847	0.944	0.931	0.141	0.929	0.016
					0.055	0.554	0.083		0.041	0.058	0.056		0.055	
Taiwan	1977	2013	1996	1993	0.449	0.711	0.635	0.290	0.600	0.773	0.729	0.279	0.701	0.301
					0.160	0.249	0.223		0.146	0.159	0.155		0.127	
United States	1917	2015	1970	1976	0.440	0.559	0.551	0.071	0.642	0.725	0.717	0.100	0.655	0.371
					0.084	0.095	0.094		0.072	0.072	0.072		0.057	

Notes: This table reports details of the estimates of ρ in the WID sample. The first four columns indicate the first and last years of the sample for each country in the post-World War II period, together with the estimated trend break dates for the mean and variance of log income. Note that three countries shown in Figure 2 also have data for a pre-World War II sub-period. The next two sets of four columns report the OLS, bias-corrected OLS, the MSE-minimizing weighted average of the two, and the MSE-minimizing weight on the OLS equation, for dynamics of the mean and variance of log income, respectively. The last two columns report the MSE-minimizing weighted average of the estimates of ρ from the two equations, together with the MSE-minimizing weight on the estimate from the equation for the dynamics of the mean of log income. Standard errors are reported below the coefficient estimates for the OLS, bias-corrected OLS, and MSE-minimizing weighted average estimators.

Appendix Table B2: Estimation Details for PovcalNet Sample

	Estimation Sample			Estimate of ρ Based on Equation for:								Preferred Estimate of ρ	
	First	Last	Freq-	*Mean of log Income*				*Variance of log income*					Weight on
	Year	Year	uency	OLS	BC-OLS	Min-MSE	Weight on OLS	OLS	BC-OLS	Min-MSE	Weight on OLS	Estimate	Estimate Based Mean of log Income
Argentina	1991	2013	1	0.327	0.816	0.553	0.538	0.319	0.637	0.586	0.208	0.581	0.129
				0.204	0.555	0.366		0.259	0.195	0.197		0.178	
Armenia	1999	2013	1	0.716	1.000	0.690	1.090	0.564	0.804	0.655	0.663	0.673	0.516
				0.232	2.330	0.042		0.268	0.309	0.281		0.138	
Belarus	1998	2012	1	0.430	0.872	0.448	0.960	0.270	0.547	0.371	0.714	0.405	0.452
				0.264	0.910	0.290		0.556	0.375	0.447		0.278	
Bolivia	2000	2013	1	-0.105	0.077	0.077	0.000	0.275	0.566	0.348	0.813	0.171	0.652
				0.310	0.403	0.403		0.595	0.411	0.507		0.316	
Brazil	1995	2013	1	0.057	0.437	0.230	0.544	0.699	1.000	0.682	1.045	0.576	0.236
				0.280	0.490	0.376		0.208	2.830	0.037		0.093	
Chile	1987	2013	2	0.730	1.000	0.703	1.082	0.826	1.000	0.816	1.045	0.907	-0.816
				0.218	1.595	0.058		0.104	0.587	0.065		0.127	
China (Rural)	1990	2012	2	0.715	1.000	0.684	1.088		0.669	0.669	0.000	0.680	0.704
				0.258	1.844	0.057			0.495	0.495		0.152	
China (Urban)	1990	2012	2	0.829	1.000	0.812	1.089	0.979	1.000	0.976	1.110	0.999	-0.140
				0.147	1.220	0.030		0.098	0.914	0.001		0.004	
Colombia	1999	2013	1	0.159	0.413	0.245	0.663	0.515	0.752	0.629	0.564	0.485	0.375
				0.298	0.424	0.341		0.285	0.291	0.283		0.218	
Costa Rica	1989	2013	1	0.341	0.499	0.428	0.454	0.667	0.792	0.753	0.334	0.661	0.281
				0.204	0.253	0.230		0.147	0.154	0.152		0.127	
Ecuador	2003	2013	1	0.414	1.000	0.379	1.060		0.540	0.540	0.000	0.467	0.451
				0.295	2.955	0.137			0.577	0.577		0.323	
El Salvador	1997	2013	1	0.308	0.574	0.484	0.338	0.199	0.457	0.255	0.848	0.455	0.873
				0.199	0.291	0.260		0.754	0.413	0.610		0.240	
Georgia	1997	2013	1	0.467	0.834	0.542	0.795	0.753	1.000	0.741	1.042	0.745	-0.019
				0.268	0.589	0.333		0.124	0.932	0.079		0.080	
Honduras	1989	2013	1	0.529	0.739	0.664	0.355	0.622	0.751	0.712	0.319	0.698	0.299
				0.177	0.247	0.222		0.151	0.155	0.154		0.127	
Indonesia (Rural)	1996	2013	1	0.118	0.300	0.210	0.494	0.763	1.000	0.757	1.023	0.531	0.413
				0.247	0.306	0.277		0.104	0.792	0.083		0.124	
Indonesia (Urban)	1996	2013	1	0.202	0.402	0.306	0.480	0.836	1.000	0.826	1.056	0.725	0.195
				0.230	0.301	0.267		0.088	0.732	0.044		0.063	
Kazakhstan	2001	2013	1	-0.104	0.103	0.103	0.000	0.774	1.000	0.759	1.058	0.604	0.236
				0.328	0.447	0.447		0.128	0.993	0.063		0.116	
Kyrgyzstan	1998	2012	1	0.331	0.669	0.525	0.426	0.684	1.000	0.686	0.994	0.609	0.482
				0.229	0.382	0.317		0.125	0.862	0.132		0.167	
Moldova	1997	2013	1	0.377	0.679	0.573	0.351		0.189	0.189	0.000	0.524	0.873
				0.205	0.319	0.279			0.782	0.782		0.263	
Mexico	1992	2012	2	0.618	1.000	0.589	1.056		0.673	0.673	0.000	0.626	0.566
				0.243	1.500	0.127			0.491	0.491		0.225	
Panama	2000	2013	1	0.208	0.514	0.385	0.422	0.600	0.883	0.642	0.875	0.490	0.591
				0.235	0.363	0.309		0.280	0.416	0.300		0.220	
Peru	1997	2013	1	0.150	0.724	0.261	0.807		0.610	0.610	0.000	0.504	0.303
				0.332	0.926	0.447			0.424	0.424		0.325	
Philippines	1985	2012	3	0.644	1.000	0.599	1.072	0.883	1.000	0.869	1.083	0.969	-0.368
				0.343	1.425	0.139		0.113	0.608	0.030		0.066	
Poland	1998	2012	1	0.404	0.801	0.555	0.619	0.905	1.000	0.895	1.095	0.932	-0.108
				0.230	0.517	0.339		0.093	0.843	0.013		0.040	
Paraguay	1995	2013	1	0.046	0.416	0.237	0.482	0.462	0.865	0.699	0.483	0.508	0.413
				0.258	0.444	0.355		0.279	0.337	0.304		0.231	
Romania	1999	2008	1	0.250	1.000	0.271	0.972	0.711	1.000	0.680	1.088	0.891	-0.517
				0.220	2.201	0.276		0.267	1.895	0.057		0.167	
Russia	1999	2012	1	0.295	0.663	0.375	0.784	0.517	0.783	0.628	0.635	0.563	0.254
				0.344	0.608	0.401		0.289	0.321	0.298		0.244	
Thailand	1988	2012	2	0.681	1.000	0.655	1.066		0.692	0.692	0.000	0.676	0.436
				0.212	1.444	0.090			0.313	0.313		0.181	
Turkey	2002	2012	1	-0.256	0.000	0.000	0.000	0.812	1.000	0.791	1.099	0.677	0.145
				0.360	0.510	0.510		0.220	1.788	0.025		0.077	
Ukraine	2002	2013	1	0.674	1.000	0.653	1.066	0.776	1.000	0.755	1.080	0.884	-1.269
				0.175	1.751	0.072		0.168	1.306	0.048		0.143	

Notes: Notes: This table reports details of the estimates of ρ in the WID sample. The first three columns indicate the first and last years of the sample for each country, and the frequency of the surveys (every 1, 2 or 3 years). The next two sets of four columns report the OLS, bias-corrected OLS, the MSE-minimizing weighted average of the two, and the MSE-minimizing weight on the OLS equation, for dynamics of the mean and variance of log income, respectively. The last two columns report the MSE-minimizing weighted average of the estimates of ρ from the two equations, together with the MSE-minimizing weight on the estimate from the equation for the dynamics of the mean of log income. Standard errors are reported below the coefficient estimates for the OLS, bias-corrected OLS, and MSE-minimizing weighted average estimators. Missing OLS estimates for the equation for the variance of log income correspond to cases where the OLS estimates are negative. In these cases the estimates are based on the bias-corrected estimator only.