RIP to HIP:

The Data Reject Heterogeneous Labor Income Profiles *

Dmytro Hryshko†

University of Alberta

Abstract

Idiosyncratic labor incomes are typically modeled either by stochastic processes featuring heterogeneous income profiles (HIP) or restricted income profiles (RIP). The HIP assumes that individual labor income grows deterministically at an unobserved rate and contains a persistent but stationary component, while the RIP assumes that income contains a random walk, a stationary component, and no unobserved deterministic growth-rate component. I show that if idiosyncratic labor income contains a persistent component, a deterministic household-specific trend, and a random walk component, then all of the components can be identified. Using data on idiosyncratic labor income growth from the Panel Study of Income Dynamics, I find that the estimated variance of deterministic income growth is zero, i.e., the HIP model can be rejected. The RIP model with a permanent component cannot be rejected. This result is important for an appropriate choice of modeling the heterogeneity in individual incomes and calibrating/estimating macro models with incomplete insurance markets and heterogeneous agents.

KEYWORDS: Idiosyncratic income processes, heterogeneity, labor income risk.

*I benefited from conversations with and/or comments from Adolf Buse, Flavio Cunha, Fatih Guvenen, Gueorgui Kambourov, Greg Kaplan, Iourii Manovskii, René Morissette, Giuseppe Moscarini, Bent Sørensen, Henry van Egteren, and Gianluca Violante. For useful comments, I thank conference participants at the 2008 SED Meetings in Boston, the 2008 CEA Meetings in Vancouver, and the 2009 NBER Summer Institute (Consumption group); and seminar participants at the EERC in Kiev, the University of Pennsylvania, the University of British Columbia, UCL, and the Federal Reserve Bank of St. Louis.

†University of Alberta, Department of Economics, 8-14 HM Tory Building. Edmonton, Alberta, Canada, T6G 2H4. E-mail: dhryshko@ualberta.ca. Phone: 780-4922544. Fax: 780-4923300.

1 Introduction

Individuals and households face substantial amounts of idiosyncratic labor market risk. Layoffs, health shocks, bonuses, promotions, demotions, and time-varying returns to the individual skills valued by labor market contribute towards fluctuating individual labor incomes. Idiosyncratic labor income risk, absent perfectly functioning credit and insurance markets, affects individual and aggregate welfare.

Two different approaches to modeling individual and household labor income risks currently stand out. The first approach, with a long-standing tradition, models each individual's income growing at the individual-specific, deterministic rate, with the level of income affected by a stochastic component with moderate persistence. Since each individual's labor income profile, even in the absence of shocks, is unique, I label this model, following Guvenen (2009), the "Heterogeneous Income Profiles" (HIP) model. The second approach models idiosyncratic labor income as the sum of a permanent random walk component, the shocks to which persist for the entire working lifetime of an individual, and a mean-reverting stationary component, the shocks to which die out quickly. Since this model abstracts from the deterministic growth-rate heterogeneity, I label it the "Restricted Income Profiles" (RIP) model. Even though variants of the RIP are currently a preferred choice in macro models, there is no consensus in the labor income processes literature on which income model best fits the earnings data. As Guvenen (2007) concludes: " ... it is fair to say that this literature has not produced an unequivocal verdict." This paper is a step towards finding a verdict in favor of the RIP model.

I start with a general income model that encompasses the RIP and HIP models. I then conduct a Monte Carlo study to explore identification of different income processes found in the literature, obtained when certain restrictions on this general process are imposed. I find that if the true income process is the RIP with a permanent random walk component and an econometrician estimates the misspecified HIP model instead, he will typically find statistically significant amounts of the growth-rate heterogeneity, of magnitudes comparable with those in the literature. I show that the general income process composed of a deterministic growth rate, a permanent random walk, and transitory components can be identified when the earnings data used in estimation are in first differences. The results of a Monte Carlo study confirm that the parameters of this general process should be precisely recovered. I then proceed by estimating the model utilizing labor income data for male household heads from the Panel Study of Income Dynamics (PSID). I find that the estimate of the variance of the deterministic growth-rate component is zero, while the variance of the shock to the random walk component is significant and substantial. Hence, the data utilized in this paper favor the RIP model with a permanent random walk component and a mean-reverting persistent process.

The results of this paper are important as they contribute to understanding a number of issues. First, they speak to the economists' choices for modeling of household consumption, savings, and wealth. If the correct model for idiosyncratic labor income is the HIP, one needs to model individuals sequentially learning about their own labor income profiles to jointly fit the features of consumption and income data. Guvenen (2007) is an example of such a model that successfully explains the profile of consumption inequality observed in U.S. micro data and the co-movement of the life-cycle profiles of earnings and consumption for households with different levels of schooling. If a substantial variation in incomes is due to permanent and persistent shocks, as is found in this paper, an appropriate model for household choices of consumption, savings, and wealth is an incomplete markets model with uninsurable persistent and/or permanent shocks. Castañeda, Díaz-Giménez, and Ríos-Rull (2003), utilizing such a model, successfully explain the U.S. wealth and earnings inequality; Scholz, Seshadri, and Khitatrakun (2006) explain more than 80% of the 1992 cross-sectional variation of household wealth observed in data from the Health and Retirement Study. Krebs (2003) is an example of a model where permanent idiosyncratic risk, absent in the estimations of the HIP processes but found to be substantial in this and some other papers,[1] reduces economic growth and individual welfare. De Santis (2007) develops a model where log-individual consumption is a random walk due to permanent uninsurable idiosyncratic income shocks and shows that such a model can potentially produce large welfare gains from eliminating business cycles.

Second, the results of this paper speak to the literature on the importance of initial conditions at the start of the individual's working career versus life-cycle shocks for the lifetime inequality in earnings and welfare (for recent contributions, see Storesletten, Telmer, and Yaron (2004) and Huggett, Ventura, and Yaron (2007)). If household incomes contain a random walk and persistent components, the marginal propensity to consume from the permanent shock should be close to one, and this reaction should translate one-for-one into consumption inequality among a cross-section of households with similar labor market experience. The contribution of the life-cycle labor income shocks will be understated if one models household incomes as the HIP, since, as shown in this paper, initial conditions at the start of the individual's working career—partly determined by the variance of the deterministic growth-rate heterogeneity—will capture the variation in incomes

[1] The prominent papers that find substantial amounts of permanent idiosyncratic labor income risk are Carroll and Samwick (1997), Meghir and Pistaferri (2004), and Moffitt and Gottschalk (1995).

due to permanent shocks. Third, the contribution of the variance due to persistent components towards the rising earnings inequality observed in the U.S. will be underestimated if the random walk component is ignored.[2]

Lastly, the idiosyncratic labor income process, best fitting the data utilized in this paper, places restrictions on the models attempting to endogenize labor incomes. A fruitful starting point can be the model in Krebs (2003), where, in equilibrium, permanent shocks to individual human capital translate into permanent shocks to individual labor incomes.

From a policy perspective, it also matters whether the true income process is the HIP or RIP. If an objective of the policymaker is to reduce consumption inequality and the true idiosyncratic income process is the HIP with a stochastic component of moderate persistence, the policymaker may want to implement policies that subsidize human capital investments by disadvantaged; self-insurance will be a sufficient shield against the shocks of moderate persistence. If, however, the true income process is the RIP with substantial permanent shocks, an appropriate policy, in addition to the above-mentioned, is to educate the public about risk-sharing instruments provided by credit institutions, stock, and insurance markets.

The rest of the paper is structured as follows. In Section 2, I present a Monte Carlo study of income processes found in the literature and introduce the HIP and RIP models. I estimate income processes on simulated data; and also discuss identification of the models containing a random walk and deterministic growth-rate components when data used for estimation are in first differences. In Section 3, I first describe the data I use and then present the empirical results. Section 4 concludes.

2 A Monte Carlo Study

In this section, I present the income processes estimated in the literature and perform a Monte Carlo study to explore identification of those income processes.

Let the true income process be:

$$y_{iht} = \alpha_i + \beta_i h + p_{iht} + \tau_{iht} + u_{iht,me} \tag{1}$$

$$p_{iht} = p_{ih-1t-1} + \xi_{iht} \tag{2}$$

$$\tau_{iht} = \theta(L)\epsilon_{iht}, \tag{3}$$

[2]See Baker and Solon (2003), which elaborates on this issue using Canadian earnings data.

where y_{iht} is the idiosyncratic log-income of individual i with h years of (potential) labor market experience at time t; β_i is individual i's growth rate of income; α_i is individual i's initial level of income; p_{iht} is the permanent stochastic component of income; ξ_{iht} is a mean-zero shock to the permanent component; τ_{iht} is the transitory stochastic component of income; ϵ_{iht} is a mean-zero shock to the transitory component; $u_{iht,me}$ is a mean-zero measurement error; L is the lag operator so that $L^k x_t = x_{t-k}$, $\forall k = 0, \pm 1, \pm 2, \ldots$; and $\theta(L)$ is a moving average polynomial in L.

The income process outlined in equations (1)–(3) encompasses most of the income processes estimated in the literature.[3] Hause (1980), Lillard and Weiss (1979), and more recently Guvenen (2009) estimate the income process that is driven by "deterministic effects," α_i and β_i; an AR(1) transitory component affected each period by the transitory shock, ϵ_{iht}; and measurement error, $u_{iht,me}$. I label this process the HIP.[4] Meghir and Pistaferri (2004), Carroll (1992), and Carroll and Samwick (1997) are examples of the studies that assume the presence of a random walk and transitory components in idiosyncratic income but assume away (or present some evidence against) the deterministic idiosyncratic growth-rate component. I label this process the RIP. My ultimate goal is to determine whether the process containing random walk, transitory and deterministic components can be identified empirically.

2.1 Simulation Details

To see whether different processes are identified, I conduct a Monte Carlo study. I simulate data for a number of individuals "observed" for at most 30 periods using the data generating process of equations (1)–(3). I purposefully do not create a balanced panel data set—to mimic the patterns of the PSID data, which I will later use in empirical analysis. The PSID may contain at most 30 consecutive records on income for each head of household (from the 1968–1997 waves), but, since many heads first enter the labor market and the survey in different years and because of attrition and non-response, they contribute one or more observations on labor income.

The details of simulations are as follows. I assume that α_i and β_i are mean-zero, possibly correlated normally distributed fixed effects, the head is endowed with when he enters the labor market. I further assume that ξ_{iht} is an i.i.d. mean-zero shock to the permanent component of income nor-

[3] Most of the studies in the literature, to account for time-varying variances and covariances, allow for time-varying variances of stochastic (transitory and, if present, permanent) disturbances. This is the strategy I adopt in Section 3.2.

[4] Note that even though, say, Guvenen (2009) does not model the permanent stochastic component of income explicitly, he allows a root of the autoregressive representation of τ_{iht} to be one. The studies not modeling the permanent component explicitly find that the largest root of the stochastic component is below unity. They interpret this as the absence of the random walk component in idiosyncratic labor income, i.e., as if $p_{iht} = 0$ for all t.

mally distributed with the variance equal to σ_ξ^2; $\epsilon_{iht} \sim iidN(0, \sigma_\epsilon^2)$; $u_{iht,me} \sim iidN(0, \sigma_{u,me}^2)$; and τ_{iht} is a moving average process of order 1, an autoregressive process of order 1, or an ARMA(1,1) process. I use these particular representations of the transitory component of earnings for the following reasons. First, RIP studies, such as Abowd and Card (1989) and Meghir and Pistaferri (2004), find that the growth rate in male earnings can be represented by a moving average process of order 2, suggesting that the transitory component is a moving average process of order 1. Second, HIP studies, such as Lillard and Weiss (1979) and Guvenen (2009), model the transitory component as an autoregressive process of order 1. The estimated AR(1) process is easy to deal with in computational models featuring incomplete insurance markets and agents with idiosyncratic earnings histories, as argued in Guvenen (2009). Third, a moving average process of order 1 with the moving average parameter of a small magnitude is hard to distinguish from an autoregressive process of order 1.[5] Finally, an ARMA(1,1) transitory process encompasses the autocovariance structure of pure AR(1) and MA(1) processes and have been used, for example, in Baker (1997) and Haider (2001).

In the first year of a simulated data set, I observe a cross section of households whose heads' potential labor market experience ranges from one to 40 years, 70 of each type;[6] heads with one year of experience in the first sample year contribute 30 observations towards the final sample, while heads with 40 years of experience are observed in the first year only. In the second year, since all of the heads have one more year of experience and those with 40 years of experience exit the sample, I add 70 households whose heads just enter the labor market and have only one year of experience. I repeat these steps until I simulate a data set with the time dimension of 30 years. Further, following the literature, I keep information for those simulated individuals who contribute at least 9 consecutive observations on incomes towards the final sample.[7]

For each estimated income model, I report the results based on 100 simulated samples. The models are identified by fitting the theoretical autocovariances to the autocovariances in the simulated data. Estimation is performed using the minimum distance method, with the identity

[5] If the true transitory process is $\tau_{it} = (1+\theta L)\epsilon_{it}$, it can be represented by an infinite order autoregressive process, $\tau_{it} = \theta\tau_{it-1} - \theta^2\tau_{it-2} + \theta^3\tau_{it-3} - \ldots + \epsilon_{it}$, and approximated by $\tau_{it} = \theta\tau_{it-1} + \upsilon_{it}$, where $\upsilon_{it} = -\theta^2\tau_{it-2} + \theta^3\tau_{it-3} - \ldots + \epsilon_{it}$. Galbraith and Zinde-Walsh (1994) show that low-order autoregressive approximations of an MA(1) process—of order 1 up to order 3—with a moving average parameter of 0.5 and less in absolute value perform the best in terms of minimizing the mean squared error.

[6] This corresponds to ages 25–64 in the empirical analysis below.

[7] The requirement of having estimation samples with long (and often consecutive) spells of income observations adopted in the literature is not necessary for identification of income processes. I use this requirement to be consistent with other studies in the literature such as Meghir and Pistaferri (2004) and Guvenen (2009).

weighting matrix.[8] I now turn to estimation results for different simulated income processes.

2.2 Identification and Results from Simulated Data

In this section, I present estimation results on simulated data transformed into first differences. I first discuss identification of the processes containing a random walk component, a transitory component, a deterministic growth-rate component, and measurement error.

Note that both the variances and autocovariances estimated from data in levels contain contributions from the growth-rate heterogeneity, the permanent component, and the mean-reverting persistent component (see Appendix A). It is therefore challenging to tell apart the growth-rate heterogeneity from the permanent stochastic variation in incomes utilizing data in levels. The autocovariance function for the data in first differences, however, can be used to identify the growth-rate heterogeneity and random walk components, if both are present in the data. Permanent shocks will contribute only to the diagonal elements of the autocovariance function, i.e., the variances, while the growth-rate heterogeneity will contribute, in addition, towards all the off-diagonal elements of the autocovariance function. This information can be used to identify all the components as is shown in detail below.[9]

2.2.1 Identification

In this section, I provide the intuition behind identification of income processes that contain individual-specific growth rates, a permanent random walk and mean-reverting transitory components when the data used for estimation are in first differences. In the next section, I confirm identification using the minimum distance method, which utilizes all the available information in the autocovariance structure of the data. I present identification for income processes with transitory components modeled as an AR(1) or MA(1) processes. As emphasized above, those are the transitory processes commonly used in the HIP and RIP studies. Identification can be achieved for the income processes with a more general class of transitory components modeled as ARMA(p,q) processes.

Income Processes with Deterministic Growth-Rate Heterogeneity and a Random Walk Com-

[8] Altonji and Segal (1996) showed that an identity weighting matrix is the best choice for weighting the moments while estimating models of autocovariance structures on micro data with small samples. Most of the papers in the literature, guided by this result, utilize this weighting matrix.

[9] In real data, the results of estimations based on moments in levels may be affected by the chosen model of initial conditions. Estimations in first differences are not likely to depend on initial conditions as emphasized, for example, in Meghir and Pistaferri (2004).

ponent

In first differences, the process (1)–(3) is:

$$\Delta y_{it} = \beta_i + \xi_{it} + \theta(L)\Delta\epsilon_{it} + \Delta u_{it,me}, \tag{4}$$

where $\Delta \equiv 1 - L$.

First, assume that the transitory component is a moving average process of order 1, i.e., $\tau_{iht} = (1+\theta L)\epsilon_{iht}$.[10] The theoretical autocovariance moments, $\gamma_k = E\left[\Delta y_{it}\Delta y_{it-k}\right]$, of this process are:

$$\gamma_0 = \sigma_\xi^2 + \sigma_\beta^2 + (1+(1-\theta)^2+\theta^2)\sigma_\epsilon^2 + 2\sigma_{u,me}^2 \tag{5}$$

$$\gamma_1 = \sigma_\beta^2 - (\theta-1)^2\sigma_\epsilon^2 - \sigma_{u,me}^2 \tag{6}$$

$$\gamma_2 = \sigma_\beta^2 - \theta\sigma_\epsilon^2 \tag{7}$$

$$\gamma_k = \sigma_\beta^2, \quad k \geq 3. \tag{8}$$

The empirical variance-covariance matrix contains $\frac{T(T+1)}{2}$ unique moments. The variance of deterministic growth, σ_β^2, can be identified from the following vector of moments:

$$E\left[\Delta y_{it}\Delta y_{it+k}\right] = \sigma_\beta^2\mathbf{1}, \quad k = 3,\dots,T-t, \;\; t = 1,\dots,T-k, \tag{9}$$

where $\mathbf{1}$ is a vector of ones of the row dimension $\frac{(T-3)(T-2)}{2}$. Empirical analogs of the moments γ_0, γ_1, and γ_2 can be further used to identify the other four parameters: σ_ϵ^2, σ_ξ^2, $\sigma_{u,me}^2$, and θ. The variance of permanent shocks is uniquely identified; to identify the variances of transitory shocks and the moving average coefficient, however, one needs to restrict the variance of measurement error. In particular, the variance of permanent shocks can be identified from the following moment:

$$\hat{\sigma}_\xi^2 = E\left[\Delta y_{it}\Delta y_{it}\right] + 2E\left[\Delta y_{it}\Delta y_{it-1}\right] + 2E\left[\Delta y_{it}\Delta y_{it+1}\right] - 5\hat{\sigma}_\beta^2,$$

[10] Absent the growth-rate heterogeneity, the income process in first differences is a moving average process of order 2. This is consistent with the results in Abowd and Card (1989) and Meghir and Pistaferri (2004).

where $\hat{\sigma}^2_\beta$ is estimated using (9).

If $\tau_{iht} = (1 - \phi L)^{-1}\epsilon_{iht}$, i.e., the transitory component is an AR(1) process, the theoretical autocovariance moments of the income process in first differences are:

$$\gamma_0 = \sigma^2_\xi + \sigma^2_\beta + \frac{2}{1+\phi}\sigma^2_\epsilon + 2\sigma^2_{u,me} \tag{10}$$

$$\gamma_1 = \sigma^2_\beta - \frac{1-\phi}{1+\phi}\sigma^2_\epsilon - \sigma^2_{u,me} \tag{11}$$

$$\gamma_k = \sigma^2_\beta - \phi^{k-1}\frac{1-\phi}{1+\phi}\sigma^2_\epsilon, \quad k \geq 2. \tag{12}$$

Intuitively, σ^2_β should be identified from higher-order autocovariances—when the contribution of the transitory component towards the autocovariances approaches zero. In particular, the parameters for the transitory process can be identified using the following set of moments: $\hat{\phi} = \frac{\hat{\gamma}_{k+1} - \hat{\gamma}_k}{\hat{\gamma}_k - \hat{\gamma}_{k-1}}$, $k \geq 3$, $\hat{\sigma}^2_\epsilon = \frac{(\hat{\gamma}_{k+1} - \hat{\gamma}_k)(1+\hat{\phi})}{\hat{\phi}^{k-1}(1-\hat{\phi})^2}$, $k \geq 3$, where $\hat{\gamma}_k = E\left[\Delta y_{it} \Delta y_{it-k}\right]$. The variance of growth-rate heterogeneity, σ^2_β, can be then identified from the set of moments (12). Empirical analogs of the first two moments, γ_0 and γ_1, can be further used to uniquely identify the remaining model parameters—the variance of permanent shocks and the variance of measurement error.

If individual incomes contain transitory components modeled as an autoregressive moving average process, identification of the model parameters will be similar to identification of the processes containing moving average and autoregressive processes. In particular, the variance of measurement error, and the moving average parameter are not separately identified.

Summarizing, if the income process contains individual-specific growth rates and intercepts, a permanent random walk component, a mean-reverting transitory component, and measurement error, it is possible to identify the variance of permanent shocks and the variance of the deterministic growth-rate heterogeneity. I will confirm this intuition in estimations using simulated data in Section 2.2.2.

Income Processes with a Random Walk Component but No Deterministic Growth-Rate Heterogeneity—the RIP Processes

What if the true variance of the growth-rate heterogeneity is zero and the income process contains a random walk component but an econometrician estimates the HIP model instead?

In a time series context, the asymptotic variance of the scaled mean of a mean-zero stationary process, $\lim_{T\to\infty} E\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta y_t\right]^2$, will be equal to $\sum_{k=-\infty}^{\infty}\gamma(k)$, or the sum of the variance and twice the sum of the non-zero autocovariances if the autocovariances are absolutely summable.[11] The moment summarizes all the information contained in the autocovariance structure of the data. For the income model in equation (4) with $\sigma_\beta^2 = 0$ and the transitory component modeled as an MA(1), the asymptotic variance of the (scaled) sample mean is:[12]

$$\lim_{T\to\infty} E\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta y_t\right]^2 = \sigma_\xi^2. \tag{13}$$

The empirical analog of equation (13), for a covariance-stationary process, can be estimated from $\frac{1}{T}\left[T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2\right] = \sigma_\xi^2 + \frac{2}{T}\left[\sigma_\epsilon^2(1+\theta^2) + \sigma_{u,me}^2\right]$.[13] The estimated moment will be closer to σ_ξ^2 for a larger time dimension of the data, T. If, however, the random walk is ignored in estimation, the theoretical autocovariance function is non-zero beyond order 2 and is equal to σ_β^2. The moment in equation (13) will be estimated as $\frac{1}{T}[T\gamma_0 + 2(T-1)\gamma_1 + 2(T-2)\gamma_2 + \ldots + 2\gamma_{T-1}] = \frac{1}{T}\hat{\sigma}_\beta^2[T + 2(T-1) + 2(T-2) + \ldots + 4 + 2] + \frac{2}{T}[\hat{\sigma}_\epsilon^2(1+\hat{\theta}^2) + \hat{\sigma}_{u,me}^2] = T\hat{\sigma}_\beta^2 + \frac{2}{T}[\hat{\sigma}_\epsilon^2(1+\hat{\theta}^2) + \hat{\sigma}_{u,me}^2]$, where $\hat{\sigma}_\epsilon^2$, $\hat{\sigma}_{u,me}^2$, and $\hat{\theta}$ will differ from their true values.

The moment estimated in the data, generated by a model that contains a random walk and no growth-rate heterogeneity, should be replicated both by the true and misspecified models. Equating the two moments, one can show that if the true data generating process consists of a random walk, a persistent moving average component, and measurement error, and an econometrician estimates the (misspecified) HIP instead, the variance of the deterministic growth component will be approximately equal to:

$$\hat{\sigma}_\beta^2 \approx \frac{1}{T}\sigma_\xi^2. \tag{14}$$

[11] See, e.g., Hamilton (1994) Chapter 7 for a proof. This moment identifies the long-run variance of $\{\Delta y_t\}$. In the context of longitudinal data, γ_k will be zero for $k \geq \min\{T-1, H-1\}$, where H is the maximum age at which an individual can be observed in the data set and T is the time dimension of the data set. In the following, I will assume that $H > T$; this assumption is satisfied in the PSID data I will use, where $h = 1, \ldots, H = 40$ ($h = 1$ corresponds to age 25 ahd $h = H$ corresponds to age 64), and $T = 30$.

[12] In a time series context, if the transitory income component is an AR(1) or ARMA(1,1) process, the moment condition (13) will be the same; it will also identify σ_ξ^2 only.

[13] The population moment condition, for the transitory process modeled as an MA(1), will be equal to $\gamma_0 + 2\gamma_1 + 2\gamma_2$, and can be also estimated from $E\left[\Delta y_{it}\sum_{k=-2}^{k=2}\Delta y_{it+k}\right]$, the moment used by Meghir and Pistaferri (2004) to uncover the variance of the permanent shock.

Major micro data sets in the U.S. have no more than 30 years of consecutive observations on individual labor income. Thus, if the true variance of permanent shocks is equal to 0.02 and $T+1=30$, then the variance of the deterministic growth will be estimated at about 0.0007—within the bounds of the typical estimates of the HIP in the literature.

The same logic holds if the transitory stochastic component of income is an AR(1) process. If the true income process is the RIP with a permanent random walk component, the empirical analog of the moment in equation (13) is equal to $\frac{1}{T}[T\gamma_0+2(T-1)\gamma_1+2(T-2)\gamma_2+\ldots+2\gamma_{T-1}] = \sigma_\xi^2+\frac{2\sigma_\epsilon^2}{1+\phi}-\frac{2}{T}\frac{1-\phi}{1+\phi}\sum_{j=1}^{T-1}\phi^{j-1}(T-j)+\frac{2}{T}\sigma_{u,me}^2$. If the random walk is ignored and the HIP is estimated instead, the moment will be estimated as $T\hat{\sigma}_\beta^2+\frac{2\hat{\sigma}_\epsilon^2}{1+\hat{\phi}}-\frac{2}{T}\frac{1-\hat{\phi}}{1+\hat{\phi}}\sum_{j=1}^{T-1}\hat{\phi}^{j-1}(T-j)+\frac{2}{T}\hat{\sigma}_{u,me}^2$.

Thus, one should expect, for any mean-reverting transitory process, that the estimated variance of growth-rate heterogeneity is inversely related to the time dimension of the data set and directly related to the variance of permanent shocks, provided the true income process contains a random walk component and the estimated process is the HIP.

Income Processes with Deterministic Growth-Rate Heterogeneity but No Random Walk Components—the HIP Processes

What if the true income process is the HIP but an econometrician estimates the RIP model instead? If the transitory component is a moving average process of order 1, the econometrician will match the sample autocovariance moments to the misspecified population autocovariance moments, estimating σ_ξ^2, θ, and σ_ϵ^2. The model restricts the population autocovariance moments of order 3 and higher to zero, while the true population moments will be equal to σ_β^2. Heuristically, the matching procedure will look for σ_ξ^2, θ, σ_ϵ^2 that minimize the squared distance between T sample and theoretical zero-order autocovariances, $2(T-1)$ first-order autocovariances, and $2(T-2)$ second-order autocovariances. The estimate of the moment $E\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Delta y_{it}\right]^2$ for the misspecified model will be equal to $\hat{\sigma}_\xi^2+\frac{2}{T}\left[\hat{\sigma}_\epsilon^2(1+\hat{\theta}^2)+\hat{\sigma}_{u,me}^2\right]$, while the estimate of the same moment for the true model, with zero restrictions placed on the autocovariances above order 2, will be equal to $\frac{1}{T}\sigma_\beta^2(5T-6)+\frac{2}{T}\left[\sigma_\epsilon^2(1+\theta^2)+\sigma_{u,me}^2\right]$. Thus, the estimated variance of permanent shocks will be approximately equal to $\frac{1}{T}\sigma_\beta^2(5T-6)$. If, e.g., the true variance of deterministic growth-rate heterogeneity is equal to 0.0004, and the time dimension of the sample is 30 periods ($T+1=30$), the variance of permanent shocks will be approximately estimated at 0.00192, even though the true variance of these shocks is equal to zero. Similar arguments apply to the case when the transitory component is an AR(1) process—the estimated variance of permanent shocks will be non-zero when the true income process contains the deterministic idiosyncratic trends and an autoregressive transitory component, but an econometrician estimates the RIP process. The main lesson is that the estimated size of the permanent random walk component is expected to be small if the true process is the HIP but an econometrician estimates the RIP process with a random walk component.

2.2.2 Simulation Results

In this section, I present the results of estimations of income processes using simulated data transformed into first differences.

In Table 1, I estimate the process that contains the individual-specific intercepts and growth rates, a random walk component, a stationary component modeled as an ARMA(1,1) process, AR(1) process or a moving average process of order 1, and measurement error. The true values of the parameters in the models are: the variance of the growth-rate heterogeneity, $\sigma_\beta^2 = 0.0004$; the variance of permanent shocks, $\sigma_\xi^2 = 0.02$; the variance of the shocks to the transitory component, $\sigma_\epsilon^2 = 0.04$; the moving average parameter, $\theta = 0.50$ for a model with the transitory component modeled as an MA(1) process; the autoregressive parameter, $\phi = 0.50$ for a model with the transitory component modeled as an AR(1) process; $\phi = 0.50$, $\theta = -0.20$ for a model with the transitory component modeled as an ARMA(1,1) process; and the variance of measurement error, $\sigma_{u,me}^2 = 0.02$.[14] The chosen values for the true parameters are within the range of the values estimated in the literature. An individual with a (deterministic) growth rate one standard deviation above the mean will have a 2% advantage in income every year relative to an observationally equivalent individual whose income does not grow. One standard deviation in the permanent shock translates into a permanent change in income of about 14%. Note that when the transitory process contains a moving average component—columns (1) and (3)—the variance of measurement error is not identified.

Regardless of the model for the transitory component, the variance of growth-rate heterogeneity, σ_β^2, and the variance of permanent shocks, σ_ξ^2, are recovered without any biases and statistically precisely by the equally weighted minimum distance method. In columns (1) and (3), I set the variance of measurement error to the value that equals 25% of the true variance of the income growth rate.[15] The ratio of the assumed to the true variance of measurement error is about 75%.

[14] The results presented below are qualitatively similar when more or less persistent transitory processes are chosen in simulations.

[15] This is consistent with Meghir and Pistaferri (2004) and the findings in the literature on measurement error in longitudinal income data surveyed in Bound, Brown, and Mathiowetz (2001).

As a result, in column (1), the autoregressive persistence is estimated at the value close to its true value of 0.50, while the (absolute) value of the estimated moving average parameter and the variance of transitory shocks are larger than their true values. It is possible to identify all the parameters when the transitory component is an autoregressive process of order 1, which is confirmed in column (2). In column (3), the transitory component is modeled as an MA(1) process. Since the variance of measurement error set in estimations differs from its true value, the estimated moving average parameter is somewhat below its true value while the estimated variance of the shocks to the transitory process is slightly above its true value.[16] The last two rows in Table 1 report the median value of the goodness of fit statistics across 100 estimations, and the frequency of model rejections at the 1% significance level. The result in the last row is quite important. The size of the χ^2 test is severely distorted: instead of 1% rejections of the true model, the test rejects the true model more than 60% of the time. It appears that the χ^2 test of the model validity is not likely to be useful in empirical applications.[17]

In Table 2, I present the results for a HIP model that does not contain a random walk component, while in estimations I allow for a permanent component. Importantly, the variance of growth-rate heterogeneity is precisely recovered in estimations, while the variance of permanent shocks is small in magnitude and not statistically different from its true value of zero. This result has important implications for empirical analysis: if incomes differ over the life cycle due to deterministic growth-rate heterogeneity and are not affected by the shocks that persist over the entire life cycle, one can expect that the model that allows for both components will recover the true variance of growth-rate heterogeneity and a small and imprecise variance of the random walk shocks.[18]

In Table 3, the true income process contains the individual-specific intercept, a random walk component, a transitory component, and measurement error. The true variance of the random walk shock equals 0.02, while the true variance of deterministic growth-rate heterogeneity equals zero. The process is estimated as the (misspecified) HIP containing a deterministic growth-rate component and an unrestricted ARMA(1,1) process (column (1)), AR(1) process (column (2)) or

[16] If the variance of measurement error was set to its true value in estimations of the models in columns (1) and (3), all the model parameters would be estimated without any biases. The results are not reported for brevity.

[17] The size of the test will depend on the cross-sectional and time dimensions of the sample size.

[18] I also performed estimations of misspecified RIP income processes, when the true income processes contain deterministic growth-rate heterogeneity while the estimated models contain a random walk component and no idiosyncratic deterministic growth rates. The results, not reported here, largely confirm identification arguments outlined above. E.g., when the true income processes are such that the variance of deterministic growth-rate heterogeneity, σ^2_β, is equal to 0.0004, the transitory process is modeled as an MA(1) process with the moving average parameter equal to 0.50, and the estimated models ignore the growth-rate heterogeneity, the estimated variance of random walk shocks is equal to 0.0091. This is in line with the theoretical prediction calculated from the formula $\frac{1}{T}\sigma^2_\beta(5T-6)$, with $T+1$=30.

MA(1) process (column (3)). The variance of the shock to the transitory component is estimated at about 0.05 in columns (2) and (3) and 0.07 in column (1), while the estimated values of the autoregressive parameter are biased upwards—columns (1) and (2). Importantly, when the random walk component is ignored in estimation, the long-run persistence of the process is captured instead by the variance of the deterministic growth, with the estimated value substantially and significantly away from its true value of zero. When the model contains a moving average transitory process, this is the estimate one can expect given the time series dimension of 29 periods for income growth rates (see Section 2.2.1).

Guvenen (2009), in a simulation exercise, shows that the tests of higher-order autocovariances equal to zero will falsely reject the growth-rate heterogeneity even when the true income process contains idiosyncratic growth rates. This test was previously used by MaCurdy (1982). Table 4 confirms this result. I first create 1,000 samples generated in accordance with the models in Table 1, which contain deterministic idiosyncratic growth rates, a permanent random walk component, and a transitory stochastic process. For each simulated sample, I calculate the empirical autocovariance function. The results in the columns are the averages of the autocovariances of a given order across 1,000 simulated samples; standard errors, in parentheses, are calculated as the standard deviations of these estimates across 1,000 simulated samples. In column (1) the true process contains an ARMA(1,1) component; in column (2)—an AR(1) component; in column (3)—an MA(1) component. As can be seen from column (3), only autocovariances of orders 0, 1, and 2 are significant. The rest are insignificant, even though the magnitude of the autocovariances of orders 3 and higher will correctly identify the magnitude of the variance of the deterministic growth-rate heterogeneity. For the model with an AR(1) (ARMA(1,1)) transitory process, the autocovariance function is significant only from order 0 to order 4 (order 0 to order 3), inclusive; the contribution of the transitory component towards the autocovariance function dissipates quickly and higher-order autocovariances will, on average, correctly identify the size of the growth-rate heterogeneity. Intuitively, what matters for identification of the growth-rate heterogeneity is the average magnitude of higher-order autocovariances, and not the precision of higher-order autocovariances. The minimum distance procedure uses the entire autocovariance function, not only the information contained in higher-order autocovariances, and its sample variability to uncover correctly and precisely the variance of the deterministic growth-rate heterogeneity—Table 1, columns (1)–(3).

Summing up, using data in first differences it is feasible to identify the variance of the shock to the random walk component, the variance of the shock to the transitory component, and the variance of the growth-rate heterogeneity. If, however, the process contains a random walk permanent component and no deterministic growth-rate component and the model is estimated as the HIP, estimation will capture the long-run variance of income growth due to permanent shocks with a significant estimate of the growth-rate heterogeneity. The magnitude of the estimate will depend on the true variance of the permanent shock and the time dimension of a data set.

3 Empirical Results

In this section, I estimate time series processes for idiosyncratic labor incomes of male household heads from the PSID. I first describe the data I utilize.

3.1 Data

I use income and demographic data from the 1968–1997 waves of the PSID. I select male household heads of ages 25–64.[19] I further drop heads with inconsistent education records and split the sample into two education groups. The first group comprises heads who dropped out of high school or just finished high school. The second group includes heads who finished some college, graduated from college, or attained a graduate degree. I will label the second group college graduates.[20]

The measure of income utilized is the head's labor income from all sources, inclusive of the labor part of farm and business income. Income data in the PSID refer to the previous calendar year; I adjust them appropriately by the consumer price index for all items normalized to 100 in 1982–1984. I set income observations to missing when the head reports to be a student or self-employed and in the year subsequent to that report. When the head reports to be retired, I set his income in that year and all subsequent years to missing. I further drop observations for the years when the percentage change of real labor income in adjacent years is above 500 or below –80. I then drop observations with zero, top-coded, and missing incomes, and select the longest consecutive spell of positive incomes with at least 9 observations. I exclude data for households from the Survey of Economic Opportunity (SEO) subsample, which over-samples the poor. Figure 1 plots the time profile of the variances of log-labor income for different PSID samples—the core subsample, the SEO subsample and the sample comprising those two subsamples. The variances in the core

[19] Age in the PSID does not necessarily change in adjacent surveys since information can be collected at different months of a year. Also, some individuals have inconsistent age series which, among other things, may reflect typing errors by interviewers. I utilize information on the year of birth to construct a cleaner measure of age for those heads who have this information in the individual file. Otherwise, I use an individual's age at the time he first appears as a head in the survey to impute his age in other years.

[20] Each education group roughly comprises a 50% of the sample.

sample—which was representative of the U.S. population in 1968, at the start of the survey—have been increasing during the last three decades, with a sharp increase in the beginning of the 1980s. The variances for the SEO subsample follow a somewhat different time pattern: the variance increased in the beginning of the 1980s but started declining afterwards. For this reason, and following most of the literature, I exclude the SEO subsample from my empirical analysis.[21] The final sample contains information for 1,916 heads with 29,753 person-year observations on labor incomes. Table 5 contains some descriptive sample statistics for select years.

The measure of the idiosyncratic head's labor income in each year is the head's residual from a cross-sectional regression of the first difference in log-labor income on a third polynomial in age, "college" dummy, and interactions between the "college" dummy and the age polynomial. This regression, specifying the deterministic component of incomes common to all heads, is similar to specifications adopted in the literature and assumes that returns to the head's experience and education are affected by the aggregate state of the economy, that is, differ by year.

Table 6 contains the results of the tests of the autocovariances of a given order being zero in all time periods. One cannot reject the null that the autocovariances of orders 4 and 5 are equal to zero. However, the null that the autocovariances of order 3 and higher, or order 4 and higher are all equal to zero can be rejected. The results can be consistent with a model containing an AR(1) or ARMA(1,1) transitory component with a small autoregressive persistence (see, e.g., the autocovariance function in Table 4 column (1)).[22]

3.2 Results

Table 7 contains my main results.[23] The models are estimated by fitting the empirical autocovariance function to the theoretical autocovariance function, utilizing the identity weighting matrix, i.e., by the equally weighted minimum distance method.

In column (1), I estimate the HIP process, which ignores the potentially important random walk component in idiosyncratic labor incomes. The variance of individual-specific growth rates is

[21]The pattern of the variances in the sample comprising both the SEO and core subsamples is similar to the pattern of variances in Figure 2 of Meghir and Pistaferri (2004) for their whole sample.

[22]The results are somewhat different from Meghir and Pistaferri (2004) who find, for their pooled sample, that the autocovariances of orders 3 and 4 are not statistically different from zero, and a p-value of 12% for the test that the autocovariances of order 3 and higher are all equal to zero. Their results are based on PSID data up to 1993 and a sample that includes the SEO subsample. If I ignore the data after 1993, my results for the tests in Table 6 are similar to the ones in Table II in Meghir and Pistaferri (2004).

[23]Most of the studies in the literature allow for time-dependent variances of permanent and/or persistent shocks. My estimates of these parameters in Table 7, columns (1)–(4), (6) should be interpreted as the unconditional variances of transitory and permanent shocks.

estimated at 0.0004, significant at the 1% level, while the persistence of the transitory component is moderate—the autoregressive parameter is estimated at about 0.70.[24]

In column (2), I allow for both a random walk and a deterministic growth-rate component in earnings. Monte Carlo results and theoretical arguments spelled out in Section 2.2.1 indicated that, if both these components are present, the process should be empirically identified. In column (2), the estimate for the variance of the individual-specific growth rates binds at zero while the estimate of the variance of the shock to the random walk component equals 0.015 and is significant at the 1% level. An autoregressive parameter of the transitory process is estimated at about 0.37, capturing the fast decline of the empirical autocovariance function of labor income growth rates beyond the first order. The variance of the shocks to the transitory component, which also comprises the contribution of measurement error, is estimated at about 0.03.

In column (3), I do not restrict the autoregressive parameter of a more persistent process to equal 1. The results are largely similar to those in column (2), but less precise. The estimated variance of growth-rate heterogeneity is still zero, while the autoregressive parameter of a more persistent process is close to 1. Column (4) reports the results of the same model, when the variance of growth-rate heterogeneity is restricted to equal zero. The results are quantitatively similar to those in column (3), while the parameters are estimated more precisely.

In column (5), I re-estimate the model of column (4), allowing for time-varying permanent and transitory variances. I report the time-averages of the estimated variances and the time-averages of their standard errors. The results are similar to those in columns (2)–(4). The full sets of the transitory and permanent variances, along with their standard errors, are presented in Table 8.

The off-diagonal elements of the empirical autocovariance matrix contain important information for identification of the variance of the growth-rate heterogeneity. For example, if the true income process contains a moving average process of order 1, higher-order autocovariances will be informative for identification of the variance of the growth-rate heterogeneity. The number of heads contributing towards the empirical autocovariance $\hat{\gamma}_k$ is, in general, smaller the larger the lag length k is, which separates the head's income observation at time t from the income observation at time $t+k$. Placing an equal weight on all the variances and autocovariances in estimation may bias an estimate of the growth-rate heterogeneity towards zero if higher-order empirical autocovari-

[24] The autoregressive parameter is estimated at about the same value if the transitory component is modeled as an AR(1) process. I choose to report the results for a model with an ARMA(1,1) transitory component since its autocovariance function encompasses the autocovariance function of both AR(1) and MA(1) transitory processes. The main result—that the variance of permanent shocks is significant and the estimated variance of the growth-rate heterogeneity is zero—holds for models with transitory components modeled as an MA(1) or AR(1) processes.

ances are very close to zero and imprecisely estimated as, indeed, is found in empirical data. To take care of this concern, following Guvenen (2009), I re-estimate the model utilizing only the first 10 empirical autocovariances and all the variances in estimation—column (6) Table 7. The main result remains unaltered: the growth-rate heterogeneity is estimated at zero while the variance of permanent shocks is precisely estimated at about 0.015.

In Figure 2, I plot the resulting time series of the estimated variances of permanent and transitory shocks. It appears that an increase in the variance of heads' incomes in early 1980s depicted in Figure 1 was largely due to the increase in the variance of permanent shocks during that period. The pattern of the variances of permanent shocks in the 1980s mirrors that in Meghir and Pistaferri (2004), for their pooled sample that includes heads of household from the SEO subsample. It is also qualitatively similar to the hump-shaped pattern of the permanent volatility of *household* incomes in the 1980s reported in Blundell, Pistaferri, and Preston (2008). The variance of transitory shocks to the heads' incomes does not show a clear trend for almost two decades.[25]

For robustness, I estimate the income process using separate samples for heads who dropped out from or just finished high school (columns (1)–(2) in Table B-1), and those who finished some college, graduated from college or have education levels beyond college degree (columns (3)–(4) in Table B-1). This is roughly a 50-50 split of the main sample that allows to precisely estimate the model parameters.

When the random walk component is ignored in estimation, the variance of the deterministic growth-rate heterogeneity is substantial and statistically significant (columns (1) and (3) in Table B-1); the estimated AR(1) persistence of the stochastic component is moderate, ranging from about 0.85 for the college sample to 0.59 for high school graduates/drop-outs. Similar to the main results, in the model that includes both the growth-rate heterogeneity and permanent random walk component, the estimated variance of the growth-rate heterogeneity equals zero—columns (2) and (4). Interestingly, the estimated variance of permanent shocks is higher for more educated heads, while the estimated variance of the shocks to the transitory component is higher for less educated heads.

Discussion in Section 2.2.1 suggested that the estimated variance of the growth-rate heterogeneity should be inversely related to the time dimension of the sample size if the true process contains a random walk component and the estimated process is estimated as HIP. To support this

[25]In 1993, the PSID switched to the electronic data collection method. Presumably, the spike of the variance of transitory shocks in 1993—see the vertical line in Figure 2—is due to this change in the data collection method. The results are qualitatively similar if I do not use the data after 1992.

result, I simulated the model of column (2) in Table 7 for samples with different time dimension but the same cross-sectional dimension. The misspecified HIP model always returns non-zero and significant estimates of the growth-rate heterogeneity—column (2) of Table 9—which are higher for samples with a smaller time dimension. I also performed simulations of the model in column (1) of Table 7. The results are in column (1) of Table 9. If the true model is HIP, the variance of growth-rate heterogeneity should not depend on the time dimension of the sample size, as the results in column (1) suggest. If, for the same cross-section of heads, an estimate of the growth-rate heterogeneity is found to be systematically different for different time dimensions of the sample, this will present some additional evidence against the hypothesis that heads' idiosyncratic income growth rates systematically and deterministically differ over the life cycle.

Figure 3 graphically presents just outlined arguments using PSID data. First, I select a sample of 1,157 PSID heads who have at least 5 consecutive income observations during 1968–1977 and estimate the HIP process for idiosyncratic incomes for that sample. This gives me the first point in the graph in Figure 3, panel (a). I then extend the time dimension of the initial sample to 1978, keeping the cross-sectional dimension fixed, and estimate the HIP process for that sample. I continue this procedure until I arrive at the sample that spans the period 1969–1997 for those 1,157 heads, the longest possible period. The results in Figure 3 are quite telling: it appears that the estimated growth-rate heterogeneity is larger for smaller time dimensions of the sample size, even though the cross-sectional dimension is the same and one would expect the estimated growth-rate heterogeneity to be the same. Next, I perform the same series of estimations, assuming that the true process contains a random walk and the transitory process modeled as an ARMA(1,1) process. The results are plotted in panel (b) of Figure 3. One could think that those are the unconditional estimates of the permanent variance for different time dimensions: the average permanent variance will increase if the marginal variance is higher than the average variance, and vice versa. In panel (c) of Figure 3, I divide the estimated permanent variances in the rightmost panel of Figure 3 by the time dimension of the sample utilized for their estimation. Remarkably, the series of the estimated growth-rate heterogeneity in the leftmost graph are quite similar to the series of the variance of permanent shocks scaled by the inverse of the time dimension of the sample, as one would expect if the true process contains a random walk component and no growth-rate heterogeneity.

There is some evidence, not relying on estimation of income processes, interpreted by some researchers as favoring income models with heterogeneous income profiles. Haider and Solon (2006) and Böhlmark and Lindquist (2006) study the association between current and lifetime income over the life cycle for U.S. and Swedish samples, respectively. Specifically, they focus on the life-cycle variation in the slope coefficient from the following regression: $y_{ia} = \beta_a V_i + \epsilon_{ia}$, where y_{ia} is individual i's log-income at age a, V_i is individual i's log-lifetime income, calculated as (the log of) the annuity value of the discounted sum of annual real incomes observed for individual i, and ϵ_{ia} is individual i's regression error at age a. Haider and Solon (2006) find that β_a is estimated at about 0.20 at age 19, steadily increases afterwards, equals one at age 34 and levels off for the rest of the life cycle. Böhlmark and Lindquist (2006), for a much larger Swedish sample, find that $\hat{\beta}_a$ starts at about 0.20 at age 19, crosses one at age 34 and peaks at 1.45 at age 48. The latter authors interpreted this result as evidence favoring the presence of heterogeneous income profiles—income is low in the beginning of the life cycle and is well below the lifetime income (which is estimated to be time-invariant by the authors); income then steadily grows until it exceeds the lifetime income in the later part of the life cycle. This result is, however, also true for income processes that contain random walks and do not have deterministic idiosyncratic trends. Using the estimates of the RIP process in this paper, I was able to replicate, in simulations not reported here, the pattern of $\hat{\beta}_a$'s found in Böhlmark and Lindquist (2006). The intuition behind this result is the following. Note that $\hat{\beta}_a = \frac{cov(y_{ia}, V_i)}{var(V_i)}$. While the denominator is constant over the life cycle, the cross-sectional covariance between current incomes and lifetime incomes will be growing over the life cycle since current incomes will accumulate random walk shocks over the life cycle and will, therefore, co-vary more strongly with lifetime incomes, which aggregate all the permanent shocks to individual incomes over the entire life cycle.

Summarizing, for the samples utilized in this study, it appears that I can reject the HIP model. The RIP model with a permanent random walk component and a transitory mean-reverting component cannot be rejected.

4 Conclusion

I estimate idiosyncratic labor income processes on simulated and empirical data. The main results of a Monte Carlo study using unbalanced panel data are the following. It is possible to identify a general process containing all the elements of the HIP and RIP models. The most important elements are the growth-rate heterogeneity and the variance of a random walk component. For simulated data in first differences I show that both these elements, if present, should be recovered precisely in empirical estimations. The results on simulated data confirm another important finding of this paper: if the true income process is the sum of a random walk and persistent components, i.e., the RIP, and the random walk is ignored in estimation, the misspecified HIP model recovers significant and substantial growth-rate heterogeneity and modest persistence.

I use data for male household heads from the 1968–1997 waves of the PSID to estimate idiosyncratic labor income processes. I find that the estimated variance of the deterministic growth-rate heterogeneity is zero; i.e., the HIP model can be rejected. The RIP model, with permanent random walk and mean-reverting components, cannot be rejected. I find that the estimated variance of the permanent component is significant and substantial. Thus, the results of the paper favor the view that the observed variation in idiosyncratic income growth rates over the life cycle is entirely due to the shocks of different "durability."

The results of this paper are important for understanding a number of issues. Among them are the choice of an appropriate model of the heterogeneity in individual and household idiosyncratic incomes used in macro models; the importance of shocks versus initial conditions for the life-cycle profiles of earnings and welfare inequality; and the importance of shocks versus initial conditions for the time series of earnings and consumption inequality. The process, best fitting the data utilized in this paper, places restrictions on the models that make earnings an endogenous variable.

In this paper, I utilize only income data to identify the variances of idiosyncratic permanent and transitory shocks. Perhaps, more accurate estimates of the variances could be obtained by jointly studying household choices and income data. For recent attempts at this approach see Hryshko (2007) and Blundell, Pistaferri, and Preston (2008) (in the context of RIP), and Guvenen and Smith (2008) (in the context of HIP) who utilize data on household income and consumption choices.

References

Abowd, J., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

Altonji, J. G., and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14(3), 353–366.

Baker, M. (1997): "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics*, 15(2), 338–375.

Baker, M., and G. Solon (2003): "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records," *Journal of Labor Economics*, 21(2), 289–321.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5), 1887–1921.

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

Bound, J., C. Brown, and N. Mathiowetz (2001): "Measurement Error in Survey Data," in *Handbook of Econometrics, Vol.5*, chap. 59, pp. 3705–3843. North-Holland, New York.

Carroll, C. D. (1992): "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 2, 61–156.

Carroll, C. D., and A. A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111(4), 818–857.

De Santis, M. (2007): "Individual Consumption Risk and the Welfare Cost of Business Cycles," *American Economic Review*, 97(4), 1488–1506.

Galbraith, J., and V. Zinde-Walsh (1994): "A Simple Noniterative Estimator for Moving Average Models," *Biometrika*, 1, 143–155.

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3), 687–712.

——— (2009): "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1), 58–79.

Guvenen, F., and A. Smith (2008): "Inferring Labor Income Risk from Economic Choices: an Indirect Inference Approach," Mimeo, University of Minnesota.

Haider, S., and G. Solon (2006): "Life-Cycle Variation in the Association Between Current and Lifetime Earnings," *American Economic Review*, 96(4), 1308–1320.

Haider, S. J. (2001): "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics*, 19(4), 799–836.

Hamilton, J. (1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

Hause, J. C. (1980): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica*, 48(4), 1013–1029.

Hryshko, D. (2007): "Identifying Household Income Processes Using a Life Cycle Model of Consumption," Mimeo, University of Alberta.

Huggett, M., G. Ventura, and A. Yaron (2007): "Sources of Lifetime Inequality," NBER Working Paper # 13224.

Krebs, T. (2003): "Human Capital Risk and Economic Growth," *Quarterly Journal of Economics*, 118(2), 709–744.

Lillard, L. A., and Y. Weiss (1979): "Components of Variation in Panel Earnings Data: American Scientists 1960–1970," *Econometrica*, 47(2), 437–454.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

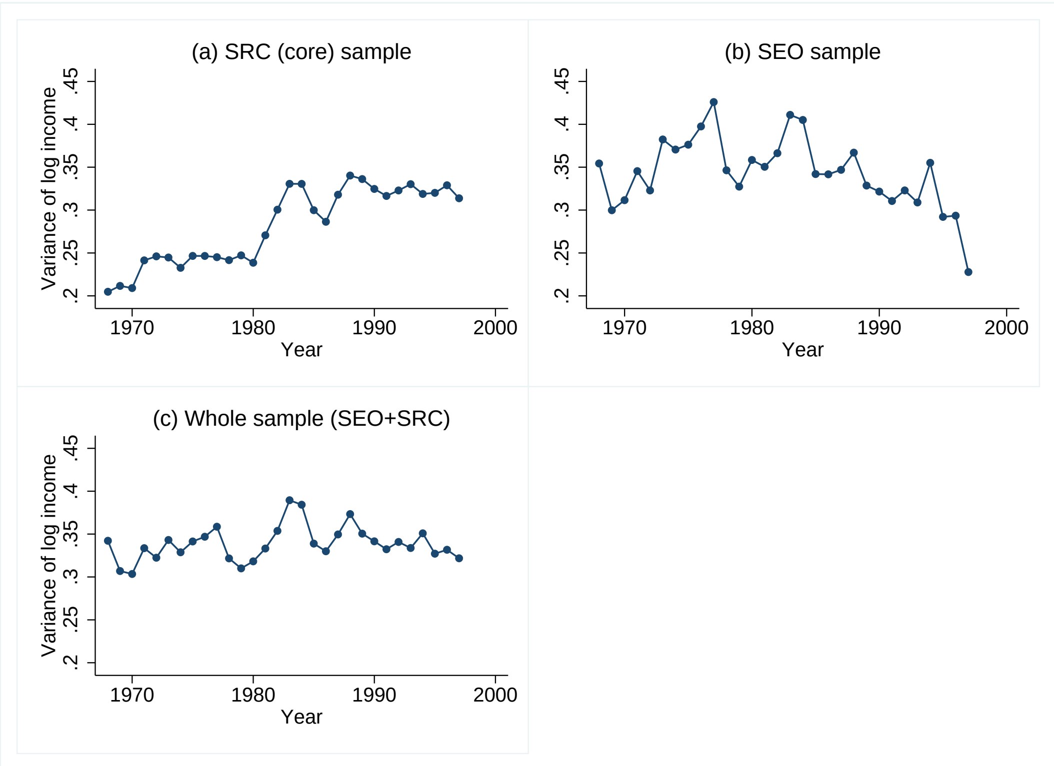
Moffitt, R. A., and P. Gottschalk (1995): "Trends in the Covariance Structure of Earnings in the U.S.: 1969–1987," Johns Hopkins University, Mimeo.

Newey, W. K. (1985): "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 29, 229–256.

Scholz, J. C., A. Seshadri, and S. Khitatrakun (2006): "Are Americans Saving "Optimally" for Retirement?," *Journal of Political Economy*, 114(4), 607–643.

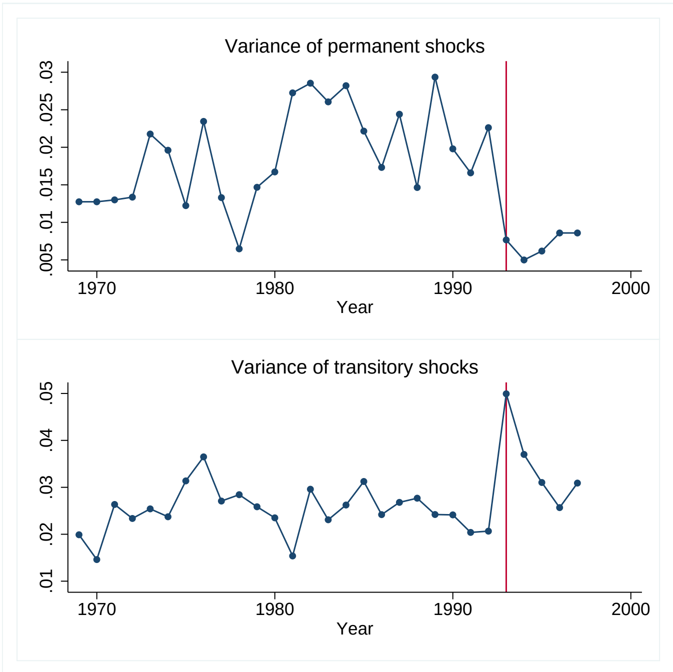
Storesletten, K., C. I. Telmer, and A. Yaron (2004): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–633.

Figure 1: The Variance of Log Labor Income by Year



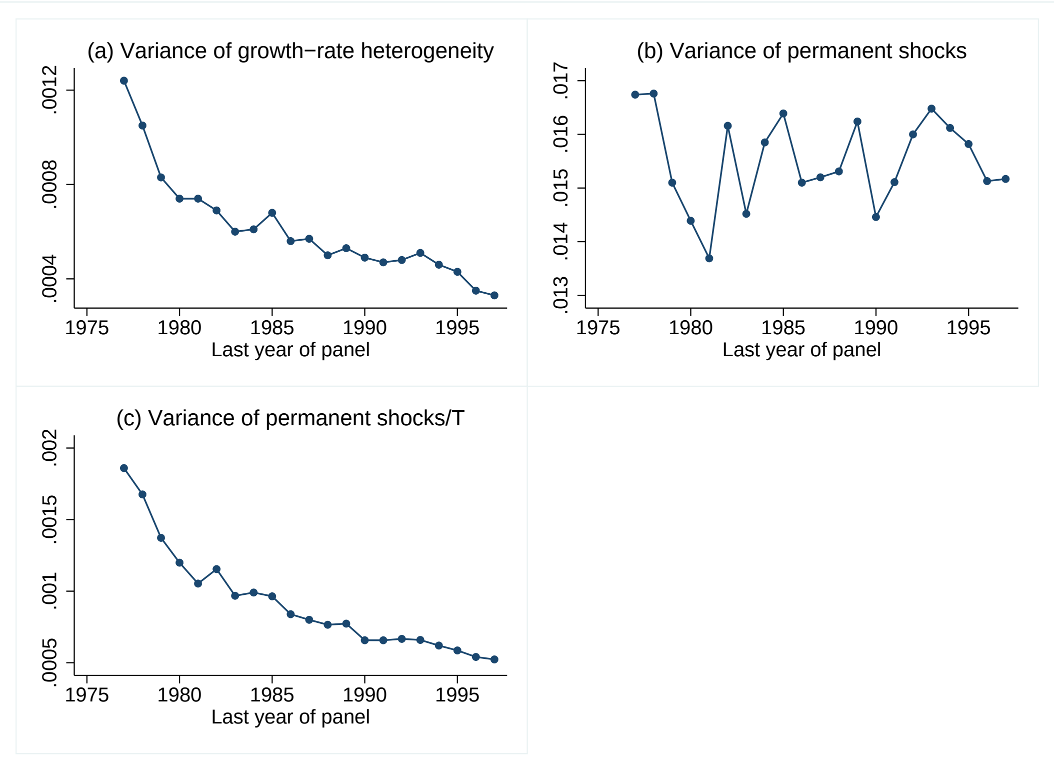
Notes: In panel (a), the graph depicts the variances of log labor income for the main sample that includes heads from the core subsample only. In panel (b), the graph depicts the variances of log labor income for a sample that includes heads from the SEO subsample only. In panel (c), the graph depicts the variances of log labor income for a sample that includes heads from the core and SEO subsamples. All samples include heads with consecutive spells of at least 9 income observations only.

Figure 2: The Variance of Shocks to Labor Income by Year



Notes: The variances of permanent and transitory shocks are estimated on the main sample fitting the model in Table 7, column (5). The vertical line is drawn for the survey year 1993, when the PSID switched to the electronic data collection method.

Figure 3: The Variance of Growth-rate Heterogeneity and the time dimension of the sample, T



Notes: In panel (a), the first point on the graph is an estimate of the variance of growth-rate heterogeneity for a sample of 1,157 PSID heads with at least 5 consecutive income observations during 1968–1977; all subsequent points are the estimates of the variance of growth-rate heterogeneity using samples that contain the same heads and income information for 1968–1978, 1968–1979, all the way up to 1968–1997. In panel (b), the same samples are used to estimate the income process that includes a permanent random walk component and an ARMA(1,1) transitory process. The estimates of the variance of permanent shocks are then divided by the time dimension of the sample size to produce the graph in panel (c).

Table 1: Estimates of the HIP with a Random Walk Component: Simulated Data in First Differences.

Parameters/Trans. comp.	ARMA(1,1) σ^2_β=0.0004, σ^2_ξ=0.02	AR(1) σ^2_β=0.0004, σ^2_ξ=0.02	MA(1) σ^2_β=0.0004, σ^2_ξ=0.02
Heterog. growth, $\hat{\sigma}^2_\beta$	0.0004	0.0004	0.0004
	(0.00007)	(0.00007)	(0.00005)
Var. perm. shock, $\hat{\sigma}^2_\xi$	0.02	0.02	0.02
	(0.002)	(0.002)	(0.0009)
AR, $\hat{\phi}$	0.495	0.497	—
	(0.079)	(0.041)	—
MA, $\hat{\theta}$	–0.223	—	0.432
	(0.049)	—	(0.007)
$\hat{\sigma}^2_\epsilon$	0.046	0.04	0.046
	(0.002)	(0.002)	(0.0008)
$\sigma^2_{u,me}$	0.015	0.02	0.015
	—	(0.002)	—
Median χ^2[d.f.]	528.89 [430]	520.93 [430]	527.10 [431]
Rejection rate at 1%	74%	67%	72%

Notes: The true income process is: $y_{iht} = \alpha_i + \beta_i h + p_{iht} + \tau_{iht} + u_{iht,me}$, with $(1-L)p_{ih+1t+1} = \xi_{ih+1t+1}$, $\sigma^2_\alpha = 0.03$, σ^2_β=0.0004, $\sigma^2_\xi = 0.02$, and $\sigma^2_{u,me} = 0.02$. In the first column, the transitory process is modeled as $\tau_{iht} = \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$, ϕ=0.50, $\theta = -0.20$. In the second column, $\tau_{iht} = \frac{\epsilon_{iht}}{1-\phi L}$, ϕ=0.50. In the third column, $\tau_{iht} = (1+\theta L)\epsilon_{iht}$, θ=0.50. In all models, the true variance of the shocks to the transitory component is σ^2_ϵ=0.04. Prior to estimation, simulated data are transformed to first differences; models are estimated by the equally weighted minimum distance method. Standard errors in parentheses calculated as the standard deviations of the estimates across 100 model simulations. Goodness of fit statistic is based on Newey (1985).

Table 2: Estimates of the HIP with no Random Walk Component: Simulated Data in First Differences.

Parameters/Trans. comp.	ARMA(1,1) σ^2_β=0.0004, σ^2_ξ=0	AR(1) σ^2_β=0.0004, σ^2_ξ=0	MA(1) σ^2_β=0.0004, σ^2_ξ=0
Heterog. growth, $\hat{\sigma}^2_\beta$	0.00038	0.00038	0.00038
	(0.00004)	(0.00004)	(0.00003)
Var. perm. shock, $\hat{\sigma}^2_\xi$	0.00045	0.00036	0.00019
	(0.001)	(0.001)	(0.0007)
AR, $\hat{\phi}$	0.480	0.492	—
	(0.068)	(0.034)	—
MA, $\hat{\theta}$	–0.233	—	0.405
	(0.046)	—	(0.006)
$\hat{\sigma}^2_\epsilon$	0.047	0.04	0.049
	(0.002)	(0.001)	(0.0007)
$\sigma^2_{u,me}$	0.013	0.02	0.013
	—	(0.002)	—
Median χ^2[d.f.]	508.95 [430]	528.67 [430]	529.93 [431]
Rejection rate at 1%	60%	73%	75%

Notes: The true income process is: $y_{iht} = \alpha_i + \beta_i h + \tau_{iht} + u_{iht,me}$, with $\sigma^2_\alpha = 0.03$, σ^2_β=0.0004, $\sigma^2_\xi = 0.02$, and $\sigma^2_{u,me} = 0.02$. In the first column, the transitory process is modeled as $\tau_{iht} = \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$, ϕ=0.50, $\theta = -0.20$. In the second column, $\tau_{iht} = \frac{\epsilon_{iht}}{1-\phi L}$, ϕ=0.50. In the third column, $\tau_{iht} = (1+\theta L)\epsilon_{iht}$, θ=0.50. In all models, the true variance of the shocks to the transitory component is σ^2_ϵ=0.04. Prior to estimation, simulated data are transformed to first differences; models are estimated by the equally weighted minimum distance method. Standard errors in parentheses calculated as the standard deviations of the estimates across 100 model simulations. Goodness of fit statistic is based on Newey (1985).

Table 3: Estimates of the Misspecified HIP: Simulated Data in First Differences.

Parameters/Trans. comp.	ARMA(1,1) σ^2_β=0, σ^2_ξ=0.02	AR(1) σ^2_β=0, σ^2_ξ=0.02	MA(1) σ^2_β=0, σ^2_ξ=0.02
Heterog. growth, $\hat{\sigma}^2_\beta$	0.00052	0.00056	0.0007
	(0.00003)	(0.00003)	(0.00003)
Var. perm. shock, $\hat{\sigma}^2_\xi$	0.00	0.00	0.00
	—	—	—
AR, $\hat{\phi}$	0.774	0.685	—
	(0.014)	(0.013)	—
MA, $\hat{\theta}$	–0.262	—	0.423
	(0.012)	—	(0.006)
$\hat{\sigma}^2_\epsilon$	0.071	0.054	0.058
	(0.0006)	(0.001)	(0.0005)
$\sigma^2_{u,me}$	0.015	0.024	0.015
	—	(0.001)	—
Median χ^2[d.f.]	612.25 [431]	678.75 [431]	2004.21 [432]
Rejection rate at 1%	100%	100%	100%

Notes: The true income process is: $y_{iht} = \alpha_i + p_{iht} + \tau_{iht} + u_{iht,me}$, with $(1 - L)p_{ih+1t+1} = \xi_{ih+1t+1}$, $\sigma^2_\alpha = 0.03$, $\sigma^2_\xi = 0.02$, and $\sigma^2_{u,me} = 0.02$. In the first column, the transitory process is modeled as $\tau_{iht} = \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$, ϕ=0.50, $\theta = -0.20$. In the second column, $\tau_{iht} = \frac{\epsilon_{iht}}{1-\phi L}$, ϕ=0.50. In the third column, $\tau_{iht} = (1 + \theta L)\epsilon_{iht}$, θ=0.50. In all models, the true variance of the shocks to the transitory component is σ^2_ϵ=0.04. Prior to estimation, simulated data are transformed to first differences; models are estimated by the equally weighted minimum distance method. Standard errors in parentheses calculated as the standard deviations of the estimates across 100 model simulations. Goodness of fit statistic is based on Newey (1985).

Table 4: Autocovariances for Income Growth Rates. Income Processes with Growth-Rate Heterogeneity and a Random Walk Component. Simulated Data.

Order	τ_{iht} ~ARMA(1,1)	τ_{iht} ~AR(1)	τ_{iht} ~MA(1)
0	0.12116	0.11365	0.12017
	(0.00069)	(0.00063)	(0.00070)
1	–0.04283	–0.03300	–0.02958
	(0.00051)	(0.00046)	(0.00042)
2	–0.0032	–0.00626	–0.01962
	(0.0005)	(0.00045)	(0.00050)
3	–0.0014	–0.00296	0.00041
	(0.00053)	(0.00048)	(0.00052)
4	–0.00048	–0.00125	0.00039
	(0.00057)	(0.00050)	(0.00052)
5	–0.00009	–0.00047	0.00039
	(0.00058)	(0.0005)	(0.00055)
10	0.00037	0.00041	0.00037
	(0.00069)	(0.00063)	(0.00066)
15	0.00039	0.00041	0.00043
	(0.0009)	(0.00081)	(0.00085)
20	0.00042	0.00038	0.00039
	(0.00123)	(0.00113)	(0.00117)

Notes: The true income process is: $y_{iht} = \alpha_i + \beta_i h + p_{iht} + \tau_{iht} + u_{iht,me}$, with $(1-L)p_{ih+1t+1} = \xi_{ih+1t+1}$, $\sigma^2_\alpha = 0.03$, $\sigma^2_\xi = 0.02$, and $\sigma^2_{u,me} = 0.02$. In the first column, $\tau_{iht} = \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$, ϕ=0.50, $\theta = -0.20$. In the second column, $\tau_{iht} = \frac{\epsilon_{iht}}{1-\phi L}$, ϕ=0.50. In the third column, the transitory process is modeled as $\tau_{iht} = (1+\theta L)\epsilon_{iht}$, θ=0.50. The true variance of the shocks to the transitory component is σ^2_ϵ=0.04. Simulated data are transformed to first differences. Autocovariances of a given order are the averages of the autocovariances in simulated data across 1,000 simulations. Standard errors in parentheses calculated as the standard deviations of the estimated autocovariances of a given order across 1,000 model simulations.

Table 5: Sample Statistics for Select Years.

	Year			
	1970	1980	1990	1997
Age	40.67	41.11	39.71	44.37
	(8.50)	(10.83)	(9.06)	(7.46)
Hours	2,258	2,224	2,270	2,281
	(460)	(496)	(498)	(495)
Nonfinancial income[a]	33,127	36,289	39,371	45,193
	(13,851)	(15,901)	(24,914)	(27,340)
White	0.89	0.91	0.93	0.93
Married	0.96	0.90	0.87	0.88
≥13 yrs. of schooling	0.35	0.46	0.53	0.56

Notes: [a] Nonfinancial income is the sum of head's and wife's real labor income from all sources, and their combined transfer income expressed in 1982–1984 dollars. The income measure excludes head's and wife's social security income. Standard deviations in parentheses.

Table 6: Test of the Null Hypothesis of Zero Autocovariance in All Time Periods.

Order	Test stat.	d.f.	p-value
1	506.81	28	0.00
2	57.88	27	0.00
3	45.35	26	0.01
4	23.65	25	0.54
5	23.76	24	0.48
≥3	465.96	351	0.00
≥4	425.18	325	0.00

Notes: The test statistic is distributed as χ^2 with degrees of freedom equal to the number of (zero) restrictions (the number of unique autocovariances of a given order in the estimated variance-covariance matrix).

Table 7: Estimates of Income Processes. PSID Data in First Differences.

	(1) HIP	(2) add RW	(3) est. pers.	(4) same as (3), set σ_β^2=0	(5) chang. perm./ trans. var.	(6) use only first 10 acfs
$\hat{\sigma}_\beta^2$	0.0004	0.00	0.00	0.00	0.00	0.00
	(0.00004)	(0.00006)	(0.001)	—	—	(0.0002)
$\hat{\sigma}_\xi^2$	0.00	0.015	0.016	0.016	0.017	0.015
	—	(0.002)	(0.002)	(0.002)	(0.005)	(0.003)
$\hat{\phi}$	0.712	0.367	0.343	0.343	0.357	0.369
	(0.029)	(0.115)	(0.194)	(0.124)	(0.114)	(0.138)
$\hat{\theta}$	–0.187	–0.091	–0.081	–0.081	–0.105	–0.092
	(0.024)	(0.08)	(0.113)	(0.087)	(0.086)	(0.088)
$\hat{\sigma}_\epsilon^2$	0.046	0.028	0.027	0.027	0.027	0.028
	(0.001)	(0.002)	(0.005)	(0.002)	(0.005)	(0.003)
$\hat{\phi}_{rw}$	0.0	1.0	0.992	0.992	1.0	1.0
	—	—	(0.158)	(0.009)	—	—
χ^2	793.32	1744.87	2184.65	696.96	492.25	636.89
(d.f.)	(431)	(430)	(429)	(430)	(376)	(430)

Notes: The estimated income process is: $y_{iht} = \alpha_i + \beta_i h + p_{iht} + \frac{1+\theta L}{1-\phi L}\epsilon_{iht} + u_{iht,me}$, where $p_{ih+1t+1} = \phi_{rw} p_{iht} + \xi_{ih+1t+1}$ and ϕ_{rw} denotes an autoregressive coefficient of a more persistent autoregressive process. Models are estimated by the equally weighted minimum distance method. Sample consists of 1,916 male household heads with at least 8 consecutive observations on labor income growth. Households from the Survey of Economic Opportunity (SEO) subsample are excluded. Standard errors in parentheses.

Table 8: The Variances of Permanent and Transitory Shocks by Year.

Year	Trans. shock	St. err.	Perm. shock	St. err.
1969	0.01989	0.00428	0.01274^a	—
1970	0.01461	0.00360	0.01274^a	0.00581
1971	0.02635	0.00535	0.01299	0.00348
1972	0.02337	0.00609	0.01336	0.00461
1973	0.02541	0.00536	0.02177	0.00540
1974	0.02373	0.00467	0.01961	0.00452
1975	0.03137	0.00549	0.01223	0.00425
1976	0.03649	0.00599	0.02345	0.00592
1977	0.02708	0.00466	0.01329	0.00466
1978	0.02843	0.00466	0.00647	0.00350
1979	0.02586	0.00502	0.01467	0.00476
1980	0.02350	0.00503	0.01672	0.00470
1981	0.01538	0.00545	0.02725	0.00655
1982	0.02959	0.00526	0.02854	0.00791
1983	0.02309	0.00430	0.02605	0.00463
1984	0.02623	0.00507	0.02821	0.00575
1985	0.03125	0.00516	0.02216	0.00465
1986	0.02418	0.00401	0.01732	0.00501
1987	0.02679	0.00453	0.02440	0.00487
1988	0.02768	0.00470	0.01464	0.00482
1989	0.02421	0.00489	0.02935	0.00547
1990	0.02413	0.00433	0.01979	0.00486
1991	0.02039	0.00407	0.01659	0.00432
1992	0.02066	0.00346	0.02261	0.00482
1993	0.04993	0.00662	0.00766	0.00499
1994	0.03701	0.00600	0.00499	0.00404
1995	0.03103	0.00508	0.00618	0.00468
1996	0.02568	0.00464	0.00859^b	0.00496
1997	0.03090	0.00699	0.00859^b	—

Notes: Estimates from the model in Table 7, column (5). ^{a},^{b} Variances of permanent shocks are restricted in estimation to be equal in those years.

Table 9: The Time Span of a Sample and Estimated Growth-Rate Heterogeneity. Simulated Data.

Time Span, T	(1) True: HIP $\hat{\sigma}^2_\beta$	(2) True: RIP with R.W. $\hat{\sigma}^2_\beta$
10	0.00036	0.00096
	(0.0001)	(0.00011)
15	0.00039	0.00074
	(0.0004)	(0.00006)
20	0.00038	0.00059
	(0.00002)	(0.00004)
25	0.00039	0.00050
	(0.00002)	(0.00003)
30	0.00039	0.00043
	(0.00002)	(0.00002)

Notes: In both columns, the estimated income process is: $y_{iht} = \alpha_i + \beta_i h + \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$. In column (1), the true process is the same as the estimated process. In simulations, the parameters are taken from column (1) of Table 7. In column (2), the true process is: $y_{iht} = \alpha_i + p_{iht} + \frac{1+\theta L}{1-\phi L}\epsilon_{iht}$, where $p_{iht} = p_{ih-1t-1} + \xi_{iht}$. In simulations, the parameters are taken from column (2) of Table 7. Models are estimated by the equally weighted minimum distance method. Standard errors in parentheses calculated as the standard deviations of the estimated growth-rate heterogeneity across 100 model simulations.

Appendix A: The Autocovariance Function and Estimation Details.

In this appendix, I briefly discuss the estimation method and present theoretical autocovariances and variances for the model (1)–(3) when the data used for estimation are in levels. For convenience, I reproduce the model equations here, assuming that the transitory component is an autoregressive process of order 1.

$$y_{iht} = \alpha_i + \beta_i h + p_{iht} + \tau_{iht} + u_{iht,me}$$
$$p_{iht} = p_{ih-1t-1} + \xi_{iht}$$
$$\tau_{iht} = (1-\phi L)^{-1}\epsilon_{iht},$$

where $(\alpha_i, \beta_i) \sim iid(0, \Omega)$, with $\Omega_{11} = \sigma^2_\alpha$, $\Omega_{12} = \Omega_{21} = \sigma_{\alpha\beta}$, $\Omega_{22} = \sigma^2_\beta$; $u_{iht,me} \sim iid(0, \sigma^2_{u,me})$; $\xi_{iht} \sim iid(0, \sigma^2_\xi)$; $\epsilon_{iht} \sim iid(0, \sigma^2_\epsilon)$. The moments used in matching estimations are:

$$var(y_{iht}) = \sigma^2_\alpha + \sigma^2_\beta h^2 + 2\sigma_{\alpha\beta} h + \sigma^2_{u,me} + var(\tau_{iht}) + var(p_{iht}), \qquad t = 1, \ldots, T,\ h = 1, \ldots, H$$
$$var(\tau_{i1t}) = \sigma^2_\epsilon \qquad var(p_{i1t}) = \sigma^2_\xi, \qquad t = 1, \ldots, T$$
$$var(\tau_{ih1}) = \sigma^2_\epsilon \sum_{j=0}^{h-1} \phi^{2j} \qquad var(p_{ih1}) = \sum_{j=0}^{h-1} \sigma^2_\xi, \qquad t = 1,\ h = 2, \ldots, H$$
$$var(\tau_{iht}) = \phi^2 var(\tau_{ih-1t-1}) + \sigma^2_\epsilon \qquad var(p_{iht}) = var(p_{ih-1t-1}) + \sigma^2_\xi, \qquad t = 2, \ldots, T,\ h = 2, \ldots, H$$
$$cov(y_{iht}, y_{ih+kt+k}) = \phi^k var(\tau_{iht}) + var(p_{iht}) +$$
$$+ \sigma^2_\alpha + \sigma_{\alpha\beta}(2h+k) + \sigma^2_\beta h(h+k),\ k = 1, \ldots, \min(H-h, T-t),\ h = 1, \ldots, H,\ t = 1, \ldots, T,$$

where H is the maximum labor market experience in the sample, and T is the time dimension of the sample.

I am assuming, in my Monte Carlo simulations and in the autocovariance function just presented, that $\tau_{i0t} = 0$ and $p_{i0t} = 0$, i.e., a head with no labor market experience entering the labor market at time $t+1$ is "endowed" with zero permanent and transitory components of earnings.

For idiosyncratic labor income growth, the above model is:

$$\Delta y_{it} = \beta_i + \xi_{it} + (1-\phi L)^{-1}\Delta\epsilon_{it} + \Delta u_{it,me}.$$

The autocovariance moments are shown in equations (10)–(12). If the transitory component is a moving average process of order 1, see the autocovariance function in the text in equations (5)–(8).

The empirical moments, taking into account that the data used in estimations are unbalanced, are calculated as:

$$vech\left(\sum_{i=1}^{N} \tilde{y}_i \tilde{y}_i'\right)/N_{tt'},$$

where $\tilde{y}_i = (y_{i(h)1}, y_{i(h+1)2}, \ldots, y_{i(h+T-1)T})$ if data are in levels; and $\tilde{y}_i = (\Delta y_{i2}, \Delta y_{i3}, \ldots, \Delta y_{iT})$ if data are in first differences; N is the total number of heads in the sample; $N_{tt'}$ is a vector with the row dimension $\frac{T(T+1)}{2}$; N_{11} is the number of heads contributing towards estimation of the variance in period 1 ($t=1$, $t'=1$); N_{12}—the number of heads contributing towards estimation of the first-order autocovariance between periods 1 and 2 ($t=1$, $t'=2$), etc. Note that if the head's income is missing, say, in period 1, this head's contributions towards the variance at time 1 and all the sample autocovariances involving this period are zero. The vector of data moments used in estimation is $m^d = vech\left(\sum_{i=1}^{N} \tilde{y}_i \tilde{y}_i'\right)/N_{tt'}$; the row dimension of m^d is $\frac{T(T+1)}{2}$. The model parameters, Θ, are recovered by minimizing a squared distance function $\left[m(\Theta) - m^d\right]' I \left[m(\Theta) - m^d\right]$, where I is an identity matrix with the row dimension $\frac{T(T+1)}{2}$.

Standard errors of the parameters are calculated as the square roots of the diagonal of $(G'_\Theta G_\Theta)^{-1} G'_\Theta V G_\Theta (G'_\Theta G_\Theta)^{-1'}$, where $G_\Theta = \frac{\partial}{\partial\Theta}\left[m(\hat{\Theta}) - m^d\right]$, a vector with the row dimension $\frac{T(T+1)}{2}$, and the column dimension equal to the row dimension of the vector of estimated parameters; V is equal to $\sum_{i=1}^{N}\left(m_i - m^d\right)\left(m_i - m^d\right)'/N_V$, where $m_i = vech(\tilde{y}_i \tilde{y}_i')$, and the kl-th element of N_V is calculated as $N_V^{kl} = N_{tt'}^k N_{tt'}^l$, where $N_{tt'}^k$ is the k-th element of $N_{tt'}$.

Appendix B: Additional Results.

TABLE B-1: ESTIMATES OF INCOME PROCESSES. PSID DATA. SAMPLE SPLIT BY EDUCATION.

	High school grad. or less		Some college or more	
	(1)	(2)	(3)	(4)
$\hat{\sigma}^2_\beta$	0.0003	0.00	0.0004	0.00
	(0.00006)	(0.00008)	(0.00007)	(0.0001)
$\hat{\sigma}^2_\xi$	0.00	0.012	0.00	0.02
	—	(0.002)	—	(0.003)
$\hat{\phi}$	0.588	0.335	0.848	0.385
	(0.050)	(0.141)	(0.029)	(0.209)
$\hat{\theta}$	–0.165	–0.073	–0.179	–0.084
	(0.041)	(0.099)	(0.025)	(0.143)
$\hat{\sigma}^2_\epsilon$	0.048	0.035	0.044	0.020
	(0.002)	(0.003)	(0.002)	(0.003)

Notes: The estimated income process is: $y_{iht} = \alpha_i + \beta_i h + p_{iht} + \frac{1+\theta L}{1-\phi L}\epsilon_{iht} + u_{iht,me}$, where $p_{ih+1t+1} = p_{iht} + \xi_{ih+1t+1}$. Models are estimated by the equally weighted minimum distance method. In columns (1) and (2) sample consists of 1,011 male household heads with at least 8 consecutive observations on labor income growth whose education levels do not exceed high school. In columns (3) and (4) sample consists of 905 male household heads with at least 8 consecutive observations on labor income growth who finished some college or graduated from college. Households from the Survey of Economic Opportunity (SEO) subsample are excluded. Standard errors in parentheses.