Learning Your Earning: Are Labor Income Shocks Really Very Persistent?*

Fatih Guvenen†

June 27, 2004

Preliminary and Incomplete. Comments Welcome.

Abstract

In this paper we examine the risk situation facing individuals in the labor market. The current consensus in the literature is that the labor income process has a large random walk component. We argue two points. First, the estimates of persistence from income data appear to be upward biased due to the omission of heterogeneity in income profiles across the population that would be implied, for example, by a human capital model with heterogeneity.in ability When we allow for differences in profiles, the estimated persistence falls from 0.99 to about 0.8. Moreover, the main evidence against profile heterogeneity in the existing literature—that the autocorrelations of income changes are small and negative—is also replicated by the profile heterogeneity model we estimate, casting doubt on the previous interpretation of this evidence. Second, we embed this process into a life-cycle model to examine how it alters individuals' consumption-savings decision. We assume that—as seems plausible—individuals do not know their profile exactly at the beginning of life, but learn in a Bayesian way with successive income observations. We find that learning is slow, and thus the uncertainty about income profiles affects consumption decision throughout the life-cycle. Consistent with empirical evidence, the model generates: (i) substantial rise in consumption inequality over the life-cycle (Deaton and Paxson 1994); (ii) a non-concave age-inequality profile of consumption (in contrast to a model with very persistent shocks); (iii) consumption profiles for college graduates that are steeper than those for high school graduates as documented by Carroll and Summers (1991). Overall this evidence indicates that income shocks may be significantly less persistent than what is currently assumed.

JEL classification: D31, D83, D91

*I would like to thank Michael Baker, Mark Bils, Jeremy Greenwood, Larry Katz, Per Krusell, Lee Ohanian, Kjetil Storesletten, Gianluca Violante, Neng Wang, seminar participants at Rochester, UCLA, USC and Western Ontario for useful comments; Dirk Krueger for providing the CEX consumption data; Damba Lkhagvasuren for excellent research assistance; and Marianna Kudlyak for expert editorial assistance. This paper originally started as a joint project with Shamim Mondal, and special thanks goes to him. The usual disclaimer applies.

†Contact address: 216 Harkness Hall, University of Rochester, Rochester, NY, 14627. E-mail: guvn@mail.rochester.edu; http://www.econ.rochester.edu/guvenen/guvenen.htm

1 Introduction

The goal of this paper is to analyze the risk situation facing individuals in the labor market, and, in particular, to question the current conventional wisdom that labor income movements are dominated by nearly-permanent idiosyncratic shocks. Understanding the nature of idiosyncratic shocks is crucial because the properties of these shocks—and their persistence in particular—have profound effects on individuals' consumption-savings decision, which lies at the heart of a range of economic problems.[1] It is fair to say that the effectiveness of self-insurance, and hence the quantitative importance of market incompleteness hinges to a large extent on the persistence of labor income shocks (Deaton (1991), Aiyagari (1994), Levine and Zame (2002)).

In response to this central role played by income shocks, a large and growing literature has emerged investigating the stochastic process for labor income (and wages) using ever more sophisticated econometric techniques (among others, MaCurdy (1982), Abowd and Card (1989); Moffitt and Gottschalk (1995), Carroll and Samwick (1997), Meghir and Pistaferri (2003), Storesletten et al. (2004)). The current consensus among these studies is that the income process contains a large random walk component. This conclusion has been further bolstered by the—more indirect—evidence that within-cohort consumption inequality increases dramatically over time, which would be implied by the permanent income model, again, only when income shocks are very persistent. Summarizing the existing evidence, Lucas (2003) states:

> The fanning out over time of the earnings and consumption distributions within a cohort that Angus Deaton and [Christina] Paxson (1994) document is striking evidence of a sizeable, uninsurable random walk component in earnings.

In this paper we consider an income process that relaxes a key (restrictive) assumption made in this literature. Although the estimated process reveals only modestly persistent shocks, the implied consumption behavior in a life-cycle framework is consistent with a number of important aspects of consumption data, including the significant rise in consumption inequality over time.

To introduce these ideas, let us first elaborate on the econometric evidence on labor income shocks mentioned above. It is clear that a discussion of "unanticipated shocks" requires the econometrician to take a stand on what is "anticipated" or predictable by individuals. In order to capture this latter (life-cycle) component, the standard approach is to posit a simple earnings function, which typically includes a polynomial in experience, an education dummy, and occasionally a few additional variables. The key assumption is that the coefficients in

[1] For example, the welfare analyses of social insurance policies depend on the amount and nature of risk that needs to be insured. Moreover, to the extent that the size of idiosyncratic risk is correlated with aggregate economic conditions—which seems to be the case, see Storesletten, Telmer and Yaron (2004)—the welfare costs of business cycle fluctuations will also depend on individual level risk (see Lucas 2003 for an extensive discussion). In a different context, Constantinides and Duffie (1996) argued that a model with idiosyncratic income shocks can succesfully explain many features of asset prices if these shocks are (nearly) permanent (see also Storesletten et al. 2002).

this earnings function are restricted to be the same across the population, so all individuals expect the same life-cycle income profile (conditional on education). Furthermore, this earnings function predicts only a small fraction of actual earnings variation—typically about 5 to 15 percent—suggesting that individuals have little idea about where they will land within the lifetime income distribution. Clearly, this simple earnings function rules out potential variation in income growth rates arising from unobserved heterogeneity in the population. But there are reasons to suspect that such heterogeneity exists and may in fact be quite important. For example, incorporating heterogeneity in learning ability into plausible versions of the human capital model will imply differences in income profiles across individuals (Becker (1965), Ben-Porath (1967)).[2] Similarly, heterogeneity in profiles may arise from variations in returns-to-experience across occupations and professions (c.f., Carroll and Summers (1991)).

This paper is not the first one to recognize that income profiles may be different for different individuals. In fact, early papers in this area, perhaps influenced by human capital theory, studied an econometric model with profile heterogeneity (henceforth, PH) as a natural starting point, and found evidence of both statistically and quantitatively significant heterogeneity in profiles (Hause (1977, 1980); Lillard and Weiss (1979)). In an influential paper, MaCurdy (1982) cast doubt on these earlier findings. He tested the simple proposition that if individuals differed systematically in their income growth rates, then the autocovariances of income changes should be positive. Instead, he found them to be close to zero, and, in fact, slightly negative (after the first lag). Subsequent work by Abowd and Card (1989), Topel (1990) and Topel and Ward (1992) tested this implication using various longer panel data sets only to confirm MaCurdy's conclusion. This body of work constitutes the main evidence against profile heterogeneity. It is important to note that none of these later papers estimated an econometric specification nesting profile heterogeneity and found it to be insignificant.

Although the described test is intuitive, and valid in principle, it has low power against the alternative of PH and is thus not well-suited to provide a verdict on this case. To see this, suppose that log income of individual i who is h years old is given by $y_h^i = \beta^i h + \varepsilon_h^i$, where β^i is the individual-specific income growth rate, and ε_h^i is a purely transitory shock. It is easy to see that $cov(\Delta y_h^i, \Delta y_{h+k}^i) = \sigma_\beta^2$, for $k \geq 2$. The key point to note is that a dispersion in β^i that is substantial—in terms of its implications for the income process—still corresponds to a value for σ_β^2 that appears minuscule. For example, with a value of σ_β^2 as small as 0.0004, heterogeneity in β^i alone is sufficient to generate the entire rise in income inequality over the life-cycle observed in the U.S. data. Thus, one should expect to see very

[2] The underlying sources of variation—such as those in communication, social, and organizational skills, among others—may not be observable to the econometrician and may be hard to proxy with variables commonly found in panel data sets, but are likely to be quite well anticipated, or learned over time by individuals themselves. Even using very noisy measures of labor market skills—such as the scores individuals get on the AFQT, which is a general aptitude test, or the income of a sibling, to proxy for genetic component of certain abilities—there is evidence of variation in the slope of earnings profiles across individuals, though the explained variation is not very large (see Heckman, Lochner and Taber 1998; Altonji and Pierret 2001)

small autocovariances even in the presence of significant heterogeneity in profiles. In addition, if the income process also contains an AR(1) component, $cov(\Delta y_h^i, \Delta y_{h+k}^i)$ will also contain a negative term, capturing mean reversion (equation 4). Thus, covariances can easily be negative as found by MaCurdy and others. In Section 3 we show that this is indeed the case. Using simulated data from an estimated income process with PH, we show that the autocovariances closely match those calculated from actual income data reported in Topel (1990) among others, and are not statistically significant (see Table 2). This finding in our view casts serious doubt on the evidence against PH. Using a similar income process featuring PH, Baker (1997) conducts a careful Monte Carlo analysis and also raises important issues about the reliability of the evidence against PH.

It is easy to see that ignoring PH (when in fact it is present) will bias the estimated persistence parameter upward, because the fanning out of the income distribution over time due to systematic differences among individuals (i.e., dispersion in β^i) will be incorrectly attributed to persistent shocks. This bias can be substantial: assuming the same process for y_h^i given above with purely transitory shocks, and $\sigma_\beta^2 = 0.0004$, the persistence will be estimated to be about 0.90 instead of the true value of zero. This result cautions against estimating a restricted econometric specification, especially given the lack of evidence against PH.

In Section 3, we estimate the parameters of an income process incorporating PH that is similar to the one used by Lillard and Weiss. Using data on labor earnings from the Panel Study of Income Dynamics (PSID) covering the period 1967 to 1993, we find statistically and quantitatively significant heterogeneity in income profiles. Furthermore, the estimated persistence falls from 0.99 down to about 0.80, a difference that is substantial for most practical purposes.

We next embed this income process into a life-cycle model to examine how the existence of PH affects individuals' consumption-savings decision. A natural question that arises in this context is how much an individual knows about his own profile. Given the complexity of factors that may give rise to this heterogeneity it seems plausible to assume that, when he enters the labor market, an individual has less than perfect information about the parameters of his profile. We capture this initial uncertainty with a prior belief over α^i and β^i, and assume that the individual updates his beliefs in a Bayesian fashion with subsequent income realizations, resulting in the gradual resolution of profile uncertainty over time. We cast the optimal learning process as a Kalman filtering problem, which allows us to conveniently obtain recursive updating formulas in the presence of AR(1) shocks to income. In a closely related model, Wang (2004) obtains closed-form solutions for optimal consumption choice when the decision-maker cannot distinguish between two separate persistent shock processes, but without learning about individual profiles.

It is often the case with Bayesian learning that most of the uncertainty is resolved quickly, with only a handful of observations. Instead, in our framework, learning is gradual and its effects on consumption choice are prolonged—extending throughout the life-cycle—for three main reasons. First, although learning reduces the posterior variance of β^i over time, the effect of β^i on income is amplified by its interaction with age in the earnings function, thus slowing down the learning process. Second, as noted above, our estimates indicate that income shocks are far from being transitory. These persistent deviations make it harder for individuals to distinguish the trend component. The combination of these two factors (together with a third discussed in Section 4) results in the slow resolution of income uncertainty. For plausible parameter values, the amount of uncertainty introduced by PH is far more important than the part arising from AR(1) shocks.

We next compare our model to the U.S. consumption data. First, in our baseline model the cross-sectional variance of log income increases by about 0.3 over the life-cycle, somewhat exceeding the rise in the U.S. data. Hence, this income process *is* consistent with substantial fanning out of the consumption distribution. Second, the empirical age-inequality profile has a *non*-concave shape. This fact has been emphasized by Deaton and Paxson (and later by Storesletten et al. (2003)) because a life-cycle model with persistent shocks implies a *concave* shape. Our baseline model instead generates a non-concave profile which also seems to match its empirical counterpart quite well. Third, a number of authors have shown that consumption tracks income over the life-cycle. For example, college graduates not only have steeper income profiles than high-school graduates but also have steeper consumption profiles (Carroll and Summer (1991)). When PH is ignored, the estimated innovation variance and persistence for each group are close to one another, resulting in similar consumption profiles for both groups. On the other hand, when PH is introduced, we find that the estimated dispersion of β^i among college graduates is more than twice that among high-school graduates. To the extent that this larger dispersion translates into more initial uncertainty about the former group's income growth rates, this will generate more precautionary savings and a steeper consumption profile for the former group. These three examples underscore the difference between the nature of labor income risk implied by permanent shocks, and that resulting from uncertainty (and learning) about income profiles.

Our results support the conclusion of Huggett, Ventura and Yaron (2003) who study a human capital model and find differences in the slope of income profiles to be critical for matching the evolution of the first three moments of the earnings distribution over the life-cycle. Similarly, Heckman, Lochner and Taber (1998) estimate a human capital model with a rich set of realistic features and reach a similar conclusion.

The rest of the paper is organized as follows. The next section describes the data and the estimation method. Section 3 presents the empirical results and relates them to the existing literature. Sections 4 and 5 study optimal learning and embeds it into a life-cycle consumption-savings model. Section 6 presents the model results and Section 7 concludes.

2 Empirical Analysis

2.1 The PSID Data

The data are drawn from the public release files of PSID covering the period from 1968 to 1997.[3] Our main sample consists of male head of households between the ages of 22 and 62.[4] We include an individual into the sample if he satisfies the following conditions for twenty (not necessarily consecutive) years: the individual has (1) reported positive labor earnings and hours; (2) worked between 520 and 5110 hours in a given year; (3) had an average hourly earnings between a preset minimum and a maximum wage rate (to filter out extreme observations). These criteria are similar to the ones used in previous studies (Abowd and Card (1989), Baker (1997), and Heathcote et al. (2004) among others). Further details of the selection criteria and summary statistics for the primary sample are contained in Appendix A.

These criteria leave us with our *main sample of 1270 individuals* with at least twenty years of data on each, yielding a total of 30,945 income observations. To study the labor income processes of different education groups separately, we further draw two subsamples: the first contains 335 individuals with at least a four-year college degree (sixteen years of education or more), and the second contains 882 individuals with at most a high school degree (fifteen years of education or less).[5] To make the text more readable, we will refer to the former group as "college-educated" and the latter as "high school educated," at the expense of a slight abuse of language.

The traditional approach to panel construction used by most previous studies (Lillard and Weiss (1978); MaCurdy (1982); Abowd and Card (1989), and Baker (1997) among others) requires an individual to satisfy the selection criteria for every year of the sample period to be included in the panel. Although this condition has the advantage of creating a balanced panel, it also has the drawback of reducing the sample size significantly as the time horizon expands This would pose a significant problem in our case since our panel extends over thirty years. Hence, our requirement for each individual to satisfy our selection criteria for twenty (out of thirty) years is intended to make our panel construction comparable to these studies, while at the same time keeping a sufficient number of observation. [6]

[3] Certain variables (such as head's annual labor earnings) refer to the previous year. So, for example, the 1968 survey contains data on earnings in 1967.

[4] Following much of the literature we exclude individuals in the Survey of Economic Opportunities subsample which oversamples low-income and non-white individuals.

[5] The remaining 53 individuals from the primary sample report a change in their education status during the sample period. We do not include them in either sample.

[6] An alternative approach pursued by some recent studies is to include an individual into the panel if certain criteria are satisfied for a few—usually two or three—years. (Heathcote et al. (2004), and Storesletten et al. (2004)). As we discuss in Section 3, our results do not seem to be sensitive to this choice (see also Haider 2001).

2.2 A Statistical Model

The process for *log earnings*, $\widetilde{y}_{it}^{h}$, of individual i with h years of labor market experience in year t is given by

$$\widetilde{y}_{it}^{h} = g\left(\boldsymbol{\theta}_t^0, \mathbf{X}_{it}^h\right) + f\left(\boldsymbol{\theta}^i, \mathbf{X}_{it}^h\right) + z_{it}^h + \phi_t \varepsilon_{it}^h \tag{1}$$

where the functions g and f denote the "life-cycle" components of earnings. The first function captures the part of variation that is common to all individuals (hence the coefficient vector $\boldsymbol{\theta}_t^0$ is not individual-specific) and is assumed to be a higher order polynomial in experience, h. Notice that the coefficients of this polynomial are allowed to be time-varying. In addition to the standard time effects (aggregate shocks) in labor income captured by year-to-year variations in the intercept of g, this flexible specification also allows us to model a number of important changes that took place in the labor market during our sample period. For example, changes in the return to experience that took place during this period (Katz and Autor (1999)) can be accounted for by the (time-varying) higher order terms in experience. Similarly, the rise in the skill premium, which is another important trend documented in the literature (Katz and Murphy (1992)) can be captured by adding an education dummy with a time varying coefficient into g. In the baseline specification though we do not pursue this strategy. Later in the paper, we analyze the labor income processes of different education groups separately to address such issues.

The second function, f, is the centerpiece of our analysis, and captures the component of life-cycle earnings that is idiosyncratic to each individual or possibly to a sub-group of individuals in the population. For example, if the growth rate of earnings varies with the ability of a worker, or is different across occupations, this variation will be reflected in an individual- or occupation-specific slope coefficient in f. In the baseline case, we assume this function to be linear in experience: $f\left(\boldsymbol{\theta}^i, \mathbf{X}_{it}^h\right) = \alpha^i + \beta^i h$, where the random vector $\boldsymbol{\theta}^i \equiv \left(\alpha^i, \beta^i\right)$ is distributed across individuals with zero mean[7] and covariance matrix

$$\mathbf{V}(\theta) = \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{bmatrix}.$$

Although it is straightforward to generalize f to allow for heterogeneity in higher order terms, Baker (1997, p. 373) finds that this extension does not noticeably affect parameter estimates or improve the fit of the model. Moreover, each term introduced into this component will appear as an additional state variable in the dynamic programming problem we solve in Section 5. In the baseline case, that problem already has four continuous state variables (and certain non-standard features described in the computational appendix), so we prefer to avoid any further complexity. We do however experiment with other extensions that do not add state variables below, and leave this extension for future research.

[7] This assumption is merely a normalization since g already includes an intercept and a linear term. Thus, in any given year, the population averages of the intercept and slope are given by the first two coefficients of g.

In specifying the stochastic component of income, we have several considerations in mind. First, the main goal of this study is to shed light on the persistence of income shocks, so it seems natural to include an autoregressive component, which does not restrict the persistence parameter (as opposed to a random walk process). This specification can capture mean-reverting shocks, such as human capital innovations that depreciate over time, or a long-term nominal wage contract whose value decreases over time in real terms, as well as fully permanent shocks as a special case. Second, recent empirical studies have documented dramatic changes in the sizes of both persistent and transitory shocks to labor income over the sample period under study (c.f. Moffitt and Gottschalk (1994), Haider (2001), Meghir and Pistaferri (2004)). To capture this non-stationarity, we write z_{it}^h as an AR(1) process with heteroskedastic shocks:

$$z_{it}^h = \rho z_{it-1}^{h-1} + \pi_t \eta_{it}^h, \qquad z_{it}^0 = 0,$$

where π_t captures potential time-variation in the innovation variance. Similarly, the transitory shock in equation (1), ε_{it}^h, is scaled by ϕ_t to account for possible non-stationarity in that component. The innovations η_{it}^h and ε_{it}^h are assumed to be independent of each other and over time (and independent of α^i and β^i), with zero mean, and variances of σ_η^2 and σ_ε^2 respectively. Furthermore, measurement error is a pervasive problem in micro datasets and income data is PSID is no exception. This measurement error will be captured in the transitory component if it is serially independent, or will be included in z_{it}^h if it has an autoregressive component (Bound and Krueger (1991)). It is useful to keep this point in mind when interpreting the empirical findings in the next section.

Our estimation strategy is based on minimizing the distance between the elements of the $(T \times T)$ empirical covariance matrix of income residuals (denote it by $\mathbf{C}$) and its counterpart implied by the statistical model described above (Chamberlain (1984)). This approach has been used extensively in this literature (including most of the studies referenced in this paper), so it is familiar enough that we relegate its details to Appendix B. In what follows we provide a brief description.

The income residual, y_{it}^h, is obtained by regressing $\widetilde{y}_{it}^h$ on the polynomial g. Since the individual-specific parameters, α^i and β^i, are not observable, f is treated as part of the random component of the income process and is included in the residual. For a given year, the cross-sectional moments of this residual for a cohort of a given age are:

$$\begin{aligned} var\left(y_{it}^h\right) &= \left[\sigma_\alpha^2 + 2\sigma_{\alpha\beta} h + \sigma_\beta^2 h^2\right] + var\left(z_{it}^h\right) + \phi_t^2 \sigma_\varepsilon^2 && (2) \\ cov\left(y_{it}^h, y_{it-n}^{h-n}\right) &= \left[\sigma_\alpha^2 + \sigma_{\alpha\beta}\left(2h-n\right) + \sigma_\beta^2 h \left(h-n\right)\right] + \rho^n var\left(z_{it-n}^{h-n}\right), && h,t > n > 0 \end{aligned}$$

where the variance of the AR(1) component is obtained recursively:

$$
\begin{aligned}
var\left(z_{it}^{1}\right) &= \pi_t^2 \sigma_\eta^2, \\
var\left(z_{i1}^{h}\right) &= \pi_1^2 \sigma_\eta^2 \sum_{j=0}^{h-1} \rho^{2j}, && t = 1, h > 1 \\
var\left(z_{it}^{h}\right) &= \rho^2 var\left(z_{it-1}^{h-1}\right) + \pi_t^2 \sigma_\eta^2, && t > 1, h > 1.
\end{aligned}
$$

Note that these equations are based on some implicit assumptions also common in the literature. In the first line, we assume that the initial value of the persistent shock is zero for all individuals. In the second line we assume that the innovation variance was constant over time before the sample started in 1968, so that the cross-sectional variance for a cohort aged h in the first year of the sample can be determined by the accumulated effect over the last h years.

Given these formulas for each (h, t) cell, one can aggregate across h to obtain the matrix **C**. For example, to obtain the element at location (2,5) in **C**—corresponding to the covariance between income residuals in 1968 and 1971—one needs to calculate the average of $\left(y_{i,71}^{h} y_{i,68}^{h-3}\right)$ across individuals of all ages who were present in these two years. Then the corresponding covariance implied by the statistical model is obtained by appropriately aggregating the counterpart in equation (2) over h. The exact formulas used as well as a discussion of the choice of weighting matrix, and related issues are contained in Appendix B.

3 Empirical Findings

To provide a benchmark, we first estimate the parameters of equation (1) by ignoring individual-specific variation in income growth rates ($\beta^i \equiv 0$) but allowing for an individual fixed-effect, α^i. We call this statistical model the "homogenous profile model" to distinguish it from the more general "heterogenous profile model," which does not impose $\beta^i \equiv 0$.

3.1 The persistence of labor income shocks

The first row in Table 1 displays the parameter estimates from the homogenous profile model.[8] The first finding is that the estimate of ρ is 0.988, which implies that income shocks follow a near-random walk process, confirming the results of previous studies that used similar parametric models. In this case, one cannot statistically reject that income shocks are permanent at

[8] The estimates reported in Table 1 are obtained using the first 26 years (1967-92) of our sample. Initially we were reluctant to use the last four years of the sample due to some concern that the data quality seems to have gone down in this later period because of reductions in the editing budgets that began in 1993 (see Haider 2001 for a discussion). PSID now encourages the use of data from this later period. We recently repeated our estimation using the full sample and obtained very similar results (reported in the next footnote). These estimates will be incorporated into the next revision.

Table 1: Estimating the parameters of the labor income process

		heter?	ρ	σ^2_α	σ^2_β	$corr_{\alpha\beta}$	σ^2_η	σ^2_ε	(ϕ, π)
(1)	A	no	.988	.058	—	—	.015	.061	yes
			(.024)	(.011)			(.007)	(.010)	
(2)	A	yes	.821	.022	.00038	$-.23$	.029	.047	yes
			(.030)	(.074)	(.00008)	(.43)	(.008)	(.007)	
(3)	C	no	.979	.031	—	—	.0099	.047	yes
			(.055)	(.021)			(.013)	(.020)	
(4)	C	yes	.805	.023	.00049	$-.70$	.025	.032	yes
			(.061)	(.112)	(.00014)	(1.22)	(.015)	(.017)	
(5)	H	no	.972	.053	—	—	.011	.052	yes
			(.023)	(.015)			(.007)	(.008)	
(6)	H	yes	.829	.038	.00020	$-.25$	.022	.034	yes
			(.029)	(.081)	(.00009)	(.59)	(.008)	(.007)	

Notes: Standard errors are in parentheses. In the second column, A = all individuals, C = college-educated group, and H = high school educated group. Time effects in the variances of persistent and transitory shocks are not reported to save space, and the reported variances are averages over the sample period.

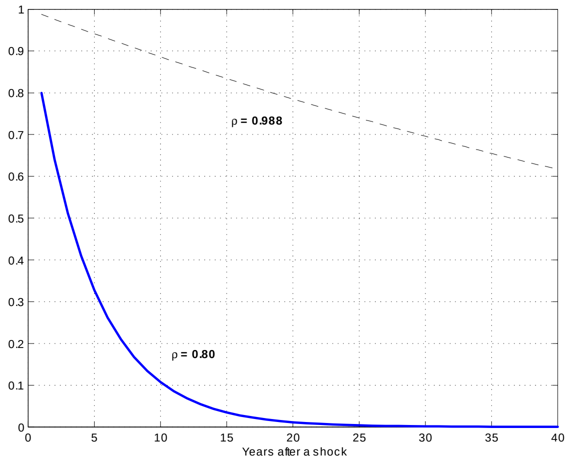
conventional significance levels. The innovation variance of z is also large so that in the long-run the persistent component dominates the cross-sectional distribution of income.

Starting in the second row, we allow for individual-specific differences in income growth rates. The first main finding is that the estimated persistence falls from 0.988 to 0.82.[9] This is a substantial difference for all practical purposes, and although this is a well-known point, it is important enough to warrant a few words. Figure 1 plots the impulse response of an AR(1) process with different persistence levels. When $\rho = 0.8$, the effect of an income shock is reduced to ten percent of its initial value in ten years, whereas for $\rho = 0.988$, it retains almost ninety percent of its effect at the same horizon. After twenty years, the effect of the former shock almost vanishes whereas the latter shock still keeps close to eighty percent of its initial impact. As can be anticipated from this comparison, optimal consumption and savings decisions are radically different for individuals facing each of these shocks, as we illustrate in Section 6.

Biased estimates of persistence: An Example. It can be readily seen why ignoring heterogeneity in income growth rates would lead to an upward bias in ρ. With heterogeneity in β^i, the income of an individual who has above (or below) average growth rate will deviate from the median in a systematic way over time. Ignoring this fact will then lead one to interpret this predictable fanning out as the result of a sequence of persistent positive (or negative) shocks to this individual's income.

[9] When we use the full sample up to 1996, the parameter estimates remain very similar to those reported: $\rho = 0.84$, $\sigma^2_\alpha = .055$, $\sigma^2_\beta = .00039$, $corr_{\alpha\beta} = -.35$, $\sigma^2_\eta = .022$, $\sigma^2_\varepsilon = .051$.

Figure 1: Impulse response functions for AR(1) income shocks for different persistence levels



The resulting bias can be substantial as can be seen in the following example. Suppose that the true persistence of income shocks in the heterogenous profile model is zero ($\rho \equiv 0$, so we only consider ε_{it}^{h}). Now suppose that, as is common, the econometrician allows for a fixed effect in the intercept, but not in the growth rates, assuming that β^i is equal to $\overline{\beta}$ for all i. In this case, the residuals are:

$$\upsilon_{it}^{h} \equiv y_{it}^{h} - \overline{y}_{it}^{h} = \left(\alpha^i + \beta^i h + \varepsilon_{it}^{h}\right) - \left(\alpha^i + \overline{\beta} h\right) = \left(\beta^i - \overline{\beta}\right) h + \varepsilon_{it}^{h}$$

Notice that, in contrast to the underlying shock, υ_{it}^{h} does not have zero mean for a given individual over time; instead it will either trend up or down. To simplify things further, assume that all individuals are observed when they are τ and $\tau+1$ years old (we relax this assumption below). Then, under the (incorrect) assumption of homogenous profiles, a consistent estimator of the persistence parameter is the minimizer of $(1/N)\sum_{i=1}^{N}\left(\upsilon_{i\tau+1}^{h+1} - \widehat{\rho}\upsilon_{i\tau}^{h}\right)^2$, which is given by

$$\widehat{\rho} = \frac{\tau(\tau-1)\sigma_{\beta}^{2}}{(\tau-1)^2\sigma_{\beta}^{2} + \sigma_{\varepsilon}^{2}}$$

Notice that $\widehat{\rho}$ is increasing in τ, and approaches unity in the limit. Substituting some plausible parameter values ($\sigma_{\beta}^{2} = 0.0004$, and $\sigma_{\varepsilon}^{2} = 0.03$) and assuming $\tau = 20$, yields $\widehat{\rho} = 0.87$. Similarly, for $\tau = 30$, one obtains $\widehat{\rho} = 0.95$, when in fact the true persistence is, again, zero. One can easily extend this calculation to show that if there is a population of individuals uniformly distributed from 25 to 65 years of age ($h = 1$ to 40), $\widehat{\rho} = 0.91$. These results suggest that it is prudent to allow for profile heterogeneity in modeling the earnings process.[10]

[10] One possible preference for considering permanent shocks to income might be the various theories of wage

3.2 Quantifying the heterogeneity in income profiles

The second key finding (in row 2 of Table 1) is that the heterogeneity in income growth rates measured by σ_β^2 is not only statistically but also quantitatively significant. For example, by age 49, the income of an individual whose growth rate is one standard deviation (≈ 0.02) above the mean will have doubled from its initial value, whereas the median income will have risen by a meager 24 percent over the same period.

As an alternative measure, we use the following equation to calculate the fraction of within-cohort income inequality that results from systematic differences in income growth rates:[11]

$$\begin{aligned} var_i(y_{it}^h) &= \left(\sigma_\alpha^2 + 2\sigma_{\alpha\beta}h + \sigma_\beta^2 h^2\right) + \left(\frac{1-\rho^{2h+1}}{1-\rho^2}\sigma_\eta^2 + \sigma_\varepsilon^2\right) \\ &= \left(\sigma_\alpha^2 + \sigma_\varepsilon^2\right) + \left(\frac{1-\rho^{2h+1}}{1-\rho^2}\sigma_\eta^2\right) + \left(2\sigma_{\alpha\beta}h + \sigma_\beta^2 h^2\right) \end{aligned}$$

In the second line, the first parenthesis contains terms that do not depend on age (i.e., the intercept of the age-inequality profile). The second parenthesis captures the rise in inequality due to the autoregressive shock. For the estimated value of $\widehat{\rho} = 0.82$, this component increases slightly in the first six to seven years and then remains roughly constant. Finally, the last parenthesis contains a decreasing linear term (since $\sigma_{\alpha\beta} < 0$) and a positive quadratic term in h. Putting these pieces together, it becomes clear that the dynamic component of income (second and third terms) mainly determines the *level* of the age-inequality profile but has little affect on the *rise* of inequality over the life-cycle, which is largely due to profile heterogeneity (last parenthesis). To see this, note for example that at age 54 ($h = 30$) heterogeneity in β contributes to inequality by $\sigma_\beta^2 h^2 = 0.00038 * (30^2) = 0.34$. This effect is mildly dampened by the negative covariance term, $2h\sigma_{\alpha\beta} = -0.04$, so the net effect is 0.30 which is 2/3 of the total at that age. Similar calculations show that heterogeneity only explains 15 percent of total dispersion at age 35, but that this fraction rises to 81 percent at retirement.

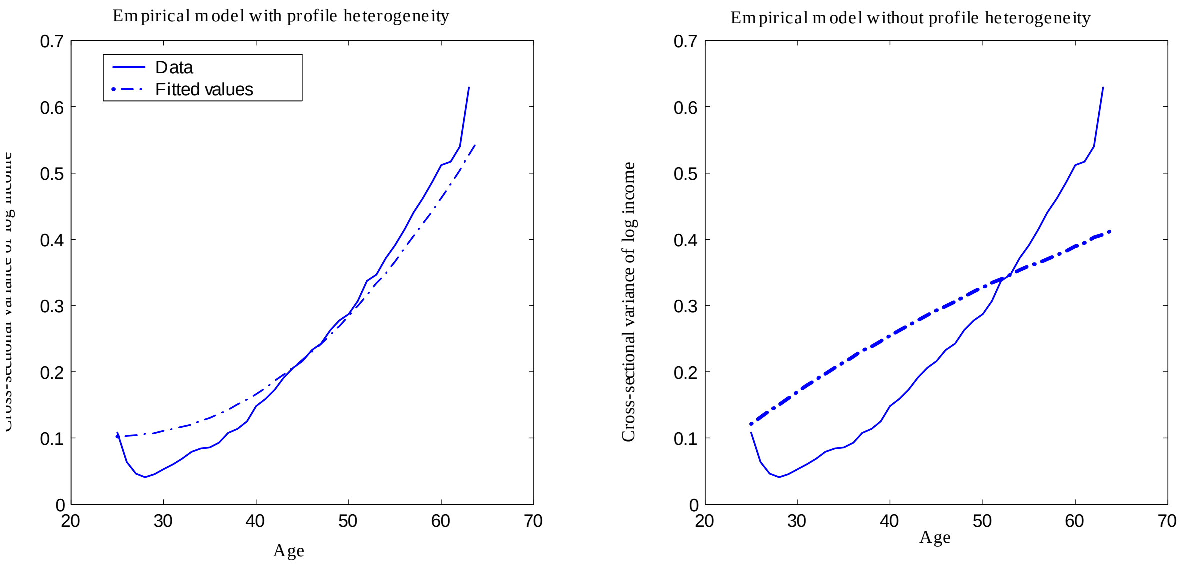
Before concluding this subsection, figure 2 displays the fit of each estimated model to the age-inequality profile of income, which is central to our analysis of consumption behavior in Section 5.[12] It is clear that allowing for heterogeneity in profiles (left panel) helps the model better account for the slightly convex rise in dispersion over the life-cycle.

determination (c.f., Jovanovic 1979). Baker and Solon (2003) allow for both profile heterogeneity *and* a random walk component in their econometric specification and still estimate statistically and quantitatively significant heterogeneity in slopes.

[11] In the rest of the paper, we set the time effects in variances π_t and ϕ_t equal to 1.

[12] The age-inequality profile is obtained by regressing the raw variances of each age-year cell, on age and cohort dummies following Deaton and Paxson (1994). The graph plots the coefficients on age dummies scaled so that the variance in the first year matches that in the respective parametric model.

Figure 2: THE FIT OF ESTIMATED ECONOMETRIC MODELS TO THE EMPIRICAL AGE-INEQUALITY PROFILE OF INCOME

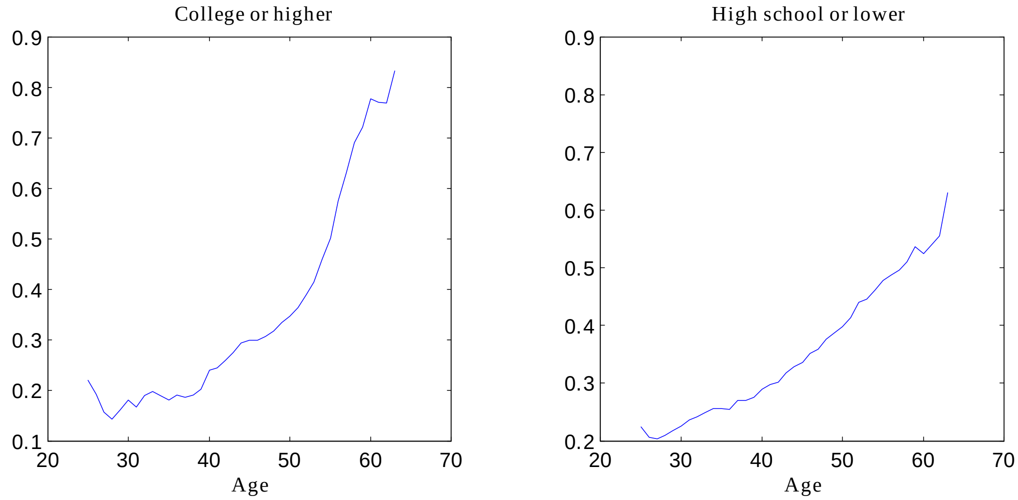


3.3 The labor income process of each education group

We next examine the labor income process of each education group separately. To our knowledge, differences between the income processes of these groups have so far only been investigated in the context of homogenous profile models (Hubbard et. al (1994); Carroll and Samwick (1997)). By and large, these earlier studies have not found significant differences across education groups in the estimated persistence parameter, but found some evidence that the innovation variance goes down with the level of education. Our first goal is then to investigate if these conclusions are robust to the introduction of profile heterogeneity. Second, by using these estimated processes as inputs into a life-cycle model, we study if the implied consumption and savings behaviors of different education groups are consistent with salient features of the U.S. data.

We first report the parameter estimates from the homogeneous profile model for the high- and low-education groups (rows three and five respectively), which broadly confirm the findings of the existing literature mentioned above. Next, we allow for profile heterogeneity. Now, the estimated persistence is significantly lower for both groups, but there is still little difference across education groups (0.81 and 0.83). There *is,* however, a major difference in a key dimension: the dispersion of income profiles $\left(\sigma_\beta^2\right)$ is significantly larger for college-educated individuals (.00049) compared to high school-educated individuals (.00020). This difference is also reflected in the age-inequality profiles of income displayed in Figure 3: the cross-sectional variance of log income rises by 0.7 for the former group compared to 0.4 for the latter.

Figure 3: Age effects in the cross-sectional variance of log income by education



Finally, note that the correlation between the slope and the intercept is negative in all rows, consistent with a human capital model: individuals who invest more early in life (perhaps in response to higher learning ability) and suffer from lower income are compensated by higher income growth later in life. Moreover, the correlation is significantly more negative for the college-educated group (-0.7) compared to the rest (-0.25), suggesting that human capital accumulation might be more important for wage growth in high-skill occupations (Mincer (1974), Hause (1980)).

3.4 A comparison to the existing literature

This paper is not the first one to recognize that income profiles may differ across individuals in a systematic way. In fact, earlier papers in this area, perhaps influenced by human capital theory, studied an econometric model featuring profile heterogeneity as a natural starting point. For example, Lillard and Weiss (1979) examined panel data on American scientists, and found evidence of both statistically and quantitatively significant heterogeneity in income growth rates. Hause (1980) reached a similar conclusion using data on Swedish males.

The main evidence *against* profile heterogeneity seems to first appear in MaCurdy (1982). Subsequent work by Abowd and Card (1989), Topel (1990) and Topel and Ward (1992) have examined the same issue using different and typically longer panel datasets—but employing essentially the same method—only to confirm MaCurdy's finding. Instead of estimating an econometric model nesting profile heterogeneity (and showing that it is not significant) these studies used a pre-test for model specification. The basic idea of the test is based on the simple observation that if individuals differ systematically in their income growth rates, then income changes of each individual should be positively autocorrelated. This can be shown easily. Using equation (1), the covariance structure of individual income *growth* is:

$$cov(\Delta y_i^h, \Delta y_i^{h+1}) = \sigma_\beta^2 - \left[\frac{1-\rho}{1+\rho}\sigma_\eta^2\right] - \sigma_\varepsilon^2 \tag{3}$$

$$cov(\Delta y_i^h, \Delta y_i^{h+k}) = \sigma_\beta^2 - \left[\rho^{k-1}\frac{1-\rho}{1+\rho}\sigma_\eta^2\right], \qquad k = 2, ...T-h. \tag{4}$$

Glancing at these formulas, one would expect the first autocovariance to be negative unless σ_β^2 is very large. More importantly, the second and higher order covariances only involve the positive term σ_β^2, and a negative second term that goes to zero at a geometric rate. Thus, one should expect the covariances after a certain lag to be significantly positive if σ_β^2 is positive after all. Hence, the main approach to testing for profile heterogeneity has been to check if higher order autocovariances are greater than zero. The first row of Table 2 reports the results of such a test conducted in Topel (1990) using fifteen years of data from PSID ($1968-1983$). For completeness the second row reports the same statistics using our longer panel. The same pattern can be seen in both samples. The first order covariance is negative as expected, and is statistically significant. However, starting from the second lag, there is no evidence of a positive covariance: they are all negative and statistically not different from zero, seemingly casting doubt on the heterogeneous profile model.[13]

Even though the described test is intuitive and valid in principle, the power may be low for the sample sizes typically found in panel datasets. To illustrate this point it is useful to start by putting some numbers into the formulas above. With the point estimates from Table 1, $cov(\Delta y_i^h, \Delta y_i^{h+k}) = 0.00038 - \left[0.0033 * 0.82^{k-1}\right] < 0$, for all lags $k < 11$, simply because σ_β^2 is quantitatively so small compared to σ_η^2. Hence, one *should* expect to find income changes to display negative and small autocovariances even in the presence of significant profile heterogeneity.

We conduct a simple Monte Carlo analysis to further investigate the power of this test. We simulate income paths with heterogeneous profiles using equation (1) and parameter values from the second row of Table 1.[14] The third row in Table 2 displays the averages of autocovariances over 500 replications along with the standard errors of the sampling distribution. As before, the first order autocovariance is negative and quantitatively large. More importantly, the higher order covariances are negative, close to zero, and statistically insignificant after the second lag, similar to their empirical counterparts. One can similarly compute the serial correlation structure and compare it to those reported in Topel (1990). Again, the same pattern is apparent here (lower panel): very weak negative autocorrelation, not significant after the first lag. In

[13] The variance in Topel (1990) is about three times smaller than ours, probably because he looks at within-job wage changes. MaCurdy (1982) and Baker (1997) report variances closer to ours.

[14] We first simulate income paths for 500,000 individuals. Then we draw 12,000 pairs of observations (Δy_i^h, Δy_i^{h+k}) without replacement for randomly selected initial age, h, and $k = 1, ..11$. We then calculate the first ten autocovariances of income changes using this sample and repeat this exercise 500 times.

Table 2: THE COVARIANCE STRUCTURE OF WAGE OR INCOME GROWTH IN THE DATA AND IN THE BASELINE MODEL

	Lag					
Sample	0	1	2	3	4	5
	Autocovariance					
(1) Data (Topel)	.0476	−.0176	.00058	−.00166	−.00014	−.00067
	(.0019)	(.0014)	(.0008)	(.0007)	(.0008)	(.0007)
(2) Data (This paper)	.1215	−.0385	−.0031	−.0023	−.0025	−.00004
	(.0023)	(.0011)	(.0010)	(.0008)	(.0007)	(.0008)
(3) Model	.0840	−.0329	−.0014	−.0011	−.0007	−.0007
	(.0013)	(.0010)	(.0008)	(.0009)	(.0008)	(.0007)
	Autocorrelation					
(4) Data (Topel)	1.00	−.394	.013	−.039	−.003	−.016
(5) Data (This paper)	1.00	−.317	−.026	−.019	−.021	−.001
(6) Model	1.00	−.391	−.016	−.012	−.009	−.009
	(.000)	(.008)	(.009)	(.010)	(.008)	(.009)

Notes: Standard errors are in parenthesis. The statistics from Topel (1990) are from Table B1 in Appendix B, which are calculated from PSID 1968-83 with 8683 observations. The counterparts from simulated data are calculated using 12,000 observations.

fact, one needs a sample size of 110,000 observations—substantially larger than any panel dataset available—to make the first five covariances statistically significant at 95 percent level.In our view, this simple comparison casts serious doubt on the empirical evidence against the heterogeneous profile model.[15]

Finally, the evidence on profile heterogeneity appears to be robust to plausible generalizations of the econometric specification. For example, one can imagine a richer income process incorporating both fully permanent and mean-reverting shocks in addition to profile heterogeneity. The results in Baker and Solon (2003, p. 313) suggest that this extension is not likely to greatly affect our conclusions: with this more general specification they find that the persistence of the AR(1) component is lower (since permanent shocks are disentangled), and the estimates of σ_α^2 and σ_β^2 are somewhat higher than before. Second, our findings are also consistent with other studies that use different datasets and sample selection criteria. In addition to Lillard and Weiss (1979), Hause (1980) and Baker (1997) mentioned earlier, Haider (2001) estimates an econometric model similar to ours using PSID data, but includes individuals into his sample if they report positive labor earnings for *two years* or more (as opposed to our choice of twenty years or more). His estimates are quite similar to ours (in particular, $\rho = 0.64$, and $\sigma_\beta^2 = 0.00041$), which is reassuring. We conclude that the best available *statistical* evidence

[15] In addition, Abowd and Card (1989) also test if covariances are jointly non-zero (using a χ^2 test). See Baker (1997) for a detailed Monte Carlo analysis of this test.

points to an income process with profile heterogeneity and a stochastic component with relatively modest persistence. In Section 6, we compare the consumption behaviors implied by these different income processes to micro data, to further distinguish between them..

4 Uncertainty about Income Profiles

The different statistical models we estimated in the last section did not require us to take a stand on how much an individual knows about his own income profile. To embed the estimated income process into a life-cycle model, however, we need to be specific about what the individual knows about $\left(\alpha^i, \beta^i\right)$. Given that profile heterogeneity may result from a complex set of factors that are unlikely to be fully known or understood by the individual, especially early on in his career, it does not seem very reasonable to assume that he would have perfect knowledge of these parameters when he first enters the labor market. A more plausible case is one in which an individual enters the labor market with some prior belief (which could incorporate some relevant information unavailable to the econometrician) about his income growth prospects. Over time, it is natural to think that a rational individual will refine these initial beliefs by incorporating the information revealed by successive income realizations. We assume that this updating ("learning") process is carried out in an optimal (Bayesian) fashion. Thus, early in life, an individual perceives a large amount of risk in his lifetime income, arising both from uncertainty about his profile and from unanticipated shocks to income. As the individual learns over time, profile uncertainty is gradually resolved and the total risk is reduced mainly to that from the latter component.

In order to formally define the learning problem we need to be specific about which components of income are observable. In the standard life-cycle model, an individual can back out the stochastic component $(z_t + \varepsilon_t)$ by observing y_t, since income profiles are identical across the population and are known by everyone.[16] Furthermore, a standard assumption in that framework is that the individuals can also observe transitory and persistent shocks separately, so he is able to compute forecasts of his future income using the actual value of the current state. Turning to our model, if $z_t + \varepsilon_t$ was observable in addition to y_t, the true income profiles would be revealed in just two periods, and there would be no role for learning. Second, although we could allow either z_t or ε_t to be separately observable (and still have non-trivial learning), it seems difficult to make a compelling case for why one component would be observable while the other is not. Hence, as the baseline case we study optimal learning about the parameter vector $\left(\alpha^i, \beta^i\right)$ using successive observations on y_t in the presence of the confounding effects of two separate stochastic shocks, z_t and ε_t.

It is convenient to express the learning process as a Kalman filtering problem using the

[16] In the rest of the paper we study a single cohort over time, so h and t are perfectly correlated. To simplify notation we drop reference to h. Also, when it is clear that we are referring to a single individual, we also drop the subscript i.

state-space representation. In this framework, the "state equation" describes the evolution of a vector of state variables that is unobserved by the decision maker. A second (observation) equation expresses the observable variable(s) in the model as a function of the underlying hidden state and some transitory shock. For this problem the state equation can be written as:[17]

$$\mathbf{S}_{t+1}^{i} \equiv \begin{bmatrix} \alpha^{i} \\ \beta^{i} \\ z_{t+1}^{i} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \rho \end{bmatrix} \begin{bmatrix} \alpha^{i} \\ \beta^{i} \\ z_{t}^{i} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \eta_{t+1}^{i} \end{bmatrix} = \mathbf{F}\mathbf{S}_{t}^{i} + \boldsymbol{\nu}_{t+1}^{i}.$$

Even though the parameters of the income profile have no dynamics, including them into the state vector allows us to obtain formulas for updating beliefs recursively using the Kalman filter. Note also that since z_t has a persistent effect on income, it is a relevant unobserved state variable that needs to be included in $\mathbf{S}_t^i$. Next, the process for log income in equation (1) can be re-written as a linear function of the state vector

$$y_t^i = \begin{bmatrix} 1 & t & 1 \end{bmatrix} \begin{bmatrix} \alpha^{i} \\ \beta^{i} \\ z_{t}^{i} \end{bmatrix} + \varepsilon_t^i = \mathbf{H}_t'\mathbf{S}_t^i + \varepsilon_t^i$$

We assume that both shocks have *i.i.d.* Normal distributions and are independent of each other, with $\mathbf{Q}$ and R denoting the covariance matrix of $\boldsymbol{\nu}_t^i$ and the variance of ε_t^i respectively. To capture an individual's initial uncertainty, we model his prior belief over $(\alpha^i, \beta^i, z_1^i)$ by a multivariate Normal distribution with mean $\widehat{\mathbf{S}}_{1|0}^i \equiv (\widehat{\alpha}_{1|0}^i, \widehat{\beta}_{1|0}^i, \widehat{z}_{1|0}^i)$ and variance-covariance matrix:[18]

$$\mathbf{P}_{1|0} = \begin{bmatrix} \sigma_{\alpha,0}^2 & \sigma_{\alpha\beta,0} & 0 \\ \sigma_{\alpha\beta,0} & \sigma_{\beta,0}^2 & 0 \\ 0 & 0 & \sigma_{z,0}^2 \end{bmatrix},$$

where we use the short-hand notation $\sigma_{\cdot,t}^2$ to denote $\sigma_{\cdot,t+1|t}^2$. After observing y_t^i in each period, an individual's *belief* about the unobserved vector $\mathbf{S}_t^i$ has a normal posterior distribution with a mean vector $\widehat{\mathbf{S}}_{t|t}^i$, and covariance matrix $\mathbf{P}_{t|t}$. Similarly, let $\widehat{\mathbf{S}}_{t+1|t}^i$ and $\mathbf{P}_{t+1|t}$ denote the *one-period-ahead forecasts* of these two variables respectively. These two variables play central roles in the rest of our analysis. Their evolutions induced by optimal learning are given by:

$$\begin{aligned} \widehat{\mathbf{S}}_{t|t}^i &= \widehat{\mathbf{S}}_{t|t-1}^i + \mathbf{P}_{t|t-1}\mathbf{H}_t \left[\mathbf{H}_t'\mathbf{P}_{t|t-1}\mathbf{H}_t + R\right]^{-1}\left(y_t - \mathbf{H}_t'\widehat{\mathbf{S}}_{t|t-1}^i\right) \qquad (5) \\ \widehat{\mathbf{S}}_{t+1|t}^i &= \mathbf{F}\widehat{\mathbf{S}}_{t|t}^i, \end{aligned}$$

[17] Vectors and matrices are denoted by bold letters throughout the paper.

[18] A "^" over a variable denotes a belief or a forecast and the subscript $t_2|t_1$ denotes forecast of (or belief about) a variable in time t_2 given the information set in t_1 (if $t_1 = t_2$).

and

$$\begin{aligned}
\mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}_t \left[\mathbf{H}_t'\mathbf{P}_{t|t-1}\mathbf{H}_t + R\right]^{-1} \mathbf{H}_t'\mathbf{P}_{t|t-1} \\
\mathbf{P}_{t+1|t} &= \mathbf{F}\mathbf{P}_{t|t}\mathbf{F}' + \mathbf{Q}.
\end{aligned} \tag{6}$$

Finally, log income has a Normal distribution conditional on an individual's beliefs:

$$y_{t+1}^i | \widehat{\mathbf{S}}_{t|t}^i \sim N\left(\mathbf{H}_{t+1}'\widehat{\mathbf{S}}_{t+1|t}^i, \mathbf{P}_{t+1|t}\right). \tag{7}$$

As is typical with Bayesian updating the covariance matrix evolves independently of the realization of y_t^i, and is also deterministic in this environment since $\mathbf{H}_t$ is deterministic. Moreover, one can show from equation (6) that the posterior variances of α^i and β^i are monotonically decreasing over time, so with every new observation beliefs become more concentrated around the true values. (This is not necessarily true for $\sigma_{z,t}^2$ which may be non-monotonic depending on the parameterization.)

As mentioned above, the income risk perceived by an individual upon entering the labor market can be quite substantial if he is sufficiently uncertain about his own income profile. However, since this uncertainty is resolved over time through learning, its quantitative importance critically depends on the speed of learning. In a variety of learning models a large fraction of uncertainty is resolved rather quickly, thus it is essential to investigate this issue in the present framework. To provide a comparison, first consider learning about the mean of an *i.i.d.* random variable with variance σ_a^2 and prior variance of mean equal to σ_v^2. Recall that in this problem the posterior variance is proportional to $1/\left(1 + n\left(\sigma_v^2/\sigma_a^2\right)\right)$ after n observations. As long as the data are not too noisy (that is, σ_v^2/σ_a^2 is not too small) this ratio shrinks rapidly with the first few observations, leaving a smaller role for learning in the remaining periods.

Learning is substantially more gradual in our model, and its effects are prolonged—extending throughout the life-cycle—for three reasons. First, although income shocks are not completely permanent, with a persistence of 0.8 they are also far from being *i.i.d.* As a result labor income can deviate from its trend for extended periods of time, slowing down learning about the profile considerably. Second, a noteworthy feature of the present model is that individuals learn about a slope coefficient $\left(\beta^i\right)$ whose contribution to income grows linearly with horizon. So, even though beliefs about β^i become more precise over time ($\sigma_{\beta,t}^2$ gets smaller), its contribution to income uncertainty is amplified by t^2. Thus, loosely speaking, unless $\sigma_{\beta,t}^2$ shrinks faster than $1/t^2$, the effect of profile uncertainty on perceived income risk will grow with horizon. Third, every period individuals update their beliefs about a three-dimensional parameter vector using a single new observation on income, which further slows down the speed of learning. In the rest of this section we elaborate on these three points.

We first quantify the effect of persistence on the speed of resolution of uncertainty. A useful measure of income risk is the mean squared error (MSE) of the forecast of future income at different horizons given by:[19]

$$E_t\left(y_{t+s}-\widehat{y}_{t+s|t}\right)^2 = \mathbf{H}'_{t+s}\mathbf{P}_{t+s|t}\mathbf{H}_{t+s}+R, \tag{8}$$

$$\text{where} \quad \mathbf{P}_{t+s|t} = \mathbf{F}^s\mathbf{P}_{t|t}\mathbf{F}'^s+\sum_{i=0}^{s-1}\mathbf{F}^i\mathbf{Q}\mathbf{F}'^i. \tag{9}$$

In the homogenous profile model, the MSE simplifies to

$$\Lambda^{hom}_{t+s|t}=E_t\left(z_{t+s}-\widehat{z}_{t+s|t}\right)^2+\sigma^2_\varepsilon,$$

which can be obtained by setting the prior variances of $\left(\alpha^i,\beta^i\right)$ to zero in equation (8). This expression makes clear that the only sources of risk in this case comes from the stochastic component of income. Turning to the heterogeneous profile model, since we are interested in measuring the risk due to profile uncertainty, it is useful to focus on the increment in MSE by subtracting the risk that is due to income shocks:[20]

$$\Lambda^{net}_{t+s|t}\equiv\Lambda^{het}_{t+s|t}-\Lambda^{hom}_{t+s|t}=\sigma^2_{\alpha,t}+\left[2(\sigma_{\alpha\beta,t})\,(t+s)+\left(\sigma^2_{\beta,t}\right)(t+s)^2\right]+\kappa_{t+s|t}, \tag{10}$$

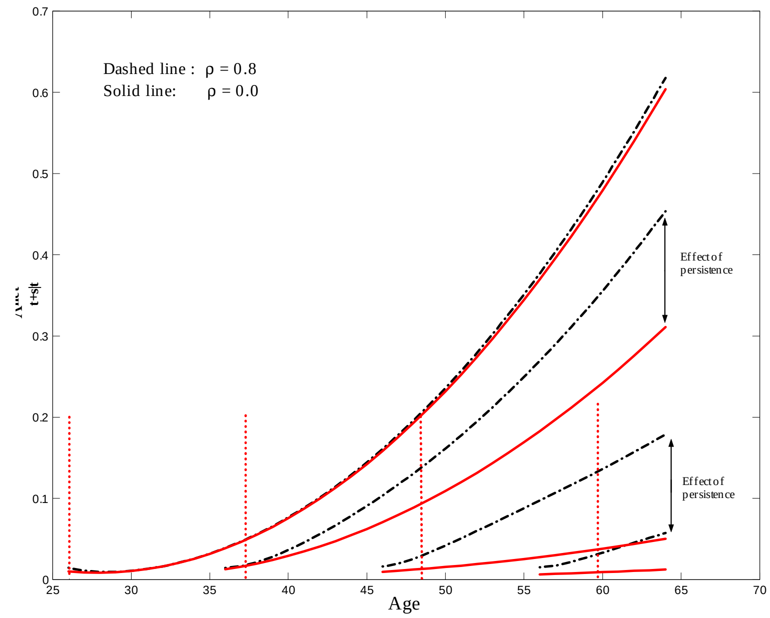
which is again obtained using equation (8). This expression gives the amount of income risk at different horizons in the future (given by s), as perceived by an individual at age t. To determine the shape of this profile, first notice that the second order moments appearing in the first three terms only depend on t and not on the horizon s. This follows from equation (9) noting that the upper 2×2 block of $\mathbf{F}$ is an identity matrix. Second, the estimated correlation between the slope and intercept terms are negative, so beliefs about their covariance $(\sigma_{\alpha\beta,t})$ will also be negative implying a linear decreasing term. And third, the quadratic term has a positive coefficient. Thus the first three components imply that—with our baseline parameterization based on Table 1—the MSE is an increasing quadratic function of horizon. Finally, while z^i_t is independent of $\left(\alpha^i,\beta^i\right)$, the joint updating of beliefs naturally induces a correlation between these two components. The last term, $\kappa_{t+s|t}$, contains the corresponding covariances; it is decreasing as a function of s for fixed t, but does not materially affect the shape of this profile.

In figure 4 the solid lines plot $\Lambda^{net}_{t+s|t}$ for $\rho=0$, and $t=0,10,20$, and 30, and the dashed lines plot the same for $\rho=0.8$. Notice that learning is slow. Even when income shocks are completely transitory, uncertainty about future income takes a long time to resolve. For example, after ten years less than half of the initial uncertainty is eliminated: the predicted variance of log income at retirement is still above 0.3; it takes roughly 20 years for this uncertainty to go down significantly. Second, and more importantly, when the persistence of income shocks go up,

[19] For example, in the homongenous profile model, this MSE will be equal to the cross-sectional variance of log income at different ages, because individuals may end up anywhere in that distribution.

[20] The superscripts *het* and *hom* indicate heterogenoug- and homogenous-profile models respectively.

Figure 4: THE EFFECT OF PERSISTENCE ON THE RESOLUTION OF PERCEIVED INCOME UNCERTAINTY



learning speed slows down even further: by the time an individual is 35 years old, less than 22 percent of income risk at retirement will have been resolved when $\rho = 0.8$, compared to almost 50 percent when $\rho = 0$. At age 45, the variance at retirement is still about 0.2, which is twice the uncertainty resulting from the persistent shock, whereas it is a mere 0.05 for the case with $\rho = 0$.

It should be noted that learning about the intercept term adds little to the perceived risk of income and its speed of resolution. Even considering an order of magnitude increase in σ_α^2 from 0.02 to 0.2, has a minor effect on the graphs in figure 4.

5 A Life-cycle Model with Optimal Learning

We now study the consumption-savings decision of an individual in an environment with profile heterogeneity and Bayesian learning as described in the previous section. We refer to this framework as the "profile heterogeneity with uncertainty" (PHU) model. For comparison we will also analyze a simplified version of this problem where uncertainty is eliminated so each individual knows his true profile ("profile heterogeneity with certainty" model, PHC).

There is a vast literature studying a variety of economic problems in realistic multi-period life-cycle models.[21] Our goal is to mainly understand the working of this model, so we prefer

[21] See for example the references cited in Hubbard, Skinner and Zeldes (1994), Storesletten, Telmer and Yaron (2003), and Gourinchas and Parker (2002).

to keep the specification simple. Specifically, consider an individual who lives for T^* years and works for the first T years of his life, after which he retires. Individuals do not derive utility from leisure and hence supply labor inelastically. While working the income process of an individual is given by equation (1). Once retired he receives a pension equal to a fraction, Φ, of his labor income in the last period of working life. While this specification is admittedly much simpler than the Social Security system, it has the advantage of abstracting from the significant redistribution and risk-sharing inherent in those more realistic pension plans, and consequently from their effects on the consumption-savings decision which may distract from the focus of our analysis. We do however investigate an extension of this pension system in the robustness analysis. Finally, there is a risk-free bond that sells at price P^b (with a corresponding interest rate $r^f \equiv \frac{1}{P^b} - 1$). Individuals can also borrow at the same interest rate up to an age-specific borrowing constraint $\underline{W}_{t+1}$, which will be specified below.

The relevant state variables for this dynamic problem are the asset level, ω_t, the current income, y_t, and the last period's forecast of the true state in the current period, $\widehat{\mathbf{S}}_{t|t-1}$. Although given the last two variables, one can obtain both $\widehat{\mathbf{S}}_{t|t}$ and $\widehat{\mathbf{S}}_{t+1|t}$ using equation (5) (which means that the individual knows the latter two vectors at the time of current period decision) our current choice is more suitable for computational reasons. Notice that only beliefs about the income processes are state variables, and not the true values. In the following equations we include the superscript i in individual-specific variables to distinguish them from aggregate variables. Then, the dynamic problem can be written as

$$
\begin{aligned}
V_t^i(\omega_t^i, y_t^i, \widehat{\mathbf{S}}_{t|t-1}^i) &= \max_{c_t^i, \omega_{t+1}^i} \left\{ U(c_t^i) + \delta E\left[V_{t+1}^i(\omega_{t+1}^i, y_{t+1}^i, \widehat{\mathbf{S}}_{t+1|t}^i) | \widehat{\mathbf{S}}_{t|t-1}^i \right] \right\} \\
& \quad s.t \\
c_t^i + P^b \omega_{t+1}^i &= \omega_t^i + y_t^i \qquad (11) \\
\omega_{t+1}^i &\geq \underline{W}_{t+1} \qquad (12) \\
& \quad eq.\ (5, 6) \qquad (13)
\end{aligned}
$$

for $t = 1, ..., T-1$. The evolutions of the vector of beliefs and its covariance matrix are governed by the Kalman recursions given in equations $(5, 6)$. Moreover, given that the only state variable that is random at the time of decision is next period's income, the expectation is taken with respect to the conditional distribution of y_{t+1}^i given by equation (7). After retirement, labor income is constant and there is no other source of uncertainty or learning, so the problem simplifies significantly:

$$
\begin{aligned}
V_t^i(\omega_t^i, y^i) &= \max_{c_t^i, \omega_{t+1}^i} \left[U(c_t^i) + \beta V_{t+1}^i(\omega_{t+1}^i, y^i) \right] \\
& \quad s.t \\
y^i &= \Phi y_T^i \\
& \quad eq.\, (11, 12)
\end{aligned}
$$

for $t = T, ...T^*$, where y^i does not have a t subscript to emphasize that it is constant over time (though it is still individual-specific), and $V_{T^*+1} \equiv 0$.

5.1 Quantitative analysis and parameterization

There is no analytical solution to the dynamic optimization problem stated in the previous section, so we solve the model using numerical methods. The numerical solution is complicated by the fact that there are five continuous state variables and four of them (excluding ω_t^i) depend on each other as a result of learning. This co-dependence poses a computational challenge, and in particular makes the solution of the problem on a rectangular grid impractical. We develop an algorithm to tackle these issues, which could be useful for solving similar problems. Further discussions of computational issues as well as the details of our algorithm are provided in the computational appendix.

Parameterization. A model period is one year of calendar time. Individuals enter the labor market (are born) at age 25, retire at 65 and are dead by age 90. The period utility function is assumed to take the CRRA form $\left(U\left(C\right) = \frac{C^{1-\phi}}{1-\phi}\right)$ with a relative risk aversion coefficient of 2. The subjective time discount rate, δ, is set equal to 0.96. P^b is also set equal to 0.96 so that in a purely deterministic world individuals would prefer a completely flat consumption profile. Finally, we set Φ equal to 0.25 in the baseline case (which is admittedly rather arbitrary), and also report results for $\Phi = 0.5$.

The parameters of the stochastic component of income are taken from Table 1. Although the estimation of the covariance matrix pins down the variances of α and β, it does not identify their means. The intercept term, α, is a scaling parameter and has no effect on results, so it is normalized to 1.5 for computational convenience. The mean of β is set to the mean growth of log income in our data set: it is equal to 0.9 percent per year for the whole sample, and 0.7 percent and 1.2 percent for the group of low and high educated individuals respectively. Since income is log-Normally distributed, together with the calibrated variances, these numbers imply that the growth rate of mean income is 1 percent for the whole sample, 1.4 percent and 0.8 percent for college- and high school-educated groups respectively.

In the baseline model, we set individuals' initial beliefs as follows. The prior mean growth rate is set equal to its true population mean, β^A. The covariance matrix of priors is

$$\mathbf{P}_{1|0} = \begin{bmatrix} 0.02 & -.00063 & 0.0 \\ -.00063 & .00038 & 0.0 \\ 0.0 & 0.0 & 0.029 \end{bmatrix},$$

where the non-zero elements are set equal to the values estimated for the *true* process in Section 3. Implicit in these choices is then the assumption that the individual does not have more information than the econometrician to predict his income profile at the beginning of his life.[22]

[22] Notice that there are further observable variables that could be included in the first stage regression and

Table 3: Baseline Parameterization

Annual model		
Parameter		Value
δ	Time discount rate	0.96
P^f	Price of risk-free bond	0.96
ϕ	Relative risk aversion	2
β^A	Avg. inc. growth for all households	0.009
β^C	Avg. inc. growth for college educ.	0.012
β^H	Avg. inc. growth for high school educ.	0.007
T	Retirement age	65
T^*	Age of death	90
Φ	Replacement rate	0.25
$\mathbf{P}_{1\|0}$	The variance of prior beliefs	See text

Note: The parameters of the income process are taken from corresponding rows of Table 1.

While this seems unlikely to be literally true, this choice provides a useful benchmark to gauge how much mileage one can get by allowing uncertainty about one's income profile. Second, recall that this assumption is also behind the homogenous profile model which attributes all income risk to unanticipated shocks, and the predictable component is given by the function g. Finally, it is important to point out that conditioning on more information does not necessarily imply less uncertainty for all individuals: for example, if a college graduate knows that the dispersion of income growth rates are different depending on education level, his prior variance $\sigma^2_{\beta,0}$ will be 0.00049 (the dispersion of college graduates) instead of 0.00038 which is the population average. We consider this case below.

As for the calibration of the borrowing constraint, we have a couple of considerations in mind. First, it is desirable to impose a loose constraint so as not to confound the effects of profile uncertainty and learning—the primary focus of this paper—with those of borrowing frictions. The loosest constraint is implied by the condition that an individual cannot have debt at the time of death. In this case, in any given period an individual can borrow up to the point where he can still pay back all of his debt even if he happens to face the lowest possible income realization for the rest of his life. In our framework, this requirement implies that each individual will face a different natural limit, unlike in a standard life-cycle with ex-ante identical individuals. However, such a specification would also imply that the constraints themselves contain information about an individual's profile which would then need to be optimally incorporated into beliefs. This would further complicate the model without providing much

increase the predictable component. Except for education (which we condition on later), these variables do not significantly increase the predictive power of the first stage regression though, suggesting that many such variables, such as ability, are hard to measure by proxies.

additional insight. As a compromise, we allow individuals to borrow up to a fraction of the natural borrowing limit implied by their prior beliefs (which is assumed to be identical across individuals). In other words, this is the natural limit that credit institutions would enforce on individuals if only time-0 information was available. Notice that since y_t is log-normally distributed, the lowest income realization can be arbitrarily close to zero, so we truncate the normal distribution at three standard deviations to provide a proper lower bound. In the next section we provide a quantitative illustration of how tight the borrowing constraint is. We should also mention that in our baseline specification this constraint is almost never binding.

6 Model Results

The analysis in this section has two main goals. We first embed the income process estimated in Section 3 in a life-cycle model to examine how the existence of (i) profile heterogeneity and (ii) Bayesian learning shape individuals' consumption-savings decision. Second, we investigate if the implied consumption behavior is consistent with empirical facts, and especially with certain findings that have commonly been interpreted as evidence supporting the high persistence of income shocks (Deaton and Paxson (1994), Carroll and Summers (1991)).

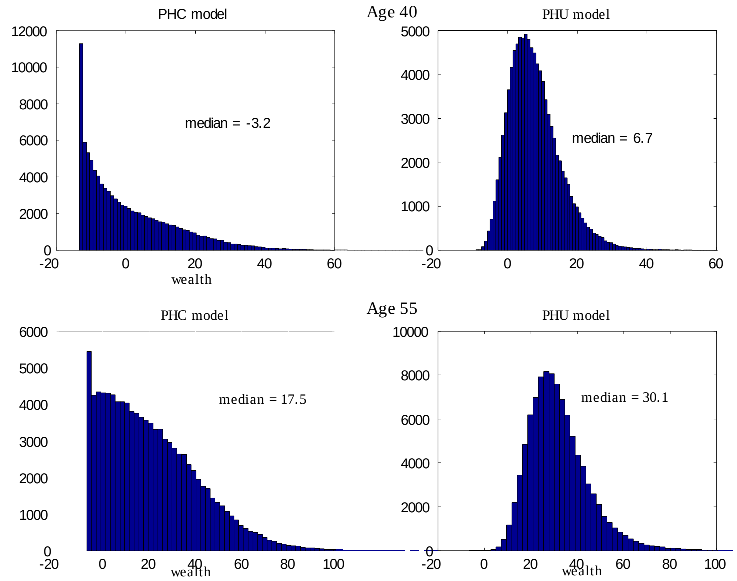
6.1 The effect of profile uncertainty on life-cycle savings

We begin by studying the precautionary savings generated in this model. The size and persistence of income shocks are key determinants of precautionary savings desired by individuals. While the autoregressive process in our model has only modest persistence, Bayesian learning introduces random walk components into the level *and* the growth rate of perceived earnings. It is important to quantify the amount of savings generated by this latter component.

In figure 5 we plot the cross-sectional distributions of wealth for two cohorts: those who are 40 (top) and 55 (bottom) years old. Those on the left are from the PHC model, where there is no uncertainty and hence no learning, and those on the right are from the PHU counterparts. The spikes at the lower bounds of the distributions in the left panel show that many individuals are borrowing constrained in the PHC model, whereas virtually no one is constrained in the PHU model. This is also true over the entire life-cycle: only 0.02 percent of the population ever hit their borrowing limit in the latter model whereas this number is 24.6 percent in the former.

One can alternatively look at differences in *aggregate* wealth accumulation, which is significantly higher in the PHU model (21.6 versus 13.4) as could be anticipated from the previous figure. This difference would be larger if so many individuals were not up against their borrowing constraints in the PHC model and could borrow further. The *median* wealth measure is robust to this problem—as long as less than fifty percent of the population is constrained. The median individual owns about twice the wealth over his lifetime in the PHU model compared to the PHC model (19.4 versus 9.8). Finally, notice that part of the wealth accumulation is for life-cycle reasons, i.e., to provide retirement income. To provide a clearer comparison of precau-

Figure 5: The effect of (Income) Profile uncertainty on wealth accumulation



tionary savings across the two models, we increase the replacement rate to 0.5, resulting in less life-cycle savings. In this case the ratio of wealth levels goes up higher, to 2.44. Overall, these comparisons show that profile uncertainty substantially alters savings behavior throughout the life-cycle, despite the fact that individuals learn to resolve this uncertainty.

Since each individual has a different income profile in this model, there is a potentially interesting relationship between an individuals' profile and his savings over the life-cycle. To provide a benchmark, in a purely deterministic world (and assuming $\alpha^i \equiv 0$), consumption smoothing implies that the wealth holdings of an individual is perfectly *negatively* correlated with the slope of his income profile. In fact, this implication typically holds true even with sizeable income uncertainty, and often yields counterfactual results, such as the prediction that college graduates will save less (or borrow more) than lower educated individuals who have lower income growth rates (Davis, Kubler and Willen (2003)). More generally, these models typically imply that the income-rich will be the wealth-poor. In the U.S. data, wealth holdings are increasing in both education and income, so both of these implications are inconsistent with empirical findings (c.f., Hurst, Luoh and Stafford (1998) and the references therein).

We begin by analyzing this question in the PHC model first. The existence of autoregressive income risk does not qualitatively change the conclusion reached above: the correlation between an individual's wealth, ω_i^h, and the slope of this profile, β^i, starts from -0.9 at age 25, and although it gradually increases over time, it remains negative until age 60, with an average value of -0.58. In contrast, in the PHU model, this correlation is positive at every age, and increases monotonically over time to reach 0.79 at retirement, with an average value of 0.37.

One can also compute the correlation of wealth at each age with life-cycle income, $\alpha^i + \beta^i h$. The average value is -0.40 in the PHC model compared to 0.54 in the PHU model. These findings show that profile uncertainty not only results in more savings on average, but more importantly, it overturns the counterfactual implication that those with high income growth are most willing to borrow over the life-cycle. This feature of the PHU model will prove important in understanding some stylized facts and we will return back to this point later below.

6.2 The age-inequality profile of consumption

Deaton and Paxson (1994) document the striking rise of within-cohort inequality of consumption and labor income over time. In particular, the cross-sectional variance of log consumption (per adult equivalent) increases by about 0.25 to 0.30—roughly corresponding to the doubling of inequality—over a cohort's life-cycle. For completeness, we replicate their finding as closely as possible using the same dataset and sample period (Consumer Expenditure Survey, $1980-90$). The broken line in figure 6 displays the resulting age-inequality profile, which is essentially the same as the one presented by Deaton and Paxson.[23]

The mere fact that there is fanning-out in the consumption distribution is not so surprising, as this would be implied, for example, by the permanent income theory. What *is* surprising though is its immense magnitude. Deaton and Paxson discuss several potential explanations and find the existence of persistent (uninsurable) idiosyncratic shocks to be the most promising candidate.[24] Recently, Storesletten et. al (2003) have tested this conjecture and have concluded that a life-cycle model can quantitatively match the rise in inequality observed in the data *if* income shocks are extremely persistent. This indirect evidence from consumption data has been interpreted as lending further support to the (earlier) estimates of high persistence obtained from income data.

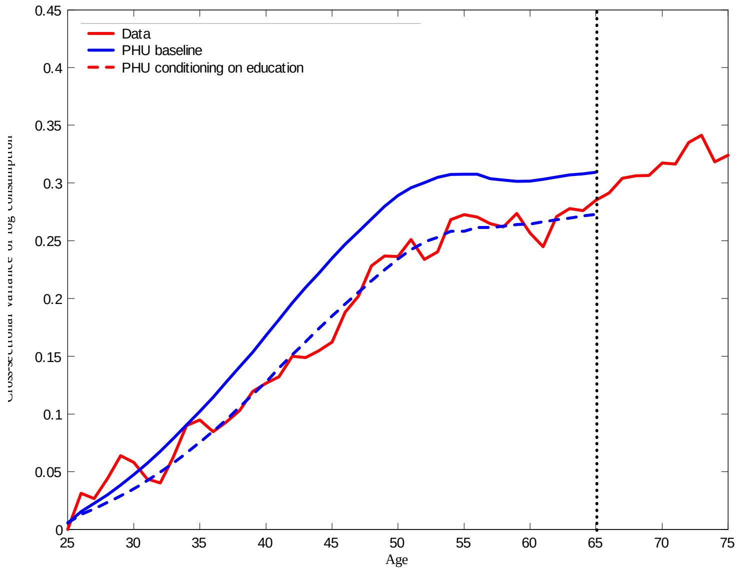
The findings in Section 3, however, indicate that—once we allow for profile heterogeneity—estimates of persistence are much lower, in the neighborhood of 0.8. Thus, income shocks are not nearly persistent enough to generate any significant increase in dispersion on their own.[25] Of course profile uncertainty introduces another source of risk that is not present in

[23] Following these authors we regress raw variances for each age-year cell on a set of age and cohort dummies and report the coefficients on the age dummies. To reduce the number of cohort dummies estimated, we group individuals between 25 and 29 as the first cohort, between 30 to 34 as the second cohort and so on. The age dummy in the first year is normalized to zero.

[24] Another explanation entertained by these authors is the possible non-separability between consumption and leisure in the utility function, combined with heterogeneity in income profiles. In this case, consumption inequality would increase over time but so would inequality in hours worked—a prediction not borne out in the data. (Storesletten et. al 2001).

[25] To quantify this assertion, consider a simplified version of our baseline model, where we eliminate profile heterogeneity (and consequently, learning). This basic framework is now essentially the same as the one studied by Storesletten et. al (2003) with some inessential differences. Using the estimates from the second row of Table 1 (and in particular $\widehat{\rho} = 0.82$), this model generates an increase in the cross-sectional variance of consumption by less than 0.07 over the life-cycle—a quarter of its empirical counterpart. Moreover, dispersion increases only in the first 5-6 years of a cohort's life and remains roughly constant thereafter, unlike in the data where it rises

Figure 6: The age-inequality profile of consumption in the CEX data and in the PHU model

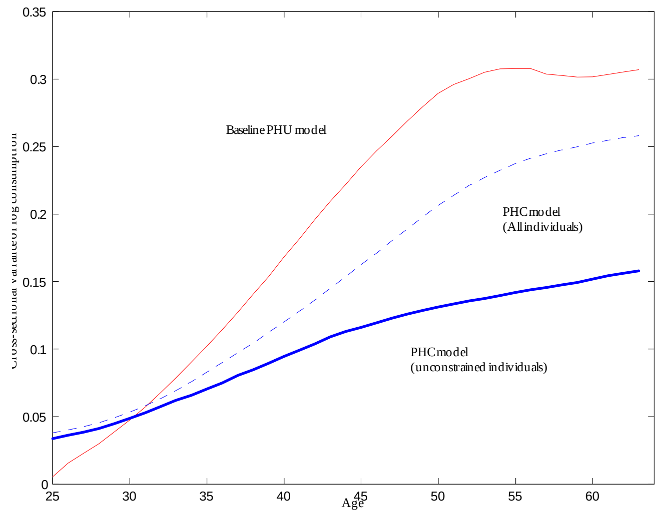


the standard life-cycle model, and one that plays a key role in understanding the empirical facts about consumption. The (top) solid curve in figure 6 plots the age-inequality profile of consumption from the baseline PHU model. There are two main points—one quantitative, and one qualitative—to observe. First, consumption inequality rises significantly, by about 0.3 over the life-cycle, which is even higher than in the U.S. data. Second, the inequality profile is approximately linear (and slightly convex) and rises through most of the life cycle, in line with the qualitative properties of the consumption data. We now discuss these two features in more detail.

First, standard life-cycle models typically include an education dummy into g in equation (1), allowing for different slope coefficients for each education group, which are known by individuals. This is not the case in our baseline model where all individuals are assumed to have the same prior belief about their income prospects. As a next step then, we let individuals condition their initial beliefs (and the mean growth rate of their income) on their education level. In addition, individuals in each education group now face the income process corresponding to that group as estimated in Section 3.3. We refer to this extension as the PHU-ED model. We solve the dynamic problem of each group separately and then compute the age-inequality profile of consumption for the entire population (dashed line in figure 6). The age-inequality profile now rises somewhat more slowly than before, and provides a very good fit to its empirical counterpart. Although this result is encouraging in terms of showing the potential of this

steadily throughout the life-cycle.

Figure 7: Comparing the age-inequality profiles in the PHU and PHC models



model to generate substantial increase in consumption inequality we should be careful not over-interpret this result: the exact amount of rise in dispersion depends on the covariance matrix of priors, $\mathbf{P}_{1|0}$, (among other parameters) and a careful calibration of these covariances is needed before a definitive empirical statement can be made.

The PHU model has two features, profile heterogeneity and profile uncertainty, that are not present in standard life-cycle models, so it is instructive to decompose the contribution of each component to increasing inequality. First, with only profile heterogeneity (PHC), inequality rises by 0.22 (dashed line in figure 7). This number seems surprisingly large given that there is little income risk (coming from the AR(1) component only) in this model. But recall that a significant fraction—24.6 percent—of the population in this model are borrowing constrained at some point in their life-cycle, mainly because there is little incentive for precautionary savings. For these individuals, consumption and income move in locksteps in those periods. Thus part of the fanning-out in the consumption distribution is generated, somewhat mechanically, by the fanning-out of income. As noted earlier, this mechanism has the counterfactual implication that those who are constrained will predominantly be the income-rich.

Incorporating profile uncertainty into this framework has two opposite effects. On the one hand, it generates more precautionary wealth accumulation, effectively relaxing borrowing constraints, and *eliminating* the rise in dispersion due to binding constraints. On the other hand, optimal learning introduces permanent innovations into the slope of the perceived income process which results in *more* consumption inequality. Hence, profile uncertainty does not only result in more fanning-out than in the PHC model (0.31 instead of 0.22) but also changes its nature. Since these two effects work in opposite directions, it is of interest to quantify each one separately. To provide a measure of the second effect (due to learning), one possible approach is to focus on unconstrained individuals in each model and compare the rise in consumption inequality among individuals in each group. In the PHC model, inequality among unconstrained individuals rises by 0.12 over the life-cycle. In the PHU model, virtually nobody is constrained, so the corresponding number is the same as before: 0.31. The difference between these two numbers ($0.31 - 0.12 = 0.19$) roughly corresponds to the second effect of uncertainty.[26] This suggests that uncertainty about income profiles has a significant quantitative effect on the rise of consumption inequality, and explains more than half of the fanning out over the life-cycle.

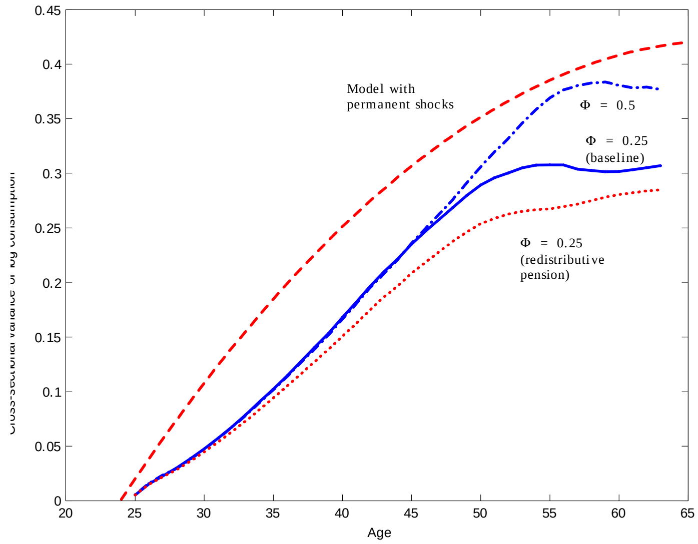
Next, we briefly examine how the age-inequality profile depends on retirement income. First, when the replacement ratio is raised to 50 percent of last period's income ($\Phi = 0.5$) , there is little change in the inequality profile until about age fifty (figure 8). However, dispersion continues to rise past that age, unlike in the baseline model, to reach a maximum of 0.38, because with a higher replacement rate individuals save less for retirement and even at older ages their accumulated wealth constitutes a smaller fraction of their remaining lifetime resources. Thus, consumption is tightly linked to the labor income (as opposed to accumulated wealth) and is strongly influenced by any uncertainty about it.

As a second extension, we modify the retirement system to incorporate redistribution (inherent in the Social Security system) in a simple way. Specifically, retirement income is given by $y^i = \Phi\sqrt{y_T^* y_T^i}$, where y_T^* is the median income at age 65, and Φ is kept at 0.25. This concave scheme implies that those with (above-) below-median income at T, will receive a (smaller) larger pension than before. The dotted line in figure 8 shows that inequality rises slightly less than in the baseline case, but still reaches 0.28 at retirement age.

Finally, while both a life-cycle model with permanent income shocks and the PHU framework are consistent with rising consumption inequality, it is useful to point out one important difference. In general, the amount of inequality generated by the former model is quite sensitive to the persistence of shocks. For example, if income shocks are permanent, an annual standard deviation of 13 percent per year is sufficient to match the fanning out in the data. However, although one can certainly think of some shocks that are truly permanent, it seems harder to imagine that this is true for the "typical" income change. In fact, even when profile heterogeneity is ignored (so $\widehat{\rho}$ is biased upward) estimates of persistence are typically less than 1. (The point estimates are between 0.94 and 0.98 in MaCurdy (1982), Hubbard, et. al (1994), Baker (1997), and Heathcote et. al (2003), Storesletten et. al (2004)). These values are quite small in terms of their implications for the rise in consumption inequality. For example, using the estimates from Table 1 for the whole population (in particular $\widehat{\rho} = 0.988$), the rise in inequality

[26] A caveat to this computation (and hence the qualification "roughly") is that, as mentioned above, constrained individuals in the PHC model are mainly those with high β^is, so by eliminating them we are effectively truncating the upper tail of the income growth distribution. As a result, income inequality rises more slowly in this subsample (call PHC-uc) than in the PHU model, making the comparison of consumption inequalities somewhat difficult. However one can re-calibrate the dispersion of βs in the PHU model to match that in the PHC-uc sample. In this case, consumption inequality rises by 0.25 in the PHU model which is still twice the value of 0.12 in PHC-uc sample.

Figure 8: The effect of retirement income on the age-inequality profile



(without profile heterogeneity) is 0.16—half of the empirical value. Moreover, the estimates for each education group are only slightly lower ($\widehat{\rho}^C = 0.979$ and $\widehat{\rho}^H = 0.972$), but generate rises in inequality that are considerably smaller (0.12 and 0.11 respectively). An advantage of the mechanism in our model is that shocks to the "perceived" income process will always be permanent, regardless of the persistence of the underlying shocks—thanks to Bayesian learning—and will thus result in significant fanning out of the consumption distribution.

6.3 The non-concavity of the age-inequality profile of consumption

A second feature of the age-inequality profile emphasized by Deaton and Paxson (1994, fig. 8) is its non-concave shape. Examining consumption data from three countries—the U.S., the U.K., and Taiwan—these authors find that the age-inequality profile increases nearly linearly in the former and is convex in the latter two countries. The same pattern also holds true in our baseline model with a slightly convex rise early on, followed by a linear segment, which tapers off after age 55. Deaton and Paxson stress this non-concavity because it seems hard to be reconciled with the existence of persistent shocks. Specifically, using the certainty equivalent version of the permanent income model they show that the inequality profile will be concave *if* the income process has a large persistent component. Although they make a number of restrictive assumptions to develop their argument, Storesletten et al. (2003) later study a more flexible model with CRRA utility and a rich set of realistic features and find concavity to be a robust feature of the life-cycle model with persistent shocks.

In the PHU model, non-concavity results mainly from learning about β^i.[27] The main intu-

[27] The fact that income shocks are not very persistent also contributes to the convexity of the profile. But since

ition can be conveyed in the certainty-equivalent version of the permanent income model (i.e., assuming quadratic utility and $\delta\left(1+r^{f}\right)=1$), with $\alpha^{i}\equiv 0$. In this case optimal choice implies that consumption growth will be given by

$$\Delta c_t = \frac{1}{\varphi_t}\left[\frac{r}{1+r}\sum_{s=0}^{T-t}(1+r)^{-s}\left(E_t - E_{t-1}\right)y_{t+s}\right]$$

where $\varphi_t = 1-\left(1/\left(1+r\right)^{T-t+1}\right)$ is the annuitization factor. Basically, this equation states the well-known intuition that consumption is readjusted every period by a fraction of the change in expected lifetime resources (the term in brackets). To simplify the problem even further, assume that income (and not log income) is a linear function of experience with *i.i.d.* innovations, and there are no fixed effects: $y_t^i = \beta^i t + \varepsilon_t^i$.[28] When an individual updates his beliefs in period t, the revision in expected future income is: $\left(E_t - E_{t-1}\right)y_{t+s} = \left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right)(t+s)$. Substituting this expression into the equation above and after performing some tedious but straightforward algebra one can show that

$$\Delta c_t = \left[\left(\frac{1-\gamma}{\gamma}\right) + \frac{(T-t+1)\,\gamma^{T-t+1}}{1-\gamma^{T-t+1}} + t\right]\left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right),$$

where $\gamma = \frac{1}{1+r}$. The second term in the square bracket is nearly linear (with slight convexity) for a range of plausible parameter values, and is increasing in t. Combined with the third term, t, they yield an increasing, approximately linear function in t. The implication is that, as cohorts get older, the response of consumption growth to a fixed amount of adjustment in beliefs about β becomes stronger. To the extent that learning is not very fast, so that $\left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right)$ does not shrink (in absolute value) too quickly with t, the inequality profile generated by slope uncertainty will be convex. This turns out to be the case for plausible parameterizations of the model, especially up to about age 50.

6.4 The co-movement of consumption and income over the life-cycle

Another interesting finding documented in the literature is that consumption tracks income over the life-cycle: it first rises and then falls with income (Carroll and Summers (1991)). Although, this relationship is not consistent with the basic certainty-equivalent version of the permanent income model, some fairly plausible extensions would generate such a hump in consumption. Perhaps the simplest one is to impose borrowing limits, which could cause consumption to rise with income. But these constraints have to be binding quite often, or bind for a large fraction

learning introduces permanent-like changes in the perceieved income process one would like to know if that may result in concavity.

[28] These assumptions are rather innocuous in this context. First, the exponential function is increasing and convex, so the log-linear specification for income in the baseline model will only reinforce the mechanism described here. Second, even though income shocks are not *i.i.d* in our model, their persistence is small enough that they do not create any concavity to overturn this conclusion.

of the population to generate a sizeable hump in average profiles. We discuss this possibility below and argue that, as before, it results in other counterfactual implications.

A second possible explanation for the hump is the precautionary savings motive that arises with more general utility functions (such as CRRA) in response to persistent income shocks (Carroll (1992), Attanasio et. al (1999), among others). In this case, the individual reduces his consumption early in life to build a buffer stock wealth for self-insurance purposes. As the individual's financial wealth grows over time, his consumption depends less on labor income (and more on financial wealth) effectively reducing the uncertainty he faces, thus allowing him to increase his consumption along with income, generating the co-movement.

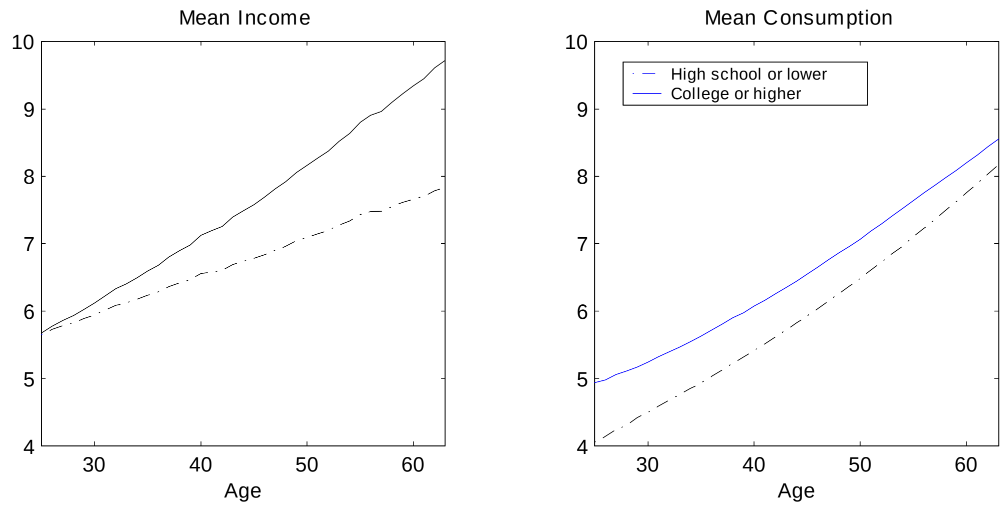
However, a second finding documented by Carroll and Summers poses a challenge to this basic story. These authors find that the consumption profile is steeper for those groups of individuals who have steeper income profiles. For example, both the income *and* the consumption profiles of the college-educated group are steeper than those of the high school-educated group. For a story based on precautionary savings alone to explain this observation, it would require the former group to face either more persistent or larger income shocks than the latter,[29] neither of which we seem to find in the data. For example, the estimates in rows three and five of Table 1 show that (when homogeneous profiles is assumed) there is little difference between the two groups in the persistence and the innovation variance of shocks, which is consistent with existing evidence (c.f., Hubbard, et al. (1994)). In fact, when certain differences are found between these groups, they turn out to be opposite of what is needed to explain the differences in humps: Hubbard et. al (1994) and Carroll and Samwick (1997) find that the variance of persistent shocks goes down with the level of education, which would generate a flatter consumption profile for those with high education.

To illustrate this point, in figure 9 we plot the average consumption and income profiles of the two education groups, implied by a life-cycle model with persistent shocks (but without profile heterogeneity). The parameter values for the income processes are taken from Table 1. The left panel displays the average income profiles, and as expected, it is steeper for college-educated individuals. However, the average consumption profile of this group (right panel) is not any steeper compared to that of high school-educated group, inconsistent with the empirical evidence.

The PHU model offers a possible explanation. Recall that in Section 3, the estimated income process of each group were very similar to each other with one exception: college-educated individuals face a much wider distribution of growth rates compared to those with high school education ($\sigma^2_\beta = 0.00049$ versus 0.0002). Thus, uncertainty about income profiles would induce a stronger precautionary response from the former group compared to the latter

[29] Clearly this is because without income shocks both groups should have the same slope of the income profiles as long as they have the same preferences. Attanasio et. al (1999) suggested that systematic differences in demographics and preferences may generate the observed differences between education groups. For example, if more highly educated individuals are more patient and tend to have larger families they would optimally choose steeper consumption profiles compared to high school graduates.

Figure 9: The average income and consumption profiles by education groups in a Lifeycle model with persistent shocks

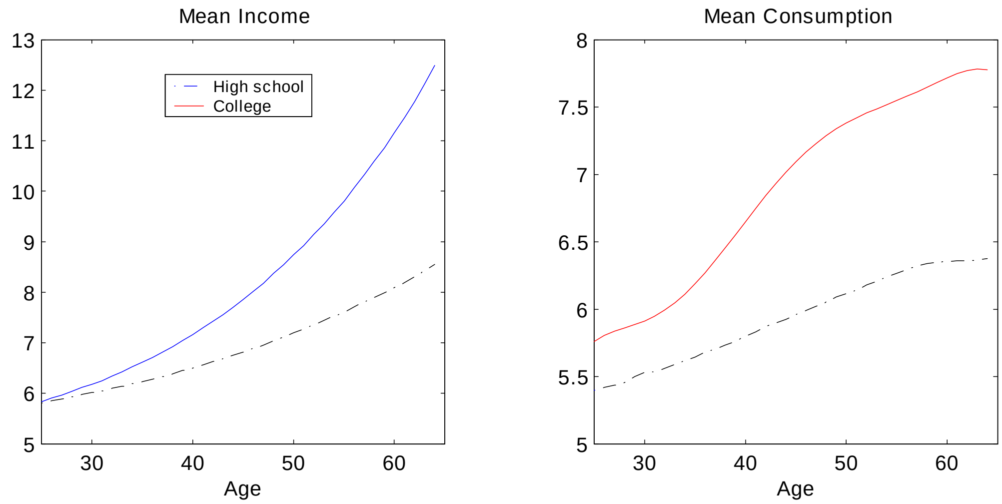


which may result in different slopes of the consumption profiles.[30] In the right panel of figure 10, the average consumption profile is steeper for the college-educated compared to the high school-educated, consistent with empirical evidence. More specifically, consumption rises twice as much over the life-cycle for the former group (35 percent) compared to the latter (17 percent). When the replacement rate is increased to 0.5 from the baseline value of 0.25, the difference in the slopes of consumption becomes even larger: it is 44 percent for higher educated individuals compared to only 18 percent for those with lower education (not shown).

Note finally that consumption would also track income (even without uncertainty about profiles) if there were frequently binding borrowing constraints. But, as discussed before, in the presence of profile heterogeneity, such a model would also imply that constrained individuals have higher income than unconstrained ones. Indeed in the PHC model, the average consumption of constrained individuals is higher throughout the life-cycle and is almost double that of unconstrained ones at retirement (10.1 versus 5.3). These comparisons between the PHC and PHU models show that profile uncertainty should be an integral part of a model with profile heterogeneity, which otherwise yields a number of counterfactual implications.

[30] It is not obvious however that more dispersion for college graduates necessarily means more uncertainty for this group. It is conceivable that higher educated individuals are better able to judge their ability, be better informed about the prospects of income growth in different occupations, etc. These issues deserve further attention in future work. Instead, here we assume that the prior variance is a fixed fraction of the true variance for each group.

Figure 10: THE AVERAGE INCOME AND CONSUMPTION PROFILES BY EDUCATION GROUPS IN THE PHU MODEL



7 Conclusion

In this paper we have studied the persistence of income shocks. We have first argued that the existing evidence against profile heterogeneity does not appear to be strong: an income process with profile heterogeneity would generate the same statistics previously used to reject it. The estimates we obtained indicate substantial heterogeneity in income profiles.

We have then examined the consumption-savings decision of individuals who face such an income process in a life-cycle model. Assuming that individuals do not fully know their profiles, but optimally learn through successive observations on their income, we have found that the model has plausible implications for consumption behavior. First, profile uncertainty is resolved only gradually, and results in significant rise in consumption inequality over the life-cycle. This is despite the fact that income shocks have low persistence. Second, the shape of the age-inequality profile of consumption is approximately linear and exhibits some mild convexity early in life. This feature fits well with the empirical evidence documented by Deaton and Paxson using data from the U.S., the U.K. and Taiwan. In contrast, models with persistent income shocks imply a concave shape. Finally, the model is also consistent with the fact that the consumption profiles of higher educated individuals are steeper than lower educated ones. This happens because the former group faces a wider dispersion of income growth rates thus possibly perceiving higher income lifetime income risk.

Overall, we conclude that income shocks are likely to be substantially less persistent than suggested by the existing literature.

A Appendix

A.1 Data Appendix [To be written]

A.2 Estimation

A.3 Computational Algorithm [To be written]

References

[1] Abowd, John, and David Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica,* Vol. 57, pp.411-445.

[2] Aiyagari, Rao S. (1994): "Uninsured Idiosyncratic Shock and Aggregate Saving," *Quarterly Journal of Economics*, Vol. 109, No.3, pp. 659-684.

[3] Altonji, Joseph G, and Chris Pierret (2001): "Employer Learning and Statistical Discrimination," *Quarterly Journal of Economics,* Vol. 116, pp. 313-350.

[4] Altonji, Joseph G, and Lewis M Segal (1996): "Small Sample Bias in GMM Estimation of Covariance Structure," *Journal of Business and Economics Statistics,* Vol. 109, pp. 659-684.

[5] Attanasio, Orazio P., James Banks, Costas Meghir and Guiglielmo Weber (1999): "Humps and Bumps in Lifetime Consumption," *Journal of Business and Economics Statistics,* Vol. 17, pp. 22-35.

[6] Baker, Michael (1997): "Growth-rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," *Journal of Labor Economics,* Vol. 15, pp. 338-375.

[7] Baker, Michael and Gary Solon (2003): "Earnings Dynamics and Inequality Among Canadian Men, 1976-1992: Evidence from Longitudinal Tax Records," *Journal of Labor Economics,* Vol. 21, pp. 289-321.

[8] Becker, Gary S. (1994): "Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education," University of Chicago Press.

[9] Ben-Porath, Yoram (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy,*.Vol. 75

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

[11] Carroll, Christopher D., and Lawrance Summers (1991): "Consumption Growth Parallels Income Growth: Some New Evidence," in *National Savings and Economic Performance,* eds. D. Bernheim and B. Shoven, Chicago: University of Chicago Press, pp. 305-343.

[12] Carroll, Christopher, and Andrew Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40, No. 1.

[13] Chamberlain, Gary (1984): "Panel Data," in *Handbook of Econometrics*, eds. Z. Griliches and M. Intriligator. Amsterdam: North Holland, pp. 1274-1318.

[14] Constantinides, George M., and Darrell Duffie (1996): "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy,* Vol. 104, pp. 219-240.

[15] Davis, Steve J., Felix Kubler and Paul Willen (2003): "Borrowing Costs and the Demand for Equity Over the Life-cycle," Working Paper, University of Chicago.

[16] Deaton, Angus (1991): "Saving and Liquidity Constraints," *Econometrica,* Vol. 59, pp. 1221-1248.

[17] Deaton, Angus, and Christina Paxson (1994): "Intertemporal Choice and Inequality," *Journal of Political Economy*, Vol. 102, pp. 437-467.

[18] Gourinchas Pierre-Olivier, and Jonathan A. Parker (2002): "Consumption Over the Life Cycle," *Econometrica,* Vol. 70, pp. 47-89.

[19] Haider Steve J. (2001): "Earnings Instability and Earnings Inequality of Males in the United Stated: 1967-1991, *Journal of Labor Economics,* Vol. 19, pp. 799-836.

[20] Harris Milton, and Bengt Holmstrom (1982): "A Theory of Wage Dynamics," *Review of Economic Studies,* Vol. 49, pp. 315-333.

[21] Hause, John C. (1980): "The Fine Structure of Earnings and the On-the-Job Training Hypothesis," *Econometrica,* Vol. 48, pp. 1013-29.

[22] Heathcote, Jonathan, Kjetil Storesletten, Gianluca Violante (2003): "The Macroeconomic Implications of Rising Wage Inequality in the United States," Working Paper, NYU.

[23] Heckman, James, Lance Lochner, and Chris Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, Vol. 1, pp.1-58.

[24] Hubbard, R Glenn, Jonathan Skinner and Steve P. Zeldes (1994): "Precautionary Saving and Social Insurance," *Journal of Political Economy,* Vol. 103, pp. 360-399.

[25] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2003): "Human Capital and Earnings Distribution Dynamics," Working Paper, Wharton School, University of Pennsylvania

[26] Hurst, Erik, Ming Ching Luoh and Frank P Stafford (1998): "The wealth dynamics of American families," *Brookings Papers on Economic Activity,* pp. 267-337

[27] Jovanovic, Boyan (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy,* Vol. 87, pp. 972-990.

[28] Katz, Larry and David Autor (1999): Changes in the Wage Structure and Earnings Inequality, in the *Handbook of Labor Economics,* Vol. 3, Amsterdam: North Holland.

[29] Keane, Michael, and Kenneth Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy,* Vol. 105, pp. 473-522.

[30] Levine David K., and William Zame (2001): "Does Market Incompleteness Matter," *Econometrica,* Vol. 70, pp. 1805-1840

[31] Lillard, Lee A. and Robert Willis (1978): "Dynamic Aspects of Earning Mobility," *Econometrica,* Vol. 46, p. 985-1012.

[32] Lillard, Lee A., and Yoram A Weiss (1979): "Components of Variation in Panel Earnings Data: American Scientists, 1960-70," *Econometrica*, Vol. 47, pp. 437-454.

[33] Lucas, Robert E., Jr. (2003): "Macroeconomic Priorities," *American Economic Review,*

[34] MaCurdy, Thomas (1982): "The Use of Time-Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics,* Vol. 18, pp. 83-114.

[35] Meghir, Costas, and Luigi Pistaferri (2001): "Income Variance Dynamics and Heterogeneity," Econometrica, forthcoming

[36] Mincer, Jacob (1974): "Schooling, Experience and Earnings," New York: Columbia University Press.

[37] Moffitt, Robert A., and Peter Gottschalk (1994): "Trends in the Autocovariance Structure of Earnings in the US, 1969-1987, Working Paper, Brown University.

[38] Storesletten, Kjetil, Chris Telmer and Amir Yaron (2001): "How Important Are Idiosyncratic Shocks? Evidence from Labor Supply," *American Economic Review,* Vol. 91, pp. 413-17.

[39] Storesletten, Kjetil, Chris Telmer and Amir Yaron (2003): "Consumption and Risk-sharing Over the Life-cycle," *Journal of Monetary Economics,* forthcoming

[40] Storesletten, Kjetil, Chris Telmer and Amir Yaron (2004): "Cyclical Dynamics in Idiosyncratic Labor-Market Risk," *Journal of Political Economy*, forthcoming.

[41] Topel, Robert H. (1990): "Specific Capital, Mobility and Wages: Wages rise with Job Seniority," NBER Working Paper No. 3294.

[42] Topel, Robert H., and Michael P. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics,* Vol. 107, pp. 439-79.

[43] Wang, Neng (2004):