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

July 3, 2005

Abstract

The current literature offers two views on the nature of the income process. According to the first view, which we call the "restricted income profiles" (RIP) model (MaCurdy, 1982), individuals are subject to large and very persistent shocks, while facing similar life-cycle income profiles (conditional on a few characteristics). According to the alternative view, which we call the "heterogeneous income profiles" (HIP) model (Lillard and Weiss, 1979), individuals are subject to income shocks with modest persistence, while facing individual-specific income profiles. While labor income data does not seem to distinguish between the two hypotheses in a definitive way, the RIP model is overwhelmingly used to specify the income process in economic models, because it delivers implications consistent with certain features of consumption data. In this paper we study the consumption-savings behavior under the HIP model, which so far has not been investigated. In a life-cycle model, we assume that individuals enter the labor market with a prior belief about their individual-specific profile and learn over time in a Bayesian fashion. We find that learning is slow, and thus initial uncertainty affects decisions throughout the life-cycle allowing us to estimate the prior uncertainty from consumption behavior later in life. This procedure implies that 40 percent of variation in income growth rates is forecastable by individuals at time zero. The resulting model is consistent with several features of consumption data including (i) the substantial rise in within-cohort consumption inequality (Deaton and Paxson 1994), (ii) the non-concave shape of the age-inequality profile (which the RIP model is not consistent with), and (iii) the fact that consumption profiles are steeper for higher educated individuals (Carroll and Summers 1991). These results bring new evidence from consumption data on the nature of labor income risk.

Keywords: Labor income risk, Incomplete markets, Inequality, Consumption-savings decision, Kalman filter.

JEL classification: D52, D91, E21

1 Introduction

When markets are incomplete, labor income risk plays a central role in many decisions that individuals make. Understanding the nature of income risk is thus an essential prerequisite for understanding a wide range of economic questions, such as the determination of wealth inequality (Aiyagari (1994)), the effectiveness of self-insurance (Deaton (1991)), the welfare costs of business cycles (Lucas (2003), and the determination of asset prices (Constantinides and Duffie (1996)), among others.

The current literature offers two views on the nature of the income process. According to the first view, which we call the "restricted income profiles" (RIP) model, individuals are subject to large and very persistent shocks, while facing similar life-cycle income profiles (that only vary across the population with a few observable characteristics). According to the alternative view, which we call the "heterogeneous income profiles" (HIP) model, individuals are subject to shocks with modest persistence, while facing life-cycle profiles that are individual-specific (and typically vary significantly across the population). As we discuss below, while *labor income* data arguably provides more support for the second view, it does not seem to distinguish between the two hypotheses in a definitive way. The goal of this paper is to use the restrictions imposed on *consumption* data by these income processes to distinguish between the two hypotheses.

It is useful to begin with a brief discussion of the empirical evidence from labor income data. For concreteness, suppose that the log income of individual i with t years of labor market experience is given by:[1]

$$\begin{aligned} y_t^i &= \beta^i t + z_t^i \\ z_t^i &= \rho z_{t-1}^i + \eta_t^i, \end{aligned} \tag{1}$$

where β^i is the individual-specific income growth rate with cross-sectional variance σ_β^2; and η_t^i is the innovation to the AR(1) process.

The early papers on income dynamics estimated (versions of) the income process given in (1) without imposing any restrictions on its parameters (Hause (1980); Lillard and Weiss (1979)). These studies found: (i) both statistically and quantitatively significant heterogeneity in income growth rates $\left(\sigma_\beta^2 > 0\right)$, and (ii) the persistence of income shocks to be significantly less than a random walk $(0.5 < \rho < 0.8)$. As noted above, these two features define the HIP model. The human capital model (Becker (1965), Ben-Porath (1967)) implies heterogeneity in income profiles, for example, if individuals differ in their ability level, providing one theoretical motivation for the HIP model.

In an influential paper, MaCurdy (1982) cast doubt on these findings. He tested the simple proposition that if individuals differed systematically in their income growth rates as the HIP model

[1] This income process is a substantially simplified version of the models estimated in the literature, but still captures the components necessary for the present discussion.

suggests, then income changes should be positively autocorrelated. Instead, he found them to be close to zero, and in fact, slightly negative. Subsequent work by Abowd and Card (1989), Topel (1990) and Topel and Ward (1992) tested the same implication using various longer panel data sets only to confirm MaCurdy's conclusion. Based on this work, most of the following literature estimated versions of the income process (1), but now *imposing* $\sigma_\beta^2 \equiv 0$ (the RIP model), and with this restriction, found the persistence of income shocks to be extremely high ($\rho > 0.95$; among others, Hubbard, Skinner and Zeldes (1994), Moffitt and Gottschalk (1995), and Storesletten et al. (2004)). Finally, more recently some authors have extended the analysis of the early HIP literature using more representative panel data sets and richer econometric specifications, and confirmed the results of this early literature (Baker (1997), Haider (2001) and Guvenen (2005)).[2]

Although taken together, the described empirical evidence does not indicate an overwhelming support for the RIP model (and could be interpreted as the opposite), this specification is overwhelmingly used as the income process in economic models. Perhaps an important reason for this preference is that the consumption-savings behavior generated in response to the RIP model is consistent with important empirical facts. For example, Deaton and Paxson (1994) have documented the significant rise in within-cohort consumption inequality over the life-cycle. As conjectured by these authors (and later verified by Storesletten et al. (2003)) a life-cycle model is consistent with this observation if idiosyncratic shocks are extremely persistent. Summarizing the existing empirical 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.

A second empirical observation, documented by Carroll and Summers (1991), is that the average life-cycle consumption profile is hump-shaped: it rises and falls together with labor income over the life-cycle. Again, a life-cycle model can generate this pattern if income shocks are very persistent, as implied by the RIP model (Carroll (1992) and Attanasio et al. (1999)).

These plausible implications of the RIP model for the consumption-savings behavior—together with the fact that the fit of the RIP model to labor income data is not dramatically worse than that of the HIP model—have made the former the preferred income process in economic analysis. Perhaps surprisingly though, there exists no corresponding study of the consumption-savings behavior when the income process is the HIP model, so the implications of such a model for consumption data are largely unknown.[3] The goal of this paper is to fill this gap.

[2]It is then curious that the conclusion reached by MaCurdy's test seems to contradict this direct estimation evidence. Baker (1997) and Guvenen (2005) argue that the HIP model is in fact consistent with MaCurdy's results.

[3]Nevertheless, there has been some suggestion in the literature that the implications of the HIP model are not

To this end, we study the consumption-savings behavior in a life-cycle model with the HIP process. Because individuals are ex ante heterogeneous in the HIP model, a key question is how much individuals know about their own profile. Rather than imposing a certain amount of knowledge, we assume that individuals enter the labor market with some uncertainty about their income profile, and we infer the amount of this prior uncertainty from consumption data as described below. More specifically, we assume that individuals form a prior belief over β^i and α^i (the intercept of the income profile), which is updated in a Bayesian fashion with subsequent income realizations. We cast the optimal learning process as a Kalman filtering problem, which allows us to conveniently obtain recursive formulas for updating beliefs. In a related model, Wang (2004) obtains closed-form solutions for optimal consumption choice when individuals cannot distinguish between two separate persistent shocks. However, he abstracts from learning about profiles.

It is often the case with Bayesian learning that most of the uncertainty is resolved quickly, sometimes with a handful of observations. Instead, in the present framework learning turns out to be very gradual and its effects on consumption behavior extend throughout the life-cycle. The key feature of our model responsible for this result is the presence of learning about the *growth rate* of income. As we show in Section 2.4, Bayesian learning about a trend parameter (β^i) has features that are inherently different (and for our purposes more appealing) than the more standard learning about a level parameter (such as α^i).

We next compare the model to the U.S. consumption data. As a first step, we show that if individuals have no private information about their own profile (i.e., the prior variances of α^i and β^i equal the population variances), then the cross-sectional variance of log income increases by about 40 log-points over the life-cycle, significantly exceeding the roughly 25 log-points rise in the U.S. data. Thus the HIP model *is* capable of generating substantial rise in consumption inequality. One way to estimate the prior variance of β^i is then to choose it such that the model generates a 25 log-points increase in inequality. This procedure yields a prior variance of β^i equal to $0.6 \times \sigma_\beta^2$. The interpretation is that the remaining 40 percent of the variability in income growth rates in the population is forecastable by individuals by the time they enter the labor market.

Second, the empirical age-inequality profile of consumption has a *non*-concave shape. This fact has been emphasized by Deaton and Paxson (1994) and Storesletten et al. (2003) because the RIP model gives rise to a *concave* shape. The HIP model instead implies a non-concave age-inequality profile—that results from learning about β^i, as we show in Section 4.2—which also provides a fairly good fit to its empirical counterpart.

---

likely to be consistent with certain consumption facts. For example, Storesletten et al. (2001) state: "Should increasing income inequality be attributable to heterogeneity which is deterministic across households, many models of consumption choice predict that consumption inequality will not increase with age (p. 416)."

Third, Carroll and Summers (1991) also document that college graduates not only have steeper income profiles than high-school graduates but also have steeper consumption profiles. In the RIP model, the estimated persistence and innovation variance of income shocks are similar for different education groups (Hubbard et al. (1994), Carroll and Samwick (1997)), resulting in consumption profiles with similar slopes for both groups. On the other hand, when HIP is introduced, we find that the estimated dispersion of β^i among college graduates is more than twice that among high-school graduates (Table 1). This larger dispersion generates more precautionary savings and consequently a steeper consumption profile for college-graduates (unless these individuals are able to predict a much larger fraction of this dispersion compared to high-school graduates). These last two examples underscore the differences between the nature of labor income risk implied by the RIP model and the HIP model with Bayesian learning.

Finally, we show that assuming a HIP process *without* the uncertainty and learning about β^i yields a number of counterfactual implications for the consumption-savings behavior. Furthermore, if heterogeneity is introduced through a larger dispersion in α^i, (instead of the dispersion in β^i), more than 80 percent of this uncertainty is resolved in the first three years of an individual's life, leaving little role for further learning. Thus, learning about income growth is an essential element in this model.

As noted earlier, to our knowledge there is no previous work on the consumption-savings behavior when the income process is the HIP model. The closest work is Huggett, Ventura and Yaron (2004) who study a human capital model with ability heterogeneity and consider idiosyncratic shocks to the human capital accumulation function. They find differences in income growth rates (induced by ability heterogeneity) to be a key element in explaining the moments of the cross-sectional distribution of income and consumption. The difference is that in their framework, individuals know their own ability (and hence their average income growth rate), so there is no learning over time. In addition, they focus on the ability of the model to explain labor income data and study only a subset of consumption facts examined in the present paper. In a different context, Cunha et al. (2005) study students' schooling choice, in a complete markets setting, to infer the amount of earnings variability that is forecastable by the time students decide to go to college. They estimate that about 60 percent of variability in returns to schooling is forecastable. Navarro (2004) extends this analysis by introducing credit constraints and consumption choice, and finds that a significant fraction of schooling returns remain forecastable.

The rest of the paper is organized as follows. In the next section we introduce the RIP and HIP models and examine the properties of Bayesian learning about profiles in the latter model. In Section 3 we present a life-cycle model of consumption-savings with optimal learning. In Section 4, we present the quantitative results of the model. Section 5 discusses possible extensions and applications of the model and presents conclusions.

2 Bayesian Learning About Income Profiles

We first specify the RIP and HIP models and discuss the specific parameterizations we use. Second, because individuals are ex ante heterogeneous in the HIP model a key question is how much individuals know, and how they learn, about their individual-specific income profiles. Thus, we investigate the properties of optimal learning in this environment. The main result of this section is that learning about income growth rate (or a "trend variable" in general) has some interesting features not present when individuals learn about the level of income (or a "stationary variable" in general). This distinction is crucial and plays a central role in the determination of consumption and savings over the life-cycle.

2.1 Two stochastic processes for labor earnings

We first introduce the two income processes. The general process for *log earnings*, $\widetilde{y}_t^i$, of individual i who is t years old is given by

$$\widetilde{y}_t^i = g\left(\boldsymbol{\theta}^0, \mathbf{X}_t^i\right) + f\left(\boldsymbol{\theta}^i, \mathbf{X}_t^i\right) + z_t^i + \varepsilon_t^i \tag{2}$$

$$z_t^i = \rho z_{t-1}^i + \eta_t^i, \qquad z_0^i = 0,$$

where the functions g and f denote two separate "life-cycle" components of earnings. The first function captures the part of variation that is common to all individuals (hence the coefficient vector $\boldsymbol{\theta}^0$ is not individual-specific) and is assumed to be a quartic polynomial in experience, t.[4]

The second function, f, is the centerpiece of our analysis, and captures the component of life-cycle earnings that is individual- or group-specific. 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: $f\left(\boldsymbol{\theta}^i, \mathbf{X}_t^i\right) = \alpha^i + \beta^i t$, where the random vector $\boldsymbol{\theta}^i \equiv \left(\alpha^i, \beta^i\right)$ is distributed across individuals with zero mean, variances of σ_α^2 and σ_β^2, and covariance of $\sigma_{\alpha\beta}$.[5]

The stochastic component of income is modeled as an AR(1) process plus a purely transitory shock. This specification is fairly common in the literature and, despite its parsimonious structure,

[4] While it is also possible to include an education dummy into g, we do not pursue this strategy in the baseline specification. Later in the paper, we will allow for a separate income process for each education group to fully control for the effect of education on the life-cycle profiles as well as its effect on the persistence and variance of income shocks.

[5] The zero-mean assumption is a normalization since g already includes an intercept and a linear term. Moreover, 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. In addition, each term introduced into this component will appear as an additional state variable in the dynamic programming problem we solve later. In the baseline case, that problem already has five continuous state variables and certain non-standard features described in the computational appendix, so we prefer to avoid any further complexity. Lillard and Reville (1999) on the other hand, provide some evidence suggesting that the quadratic term may be important, so this seems to be an extension worth considering in future work.

it appears to provide a good description of income dynamics in the data (Topel (1990), Hubbard et al. (1994), Moffitt and Gottschalk (1995), Storesletten et al. (2004)).[6] The innovations η_t^i and ε_t^i 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.

The RIP and HIP models are distinguished by their assumptions about f. The HIP model refers to the general (unrestricted) process given in equation (2). The RIP model, on the other hand, refers to the same process estimated with the restriction $\beta^i \equiv 0$ imposed.

To calibrate the model that we present in the next section, we use the parameter values displayed in table 1 taken from Guvenen (2005). The first two rows display the estimates for the whole population from the RIP and HIP models respectively. The HIP model implies a significantly lower persistence for the AR(1) process (0.82 compared to 0.988) and a statistically (and as shown below, quantitatively) significant heterogeneity in income growth rates ($\sigma_\beta^2 = 0.00038$ with a t-value of 4.9). For comparison, table 5 in Appendix A presents the estimates from the HIP model obtained in the previous literature. Overall, the parameter values we use are consistent with this earlier work with one exception: the variance of the fixed effect, σ_α^2, is much smaller in our estimates (0.02 compared to 0.14 in Baker (1997) and 0.29 in Haider (2001). In Section 2.4 we show that using a value of 0.2 would have no appreciable effect on our results. Finally, the subsequent rows of table 1 display the parameter estimates for college-educated and high school-educated individuals that is used in Section 4.3. To our knowledge, the parameter estimates of the HIP model for each education group is only available in Guvenen (2005).

2.2 Quantifying the heterogeneity in income profiles

While the point estimate of σ_β^2 of 0.00038 may appear small, this value implies substantial heterogeneity in income growth rates. To see this, we first define the income residual, $y_t^i \equiv \widetilde{y}_t^i - g$, and use the following equation (derived from eq. (2)) to decompose the within-cohort income inequality into its components:

$$var_i(y_t^i) = \left(\sigma_\alpha^2 + \sigma_\varepsilon^2\right) + \left(\frac{1-\rho^{2t+1}}{1-\rho^2}\sigma_\eta^2\right) + \left(2\sigma_{\alpha\beta}t + \sigma_\beta^2 t^2\right).$$

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 accumulated effect of the autoregressive shock. Finally, the last parenthesis contains two terms that vary with

[6] Alternatively, it is possible to use an unrestricted ARMA (1,1) or (1,2) process (MaCurdy (1982), Abowd and Card (1989), Moffitt and Gottschalk (1995)). Although this specification provides more flexibility, it also introduces additional parameters that appear as state variables in dynamic decision problems (as the one we study in Section 3) expanding the state space. Consequently, economic models that use individual income processes as inputs typically opt for more parsimonious specifications similar to the one used here.

age, which are due to profile heterogeneity: a decreasing linear term in t (since $\sigma_{\alpha\beta} < 0$), and more importantly, a quadratic term in t. It is easy to see that even when σ_β^2 is very small, the effect of profile heterogeneity on income inequality will grow rapidly with t^2, as the cohort gets older. Table 2 illustrates this point by displaying the value of terms in each parenthesis over the life-cycle. As can be seen in column 4, the contribution of profile heterogeneity to income inequality is very small early in the life-cycle. In fact, up to age 47 more than half of the cross-sectional variance of income is generated by the fixed effect, and the transitory and persistent shocks. The effect of profile heterogeneity continues to rise however, and accounts for almost 80 percent of inequality at retirement age.

2.3 The Kalman Filtering Problem

The key feature of the HIP model is that individuals are ex ante different in their income profiles, which—as the analysis above illustrates—accounts for a large fraction of the rise in within-cohort income inequality. Hence, to embed the estimated income process into a life-cycle model, we need to be specific about what the individual knows about (α^i, β^i). A plausible scenario is one in which an individual enters the labor market with some prior belief about his income growth prospects. This prior could incorporate some relevant information unavailable to the econometrician as we discuss below. Over time, 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.

In order to formally define the learning problem we need to specify which components of income are observable. If the stochastic component, $z_t + \varepsilon_t$, were observable in addition to y_t^i, individual income profiles (α^i, β^i) would be revealed in just two periods, leaving no role for further learning. 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. Thus as a benchmark case, we assume that individuals only observe total income, y_t^i, and not its components separately.

It is convenient to express the learning process as a Kalman filtering problem using the state-space representation. In this framework, the "state equation" describes the evolution of the vector of state variables that is unobserved by the decision maker:[7]

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

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

Even though the parameters of the income profile have no dynamics, including them into the state vector yields recursive updating formulas for beliefs using the Kalman filter. A second (observation) equation expresses the observable variable(s) in the model—in this case, log income—as a linear function of the underlying hidden state and a transitory shock:

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

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.[8] 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:[9]

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

where we use the short-hand notation $\sigma_{\cdot,t}^2$ to denote $\sigma_{\cdot,t+1|t}^2$. After observing $\left(y_t^i, y_{t-1}^i, ...y_1^i\right)$, 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:

$$\begin{aligned} \widehat{\mathbf{S}}_{t|t}^i &= \widehat{\mathbf{S}}_{t|t-1}^i + \mathbf{P}_{t|t-1}\mathbf{H}_t \left[\mathbf{H}_t'\mathbf{P}_{t|t-1}\mathbf{H}_t + R\right]^{-1} \left(y_t^i - \mathbf{H}_t'\widehat{\mathbf{S}}_{t|t-1}^i\right), \\ \widehat{\mathbf{S}}_{t+1|t}^i &= \mathbf{F}\widehat{\mathbf{S}}_{t|t}^i, \end{aligned} \tag{3}$$

$$\begin{aligned} \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{H}_t \left[\mathbf{H}_t'\mathbf{P}_{t|t-1}\mathbf{H}_t + R\right]^{-1} \mathbf{H}_t'\mathbf{P}_{t|t-1}, \\ \mathbf{P}_{t+1|t} &= \mathbf{F}\mathbf{P}_{t|t}\mathbf{F}' + \mathbf{Q}. \end{aligned} \tag{4}$$

Notice that 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 (4) that the posterior variances of α^i and β^i are monotonically decreasing over time, so with every

[8] The normality assumption is not necessary for the estimation of the parameters of the stochastic process (2), and is not made in Guvenen (2005) to obtain the parameters in table 1.

[9] The notation $\widehat{X}_{h_2|h_1}$ denotes the forecast of (alternatively, belief about) X_{h_2} given the information available at time h_1 if $h_2 > h_1$ (if $h_2 = h_1$).

new observation beliefs become more concentrated around the true values. (This is not necessarily true for $\sigma^2_{z,t}$ which may be non-monotonic depending on the parameterization.) Finally, log income has a Normal distribution conditional on an individual's beliefs:

$$y^i_{t+1}|\widehat{\mathbf{S}}^i_{t|t} \sim N\left(\mathbf{H}'_{t+1}\widehat{\mathbf{S}}^i_{t+1|t}, \mathbf{H}'_{t+1}\mathbf{P}_{t+1|t}\mathbf{H}_{t+1} + R\right). \tag{5}$$

2.4 The speed of resolution of profile uncertainty

The results presented in Section 2.2 suggest that a substantial fraction of income differences over the life-cycle is due to HIP. Consequently, the *initial* income risk perceived by an individual upon entering the labor market can be substantial if the individual is sufficiently uncertain about his income profile. However, since individuals learn their profile over time, the contribution of profile uncertainty to perceived income risk *later* in the life-cycle depends on the speed of learning. It is often the case with Bayesian learning that a large fraction of prior uncertainty is resolved quickly, so it is essential to investigate this issue in the present framework.

As we quantify below, learning is very gradual in our model and its effects extend throughout the life-cycle for two reasons. The first and main reason is that early in life the contribution of the $\beta^i t$ term to income is very small—most of the variation in income is due to shocks as can be seen in Table 2—so income observations are not very informative about the growth rate of income, slowing down learning. Second, later in life, when observations become potentially more informative, the moderate persistence of shocks makes it difficult to disentangle them from the trend component, again slowing down learning. In the rest of this section we make these points more rigorous.

We begin by defining a convenient measure of income uncertainty, the forecast variance—the mean squared error (MSE) of the forecast—of future income:

$$MSE_{t+s|t} \equiv E_t\left(y_{t+s} - \widehat{y}_{t+s|t}\right)^2 = \mathbf{H}'_{t+s}\mathbf{P}_{t+s|t}\mathbf{H}_{t+s} + R, \tag{6}$$

$$\text{where} \quad \mathbf{P}_{t+s|t} = \mathbf{F}^s\mathbf{P}_{t|t}\mathbf{F}'^s + \sum_{i=0}^{s-1}\mathbf{F}^i\mathbf{Q}\mathbf{F}'^i. \tag{7}$$

If individuals know their profile with certainty (i.e., $\sigma^2_{\alpha,t} = \sigma^2_{\beta,t} = 0$), the forecast variance in equation (6) reduces to $MSE^{idio}_{t+s|t} = E_t\left(z_{t+s} - \widehat{z}_{t+s|t}\right)^2 + \sigma^2_\varepsilon$, where the superscript *idio* indicates that the only source of risk in this case is idiosyncratic shocks. Notice that an income process with RIP is a special case of this, so the same expression characterizes the forecast variance for such processes. In the more general case where individuals are uncertain about their profile ($\sigma^2_{\alpha,t}, \sigma^2_{\beta,t} > 0$), the forecast variance can be written as:

$$MSE^{total}_{t+s|t} = MSE^{idio}_{t+s|t} + \left\{\left[\sigma^2_{\alpha,t} + 2\sigma_{\alpha\beta,t}(t+s) + \sigma^2_{\beta,t}(t+s)^2\right] + \kappa_{t+s|t}\right\}, \tag{8}$$

which is again obtained using equation (6). The first term captures the risk resulting from idiosyncratic shocks as before. The remaining terms in parenthesis (call it $MSE^{net}_{t+s|t}$) is the *net contribution of profile uncertainty* to income risk at different horizons (given by s) as perceived by an individual at age t. For a given t, the terms in the square bracket imply that the forecast variance (due to profile heterogeneity) is an increasing quadratic function of horizon $(t+s)$. In addition, although 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 an increasing function of s for fixed t, but does not materially affect the shape of this profile.

In the left panel of figure 1 we plot $MSE^{net}_{t+s|t}$, $s = 1, 2, ...$, for an individual at ages $t = 25, 35, 45$, and 55, who faces the HIP process estimated on row 2 of Table 1.[10] The top curve ($t = 25$) shows that the future income risk perceived by this individual upon entering the labor market is substantial, as can be expected from the fact that HIP accounts for a large fraction of income inequality and the individual does not initially know his true profile. As the individual gets older, the successive MSE curves shift downward reflecting the resolution of profile uncertainty. The main point to notice in this graph is that the resolution of uncertainty is slow: by the time the individual is 35 years old (the second curve from the left) only 26 percent of income risk at retirement will have been resolved. At age 45, the forecast variance of income at retirement is still about 0.22. For comparison, at the same age, the forecast variance at retirement that is due to idiosyncratic shocks ($MSE^{idio}_{65|45}$) is only 0.045.

The main reason for the slow learning is that individuals learn about a *slope* parameter, β^i, whose contribution to income is small when individuals are young, but grows monotonically with age. Figure 2 illustrates the implications of this feature for the speed of learning. Specifically, the vertical axes plot $(\log(1/\sigma^2_{x,t+1|t}) - \log(1/\sigma^2_{x,t|t-1}))$, for $x = \alpha^i$ (left panel) and β^i (right panel), which can be interpreted as the *percentage* improvement in precision—or equivalently, the percentage reduction in the posterior variance—at each age. In the left panel the resolution of uncertainty about α^i follows the familiar pattern: most of the learning takes place early on, and after the first five or so years each subsequent observation brings little fresh information about the intercept term. In contrast, in the right panel, the information provided about β^i by each additional observation *increases* over time, up to about age 50. Using the terminology of signal extraction problems, the signal-to-noise ratio increases—resulting in faster learning—as the individual gets older. In fact this can be seen in figure 1, where the MSE curves are shifting to the right faster as the individual gets older.

It is useful to contrast the resolution of uncertainty above to the hypothetical case where the

[10] To calculate the MSE we need to specify the prior covariance matrix, $\mathbf{P}_{1|0}$. We discuss the specification of the priors more fully below. As a simple benchmark, here we assume that the individual does not have more information than the econometrician so that the prior variance of each variable is equal to its population value (that is, $\sigma^2_{\alpha,0} = \sigma^2_{\alpha}$, etc.)

main source of uncertainty (and hence learning) is about the *level* of income, α^i. This comparison is also helpful because our baseline estimate of σ_α^2 is around 0.02 whereas the corresponding point estimate is 0.14 in Baker (1997) and 0.29 in Haider (2001) (see table 5). Figure 3 plots the change in precision of beliefs about α^i when σ_α^2 (and correspondingly the prior variance) is set to 0.20. The two lines plot the precision when the dispersion in β^i is fixed at its baseline value ('-^'), and alternatively, when it is set to zero ('- -') (The two lines are almost indistinguishable in the first four years). In both cases, the log precision improves by 130 log points with the first observation, implying that the posterior variance of α^i falls (by $\approx e^{1.3}$) from 0.20 to 0.054 after the first year, and to below 0.04 after the third year. The reason for this fast learning is clear: since $\beta^i t$ is very small early in life, and the stochastic shocks have much smaller variances and lower persistence than α^i, the latter stands out (i.e., the signal-to-noise ratio is high) and is detected easily. Hence, even when there is significantly more initial uncertainty about the level of income, it has little effect on the behavior of individuals after the first few years, unlike the effect of learning about the growth rate of income.

A second reason for the slow learning is the *moderate* persistence of income shocks. We illustrate this point in the right panel of figure 1. The bottom curve plots the net forecast variance of income at retirement by an individual who is 35 years old ($MSE^{net}_{65|35}$) as a function of the persistence of z_t, normalized by its value at $\rho = 0$. The two curves above that are constructed similarly for $t = 45$ and 55 respectively. When constructing these graphs, we adjust the innovation variance of z_t as we vary ρ to keep the unconditional variance of the AR(1) process unchanged.

One conclusion that is clear from this graph is that the speed of resolution is not a monotone function of persistence: as ρ increases from zero up to about 0.85, the resolution of uncertainty slows down (reflected in a larger forecast variance at retirement), but then speeds up again as persistence increases further towards a unit root. In particular, learning is faster when income shocks follow a random walk than for any other value of ρ.[11] Interestingly, the values of ρ where learning is slowest coincides with the empirical estimates of persistence reported in table 1 (although the figure also makes clear that the resolution of uncertainty is not dramatically different for values of ρ roughly between 0.7 and 0.9).

The second feature apparent in the right panel of figure 1 is that the impact of persistence on the speed of learning increases with age. For example, at age 35, increasing the persistence from zero to 0.8 results in a 30 percent rise in $MSE^{net}_{65|35}$. At age 55, the same experiment raises the forecast variance by 180 percent. Thus, the relatively high persistence of income shocks in the data is important for the slow resolution of uncertainty especially later in the life-cycle.

Before concluding this section, it should be noted that slow learning is also important for another

[11] Loosely speaking this is because when income shocks are random walk, income *growth* becomes very informative about β^i since $\Delta y_t^i = \beta^i + (\rho - 1)\, z_{t-1} + \eta_t + \Delta\varepsilon_t$ reduces to $\beta^i + \eta_t + \Delta\varepsilon_t$ in this case.

reason: In the next section we infer the amount of prior uncertainty from the consumption-savings behavior of individuals over the life-cycle. But if learning were quick, individuals' behavior *later* in life would contain little information about their prior uncertainty, making this setup unsuitable for this exercise. In other words, life-cycle behavior is informative about prior uncertainty to the extent that it is not resolved very quickly.

3 A Life-Cycle Model of Consumption and Savings

We now study the consumption-savings decision of an individual in an environment with HIP and Bayesian learning as described in the previous section. We 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 is given by equation (2). Once retired the individual receives a pension equal to a fraction, Φ, of his income at age T. Although this specification is admittedly much simpler than the Social Security system in the U.S., it has the advantage of abstracting from the significant risk-sharing inherent in that system, and consequently from its effects on the consumption-savings decision which may confound the main focus of our analysis. Moreover, introducing a more realistic pension scheme where the payments depend on the average earnings over the life-cycle would add an extra state variable to the dynamic program. Thus, to avoid further complication we choose this simpler specification. In the robustness analysis, we examine the effect of a redistributive pension system on our results. Finally, there is a risk-free bond that sells at price P^b (with a corresponding interest rate $r^f \equiv 1/P^b - 1$). 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, $\widehat{\mathbf{S}}_{t|t-1}$.[12] 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

$$\begin{aligned} V_t^i(\omega_t^i, y_t^i, \widehat{\mathbf{S}}_{t|t-1}^i) &= \max_{c_t^i, \omega_{t+1}^i} \left\{ U(c_t^i) + \delta E\left[V_{t+1}^i(\omega_{t+1}^i, y_{t+1}^i, \widehat{\mathbf{S}}_{t+1|t}^i) | \widehat{\mathbf{S}}_{t|t-1}^i \right] \right\} \\ & \quad\; s.t \\ c_t^i + P^b \omega_{t+1}^i &= \omega_t^i + y_t^i \qquad (9) \\ \omega_{t+1}^i &\geq \underline{W}_{t+1} \qquad (10) \\ & \quad\; eq.\ (3,4) \end{aligned}$$

[12] Although given the last two variables one can obtain both $\widehat{\mathbf{S}}_{t|t}$ and $\widehat{\mathbf{S}}_{t+1|t}$ using equation (3) (which means that the individual knows the latter two vectors at the time of decision) our current choice turns out to be more convenient for computational reasons.

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 (3, 4). 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^i_{t+1} given by equation (5). After retirement, labor income is constant and there is no other source of uncertainty or learning, so the problem simplifies significantly:

$$\begin{aligned} V^i_t(\omega^i_t, y^i) &= \max_{c^i_t, \omega^i_{t+1}} \left[U(c^i_t) + \delta V^i_{t+1}(\omega^i_{t+1}, y^i)\right] \\ s.t \qquad y^i &= \Phi y^i_T, \text{ and } eqs.\,(9, 10) \end{aligned}$$

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

3.1 Baseline parameterization

There is no analytical solution to the dynamic optimization problem stated in the previous section, so we resort to numerical methods. The numerical solution is complicated by the fact that there are five continuous state variables and four of them (excluding ω^i_t) depend on each other as a result of learning. In particular, this inter-dependence makes the solution of the value function 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.

A model period is one year of calendar time. Individuals enter the labor market (are born) at age 25, retire at 65 and die at age 95. The period utility function is assumed to take the CRRA form with a relative risk aversion coefficient, ϕ, equal to 2. The bond price, P^b is set equal to 0.96 implying an annual interest rate of 4.16 percent. We set the pension replacement rate, Φ, equal to 0.34 in the baseline case. This value is chosen so that the pension income of the average individual in the model is equal to 50 percent of the average of his income over the life-cycle as in the current U.S. Social Security system. Finally, we set the time preference rate, δ, to match the average wealth-to-income ratio in the U.S. data. Budria-Rodriguez et al. (2002) calculate this ratio both from the Survey of Consumer Finances data set and from the National Income and Products Account data, and obtain values between 4.14 and 5.26. However, it is not immediately clear how to treat housing in this calculation, which is included in their calculation but not explicitly modeled in our framework. With this in mind, we target a wealth-to-income ratio of 3.5, somewhat lower than these reported values. Moreover, because the amount of precautionary savings depends on the amount of uncertainty, in the next section when we make comparisons across versions of the model, we adjust δ to keep the wealth-to-income ratio on this target.

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 the PSID sample of Guvenen (2005): 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.

Determining the priors.—Empirical evidence is not particularly helpful for setting $\mathbf{P}_{1|0}$. The difficulty is that the econometrician is only able to measure the population distribution of $\left(\alpha^i, \beta^i\right)$ conditional on a few observable characteristics. But it is conceivable that each individual could have some information, unavailable to the econometrician that can provide a better prediction of their income profile. Thus, rather than imposing a certain amount of prior knowledge on the individual, we infer it from the observable actions over the life-cycle.

We begin by describing how an individual's prior belief about β^i is determined. Suppose that the distribution of income growth rates in the population is generated as $\beta^i = \beta^i_k + \beta^i_u$, where β^i_k and β^i_u are two random variables, independent of each other, with zero mean and variances of $\sigma^2_{\beta_k}$ and $\sigma^2_{\beta_u}$. Clearly then, $\sigma^2_\beta = \sigma^2_{\beta_k} + \sigma^2_{\beta_u}$. The key assumption is that individual i observes the realization of β^i_k, but not of β^i_u (hence the subscripts indicate *k*nown and *u*nknown, respectively). Under this assumption, the prior mean of individual i is $\widehat{\beta}^i_{1|0} = \beta^i_k$, and the prior variance is $\sigma^2_{\beta,0} = \sigma^2_{\beta_u} = (1-\lambda)\,\sigma^2_\beta$, where we define $\lambda = 1 - \sigma^2_{\beta_u}/\sigma^2_\beta$, as the fraction of variance known by individuals.

Two polar cases deserve special attention. If $\lambda = 0$, individuals do not have any private prior information about their income growth rate (i.e., $\sigma^2_{\beta,0} = \sigma^2_\beta$). This case provides a useful benchmark (or an upper bound) to gauge how much mileage one can get by allowing uncertainty about individual income profiles. On the other hand if $\lambda = 1$, each individual observes β^i completely and faces no prior uncertainty about its value. This case provides a useful comparison to illustrate how profile uncertainty and learning alter individuals' consumption-savings decision. Finally, as noted earlier, the dispersion of α^i is not large according to our parameterization and does not materially affect the results of the model even when it is larger. For simplicity we assume that individuals have no private prior information about their intercept so that $\sigma^2_{\alpha,0} = \sigma^2_\alpha$. Then the prior covariance matrix is:

$$\mathbf{P}_{1|0} = \begin{bmatrix} \sigma^2_\alpha & \sqrt{1-\lambda}\sigma_{\alpha\beta} & 0.0 \\ \sqrt{1-\lambda}\sigma_{\alpha\beta} & (1-\lambda)\,\sigma^2_\beta & 0.0 \\ 0.0 & 0.0 & \sigma^2_\eta \end{bmatrix},$$

We refer to this general framework as the "profile heterogeneity with uncertainty" (PHU(λ)) model, and to the special case with $\lambda = 0$ as the "profile heterogeneity with certainty" (PHC) model. Notice of course that in both cases the income process is the HIP model.

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 face a different natural limit, unlike in a standard life-cycle model with ex-ante identical individuals. However, in this case 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 probably without providing much additional insight. Instead, we allow all individuals to borrow up to a common borrowing limit determined as the natural limit of an individual based on *public* prior information (that is, $\sigma^2_{\beta,0} = \sigma^2_{\beta}$). In other words, this is the natural limit that credit institutions would enforce on individuals when only time-0 public information is 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 of income at 2.5 standard deviations to provide a proper lower bound. As we discuss further below in our baseline specification this constraint is almost never binding.

To simulate the model we draw 1000 (α, β) combinations (one for each type of individual) from a bivariate Normal distribution whose covariance matrix is taken from row 2 of Table 1 and simulate 100 paths for each type. The reported statistics are averages over these simulated data.

4 Quantitative Results

In this section, we investigate if the consumption behavior implied by the HIP model combined with Bayesian learning about income profiles is consistent with empirical facts, and especially with certain findings that have commonly been interpreted as evidence supporting the RIP model (Deaton and Paxson (1994), Carroll and Summers (1991)).

4.1 The age-inequality profile of consumption

Deaton and Paxson (1994) document the striking rise in within-cohort consumption inequality (along with income inequality) over time. In particular, the cross-sectional variance of log consumption (per adult equivalent) increases by about 25 log points in the U.S. data—roughly corresponding to the doubling of inequality—over a cohort's life-cycle. For completeness, we replicate their finding as closely as possible.[13] The broken line in figure 4 displays the resulting age-inequality profile, which

[13] We use data on consumption expenditures from the Consumer Expenditure Survey (CE), obtained from Krueger and Perri (2004). We choose our consumption measure and sample period following Deaton and Paxson (1994) to make comparison easier. Thus, consumption refers to household non-durable expenditures during the last quarter. The sample period is 1980-90, but unlike these authors who concentrate on urban households, we include all households into our sample, which requires us to exclude the 1982-83 period since data on rural households is not available during

is very similar to the one presented by Deaton and Paxson.

While the mere existence of fanning-out in the consumption distribution is not surprising—as it is implied for example by the permanent income theory—the large magnitude of the increase is. Deaton and Paxson discuss several potential explanations and find the existence of persistent (uninsurable) idiosyncratic shocks to be the most promising candidate. Recently, Storesletten et al. (2003) have shown that a life-cycle model can quantitatively match the rise in inequality observed in the data *if* income shocks are extremely persistent.

We now examine the evolution of within-cohort consumption inequality in the PHU(λ) model. To provide a benchmark, we begin with the case where individuals have no private prior information about β^i (that is, $\lambda = 0$). The (top) solid line in Figure 4 plots the age-inequality profile in this case, which shows a substantial rise in consumption inequality over the life-cycle—roughly 40 log points, compared to 25 log points in the U.S. data.[14] Thus, if anything, the model generates *too much* rise in inequality unless individuals have more prior information than assumed in this case.

This observation suggests that one way to measure λ is to choose it to generate the same increase in consumption inequality over the life-cycle observed in the data.[15] This procedure yields $\lambda^* = 0.46$. The interpretation of this estimate is that 46 percent of the variability in income growth rates is forecastable by individuals at the time they enter the labor market. Moreover, although λ is chosen to match the total rise in inequality, the overall *shape* of the resulting age-inequality profile provides a nice fit to its empirical counterpart. We return to this point in the next section.

There are several issues related to the interpretation of the estimate of λ. First, recall that in equation (2) we did not include an education dummy into the common life-cycle profile g, so any variation in income growth rates between education groups is also captured in σ^2_β. As a result λ also contains any forecastability in β^i that is due to differences in education level. It is possible to quantify the amount of forecastability due to education, and the amount due to other (possibly unobservable) variables known to the individual. To this end, we first solve the consumption-savings problem for each education group separately under the assumption that individuals know the average value of β^i and its dispersion for their education group, but are *not* able to predict their own β^i

this time. We use Census equivalence scale to convert household consumption into per-adult-equivalent units to make comparable to income data from PSID which is used to estimated the income processes in table 1. See Krueger and Perri (2004) for further details of sample selection and variable construction.

The differences in our data construction from Deaton and Paxson (1994) do not seem to make a noticeable difference for the age-inequality profile we obtain, which is very similar to theirs. To obtain the graph in figure 4, we follow Deaton and Paxson and 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 dummies are scaled so that the average inequality matches that in the sample.

[14] Only a small fraction of this fanning-out is directly attributable to idiosyncratic shocks: if we eliminate profile heterogeneity (and consequently learning) from this model, the rise in inequality would only be 8 log points.

[15] When we calibrate the parameters of different education groups separately as we do below, we keep δ identical across these groups and adjust them by the same amount to keep the aggregate wealth-to-income ratio unchanged.

beyond this information. In addition, individuals in each education group are now assumed to face the idiosyncratic shock process corresponding to their group as reported on rows 4 and 6 of Table 1.

The dashed line in figure 4 plots the result. Now inequality rises by 32 log points, compared to 40 log points when individuals do not condition on education. To translate this difference into the fraction of forecastability due to education information, we choose λ in the benchmark model above to generate an increase in inequality of 32 log points, which yields $\lambda^* = 0.27$. The difference between the two estimates of λ $(0.46 - 0.27 = 0.19)$ measures the fraction of forecastability due to information other than education available to the individual.

Redistributive Social Security.—The U.S. retirement pension system features significant redistribution, thereby providing risk-sharing within each cohort. The extent of risk-sharing in turn is critical for the rise in consumption inequality over the life-cycle. For example, with complete risk-sharing in the baseline model the age-inequality profile would be flat—consumption inequality would be constant over the life-cycle—regardless of the amount of prior uncertainty about income profiles. Therefore our estimate of λ partly depends on the assumed pension system. To examine the sensitivity of the previous estimate, we next introduce a redistributive pension system which captures the salient features of the U.S. Social Security system as described in Storesletten et al. (2003). Specifically, the retirement replacement rate is a concave function of an individual's income at age T given by:

$$\Phi\left(Y_T\right) = 0.715 \times \begin{cases} 0.9Y_T & \text{for } Y_T < 0.3\overline{Y}_T \\ 0.27 + 0.32Y_T & \text{for } Y_T \in (0.3\overline{Y}_T, 2\overline{Y}_T] \\ 0.81 + 0.15Y_T & \text{for } Y_T \in (2\overline{Y}_T, 4.1\overline{Y}_T] \\ 1.1 & \text{for } Y_T > 4.1\overline{Y}_T \end{cases}$$

where $\overline{Y}_T$ is the average income at age T.[16]

With this modification to the pension system, consumption inequality would rise by 22 log points over the life-cycle, if λ was kept at its baseline estimate of 0.46 (and δ was re-set to 0.957 to keep the wealth-to-income ratio unchanged). As predicted, the concave pension function reduces the differences in life-time income compared to the baseline model thereby reducing consumption inequality along with it. To match the increase of 0.25 log points as in the data, the value of λ must be 0.37. The solid line in Figure 5 plots the age-inequality profile along with the one from the baseline model for comparison. With this re-calibrated value of λ, the shape of the profile changes

[16] There is one difference between this specification and the one in Storesletten et al. (2002): Φ here is a function of Y_T instead of the average income over an individual's life-cycle (which would require us to track one more state variable). However, because income shocks are not very persistent in our model, Y_T is highly correlated with an individual's average income (correlation: 0.89), so the difference may not be crucial. Moreover, because Y_T is 40 percent higher than average income over the life-cycle, we need to multiply our pension schedule by $1/1.40 \approx 0.715$ to match the average level of benefits in their specification.

little and the only noticeable change happens after age 55.

We conclude from these results that a model with HIP generates substantial rise in consumption inequality—in fact, more than what is observed in the U.S. data—which suggests that some part of the heterogeneity in income growth rates is known by individuals by the time they enter the labor market. As our benchmark figure, we take the estimate from the model with Social Security, $\lambda^* = 0.37$.

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

A second feature of the age-inequality profile emphasized by Deaton and Paxson is its *non*-concave shape. Examining consumption data from three countries—U.S., U.K., and Taiwan—these authors find that the age-inequality profile increases approximately linearly in the former and is convex in the latter two countries. The same pattern also holds true in the PHU(λ) 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. Storesletten et al. (2003) later study a more general 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. For completeness, the dashed line in figure 5 plots the age-inequality profile from the baseline model studied by these authors where this concavity can be seen.

The effect of Bayesian learning on the shape of the age-inequality profile mainly depends on whether learning is about the intercept or the slope of the income profile. The main intuition can be conveyed in the certainty-equivalent version of the permanent income model (i.e., assuming quadratic utility and $\delta\left(1+r^f\right)=1$, and no retirement). In this case optimal consumption choice implies:

$$\Delta c_t = \frac{1}{\varphi_t}\left[(1-\gamma)\sum_{s=0}^{T-t}\gamma^s\left(E_t - E_{t-1}\right)Y_{t+s}\right] \tag{11}$$

where $Y_{t+s} \equiv \exp\left(y_{t+s}\right)$ is the level of income, $\gamma = 1/\left(1+r^f\right)$, and $\varphi_t = \left(1-\gamma^{T-t+1}\right)$ is the annuitization factor. To simplify the problem further, assume that income (instead of *log* income) is a linear function of experience with *i.i.d.* innovations: $Y_t^i = \alpha^i + \beta^i t + \varepsilon_t^i$.[17]

[17] These assumptions are rather innocuous in this context. First, the exponential function is increasing and convex, so the Log-normal specification for income in the baseline model will only reinforce the mechanism described here. In fact, we can obtain a closed-form solution for the consumption decision for the model described in this section (including when income is specified as Log-normal). This explicit solution allows one to easily verify this assertion as well as showing that the convexity result does not depend on the existence of borrowing constraints. These results are available upon request. Second, even though income shocks are not *i.i.d* in our model, their persistence is small enough that they do not create strong enough concavity to overturn this conclusion.

First, consider the case where individuals learn about their intercept α^i, but know their β^i exactly. It can easily be shown that (11) reduces to $\Delta c_t = \widehat{\alpha}_{t|t} - \widehat{\alpha}_{t-1|t-1}$. Because the right hand side of this expression is shrinking with age due to learning, the age-inequality profile will be unambiguously concave.

Now consider the opposite: individuals learn about β^i, but know α^i. When an individual updates his beliefs in period t, the revision in expected future income is: $(E_t - E_{t-1})\, Y_{t+s} = (\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1})\,(t+s)$. Substituting this expression into (11) and after performing some algebra one can show that:

$$\Delta c_t = \left[\left(\frac{\gamma}{1-\gamma}\right) + \frac{t - (T+1)\,\gamma^{T-t+1}}{1-\gamma^{T-t+1}}\right]\left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right). \tag{12}$$

For a range of plausible values for r^f and T, the term in the square bracket is an approximately linear (and slightly convex) *increasing* function of t. Moreover, recall—from the graph in the right panel of figure 2—that the speed of learning about β^i increases over time so the *absolute value* of $(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1})$ is getting larger on average up to about age 50.[18] As a result, consumption changes will be larger in absolute value as a cohort gets older implying a convex shape for the age-inequality profile.

Although in the PHU(λ) model individuals learn about both the intercept and the slope of their income profile, learning about α^i happens very quickly (even when the prior variance is much larger than what is assumed in the baseline calibration), and hence has no significant effect on the shape. Instead, the shape is mainly determined by learning about β^i, which gives it the non-concave form.

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

A second well-documented empirical finding is that consumption tracks income over the life-cycle: it first rises and then falls with income (Carroll and Summers (1991)). It is possible to generate this behavior in a life-cycle model by assuming idiosyncratic shocks with high persistence, and a utility function that induces precautionary savings behavior (such as CRRA.) In this case, individuals reduce their consumption early in life to build a buffer stock wealth for self-insurance. As they get older, persistent shocks have fewer periods left to affect income, effectively resulting in less uncertainty. In response, individuals reduce their savings rate, allowing consumption to rise along with income, generating the empirical co-movement (Carroll (1992), Attanasio et al. (1999).

However, another finding documented by Carroll and Summers poses a challenge to this basic story. These authors found that consumption also tracks income within education groups: that is, college-educated individuals not only have steeper income profiles, but also have steeper consumption profiles than high-school educated individuals. For example, Krueger and Fernandez-Villaverde

[18] Although, more precisely figure 2 shows that the change in the *logarithm* of the belief about β^i is increasing.over time, a similar plot also holds true for the *level* of beliefs.

(2004) show that over the life-cycle consumption (in per-adult equivalent units) rises by about 50 percent for the former group and by about 10 percent for the latter. For a model based on precautionary savings alone to explain this observation, it would require the former group to face income shocks that are either more persistent or more volatile than the latter group,[19] neither of which we seem to find in the data. For example, in rows 3 and 5 of table 1 which display the estimates from the RIP model, there is little difference between the idiosyncratic shock process faced by each group. To examine the consumption profiles that would be implied by these income processes, we solve the life-cycle model described in Section 3 for each education group separately. For each group we use the income process with RIP from rows 3 and 5 of Table 1. Figure 6 displays the results. The left panel plots 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 noticeably steeper than that of the high school-educated group: it rises by 38 percent over the life-cycle for the former group compared to 36 percent for the latter.

Next we solve the PHU(λ) model for each education group. As can be seen in rows 4 and 6 of table 1, while the estimates of the idiosyncratic shock processes for each group remain similar to each other when HIP is introduced, a major difference arises in the dispersion of β^i: college graduates face a much wider dispersion of income growth rates ($\sigma_\beta^2 = 0.00049$) than lower educated individuals ($\sigma_\beta^2 = 0.00020$). To translate these numbers into the amount of prior uncertainty about β^i, we assume that λ is equal to 0.19 for both groups, which is the value we obtained in Section 4.1 *after* conditioning on education information.[20] Now the average consumption profile (right panel of Figure 7) rises twice as much (by 39 percent) for the college-educated compared to the high school-educated (by 19 percent). More prior uncertainty about income profiles generates more precautionary savings for the former group resulting in steeper rise in consumption.

Note finally that consumption would also track income (even without uncertainty about income profiles) if there were frequently binding borrowing constraints. In the presence of HIP, however, constrained individuals would typically be those with high income growth rates. Indeed, in the PHC model ($\lambda = 1$), the average consumption of constrained individuals is higher throughout the life-cycle and is almost double that of unconstrained ones at retirement. This comparison also shows that profile uncertainty should be an integral part of a model with HIP, which otherwise (with a very high λ) yields counterfactual implications.

[19] Clearly this is because without income shocks both groups should have the same slope of the income profiles unless they differ systematically in some other respect. 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.

[20] It is not obvious however that λ should be the same for both groups. It seems possible to make a case for more or less forecastability of income growth prospects for each education group, and these issues deserve further attention in future work.

4.4 The effect of profile uncertainty on life-cycle savings

The flip side of the consumption choice that we have focused upon so far is the savings decision, and therefore, the wealth distribution also contains useful information that can shed light on the nature of income risk faced by individuals. In particular, we focus on the relationship between an individual's wealth holdings and his labor income, which are positively correlated in the U.S. data.[21] This positive correlation has implications for whether or not the heterogeneity in income growth rates also represent uncertainty from the individuals' point of view. For example, in a purely deterministic world (and assuming $\alpha^i \equiv 0$ for simplicity), individuals with fast income growth will borrow more (or save less) than those with slower income growth in order to smooth consumption over the life-cycle. Consequently, in this simple model wealth holdings and income will be perfectly *negatively* correlated. This negative relationship typically holds true even with sizeable income uncertainty, and is present even when one only allows for a limited amount of heterogeneity in income profiles. For example, when calibrating life-cycle models, researchers often allow β^i to vary *between* education groups, but restrict it to be the same *within* each group (among others, Hubbard, Skinner and Zeldes (1994), Campbell et al. (2001), Davis, Kubler and Willen (2003)), resulting in the counterfactual implication that wealth holdings fall with education level. More generally, such models will typically predict that the income-rich will be the wealth-poor, inconsistent with empirical evidence.

Turning to the PHC model (i.e., $\lambda = 1$), the correlation between an individual's wealth, ω_t^i, and the slope of this profile, β^i, starts from -0.88 at age 25, and while it gradually increases over time, it remains negative until age 60, with an average value of -0.58 (Table 4). As before, individuals with high income growth rates have low wealth holdings in this model as well. Moreover, the correlation of wealth with the *level* of the income profile, $\alpha^i + \beta^i t$, is also negative, averaging -0.40 over the life-cycle. In contrast, if one allows for uncertainty about income growth rates as in the baseline PHU($\lambda = 0.37$) model, the average correlation of ω_t^i with β^i becomes positive (0.26), and the correlation with the level of income profile is 0.39.

These results reiterate the conclusion of the previous section that a life-cycle model with HIP yields counterfactual implications unless uncertainty about these profiles is also taken into account.

5 Conclusion

In this paper, we have studied the consumption-savings behavior when the labor income process is the HIP model, and individuals learn about their profile in a Bayesian fashion. The first finding is that profile uncertainty is resolved very gradually, mainly due to learning about the growth rate of income, which starts slow and becomes faster over time. A second reason is the moderate persistence

[21] Budria-Rodriguez et al. (2002) calculate this correlation to be 0.47 using SCF data. Hurst et al. (1998) provide regression evidence where income enters as a significant determinant of wealth with a very high t-statistic.

of income shocks which results in slower learning compared to both *i.i.d* shocks as well as random walk shocks. Because of slow learning consumption behavior over the life-cycle is informative about initial uncertainty which we used to estimate the prior information individuals have about their income growth rate.

The resulting life-cycle model displays plausible behavior and also shows how the nature of income risk implied by the HIP model (combined with Bayesian learning) is different from that implied by the RIP model. For example, the HIP model generates steeper consumption profiles for college-educated individuals if the larger dispersion of income growth rates translate into higher initial uncertainty and hence more precautionary savings for the former group. Instead, in the RIP model one needs to also assume systematic differences across education groups in preferences and demographics to explain these facts. Moreover, the HIP model generates an age-inequality profile of consumption that is slightly convex, as documented by Deaton and Paxson (1994), as opposed to the concave shape resulting in a life-cycle model with a RIP process.

Another conclusion that we draw from this analysis is that if individuals face the HIP process without uncertainty about income growth rates, the resulting consumption behavior is counter-factual: an individual's income and wealth becomes negatively correlated; borrowing constrained individuals are the income-rich and as a result, the average consumption of constrained individuals is twice that of unconstrained individuals.

In addition, uncertainty about the level of income (α^i), even when it is very large, does not play a significant role in the consumption-savings behavior, because it is resolved very quickly. Thus in future applications of this framework it seems reasonable to eliminate the heterogeneity (or uncertainty) in α^i and reduce the number of state variables by one, which will bring significant computational gain.

An important question that has not been addressed in this paper concerns the origins of the heterogeneity in income growth rates. One plausible framework that we suggested earlier is the human capital model (Ben-Porath 1967) with heterogeneity in the ability to accumulate human capital. One difference however is that in the basic human capital model individuals are assumed to know their ability (and hence income growth) in contrast to our assumption of learning about the growth rate. Fortunately, it is possible to obtain an analytical solution to the Ben-Porath model with Bayesian learning about ability in the presence of *i.i.d* shocks (these results are available upon request). The main finding from this exercise is that learning mainly affects the timing of the dispersion in income growth rates: individuals invest in similar rates early on because they believe they have similar ability levels. As a result, heterogeneity in income growth is small early in the life-cycle, because it is only driven by ability differences. Over time, as individuals find out more about their true ability, they also adjust their investment levels which increases the dispersion of income growth rates further. In future research we intend to introduce a richer shock process, and study the properties of consumption decision in this model.

An appealing feature of the present framework with slow learning is that it could provide a setup for estimating λ from a broader set of economic actions of individuals over the life-cycle. For example, the discussion of equation 12 shows that the dynamic response of consumption to income shocks contains useful information about the amount of profile uncertainty. Similarly, one could augment the current model with other economic decisions, such as labor supply and/or portfolio choice, to bring a wide range of evidence to bear on the estimation of λ. We intend to pursue these issues in future research.

A Appendix: Estimates of the HIP model in the Literature

Table 5 presents the estimates of the HIP model from the U.S. data in the previous literature. As can be seen here, the estimates of σ_β^2 range from 0.00018 in Lillard and Weiss (1979) to 0.00041 in Haider (2001). The former paper estimates a separate income process for each finely defined occupation category (such as chemists, psychologists, etc.), which could be partly responsible for the smaller estimate of profile heterogeneity. However, all the estimates of σ_β^2 are statistically significant, and the latter two papers point estimates are rather close to each other. Baker also report estimates as high as 0.00082; his lowest estimate is 0.00031. Second, the persistence parameter in these studies are around 0.6 to 0.7, indicating significantly lower persistence than a unit root.

B Appendix: Computational Algorithm

This appendix describes the algorithm used to solve the consumption-savings problem described in Section 3. The first point to observe is that since the value function does not explicitly depend on the type of individual we need to solve for only one value function for all individuals. The true type only determines the probability distribution of income (induced by the probability distributions of η and ε for a given (α^i, β^i)) which then determines the probability distribution of the belief vector, $\widehat{\mathbf{S}}^i_{t|t-1}$, for a given agent. In turn, this determines which region of the state space will be most visited for a given individual. To solve the model for a large number of types we need to get a good approximation of the value function for the union of the supports for these different types, which is the challenging part.

We first describe the algorithm for $\lambda = 0$ so that all individuals begin life with the same prior information. A slight modification then will solve the model for different λ values. The critical part of the algorithm is the construction of a convenient grid over which the dynamic problem is solved. Once this is accomplished, solving the model is straightforward.

Step 0: Grid construction

1. Draw I types $\left\{(\alpha^i, \beta^i), i = 1,..,I\right\}$ from a Normal distribution with second moments $(\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta})$ reported in Table 1. In the baseline case, we chose $I = 1000$

2. For each i, simulate J income paths $\left\{\widetilde{y}_t^{ij},\ t = 1,..,T;\ j = 1,..,J\right\}$ using equation (2) to obtain an empirical approximation to the distribution of $\widetilde{y}_t^i$. We chose $J = 100$.

3. For each of the $N \equiv I \times J$ income paths, use equation (3) to obtain a sequence of $\widehat{\mathbf{S}}^{ij}_{t|t-1}$ for $t = 1,..,T$. Thus, for each t, we have $N = 100,000$ points distributed over the 3-dimensional space of beliefs, $(\widehat{\alpha}^i_{t|t-1}, \widehat{\beta}^i_{t|t-1}, \widehat{z}^i_{t|t-1})$. Instead of choosing independent grids in $\widehat{\alpha}^i_{t|t-1}, , \widehat{\beta}^i_{t|t-1}$ and $\widehat{z}_{t|t-1}$ directions and taking the Cartesian product of these intervals, we directly choose points in this 3-dimensional space as follows. We divide the space $[\widehat{\alpha}_{\min}, \widehat{\alpha}_{\max}] \times [\widehat{\beta}_{\min}, \widehat{\beta}_{\max}] \times [\widehat{z}_{\min}, \widehat{z}_{\max}]$ (with appropriately chosen lower and upper bounds) into cubes by taking 21 points in each direction (and get $20 \times 20 \times 20$ cubes). For every t, if there are any points (among the 100,000 realizations of $\widehat{\mathbf{S}}^{ij}_{t|t-1}$) that fall into a cube, we assign a grid point to the center of that cube (and eliminate all empty cubes). This procedure picks a subset of the 3-dimensional space that contains state points that have a non-negligible probability of being realized when we simulate the model. (It is important to emphasize that we do *not* do this for efficiency reasons. Our experience is that attempts at solving for the value function over a Cartesian state spaces runs into a number of difficulties and this is one approach we found to work). We enumerate these triplets $\{\widetilde{\mathbf{S}}_t^q = (\widehat{\alpha}, \widehat{\beta}, \widehat{z})^q,\ q = 1,..,Q_t\}$, where Q_t is the total number of non-empty cubes and hence grid points at age t (From this point on, we drop the reference to t and describe the grid construction for a given age. The same procedure is repeated for each t.)

4. The grid for $\widetilde{y}^{ij}$ needs to be consistent with the probability distribution implied by the type of individual, otherwise one runs into a number of problems.[22] However, since (α^i, β^i) is not a state variable it is

[22] For example, if we attempt to solve the dynamic problem with a $\widetilde{y}_{it}^h$ that is much larger than what would be

not possible to literally have the grid for $\widetilde{y}^{ij}$ depend on the type. Instead then, we choose a different grid for each possible belief vector, $\widetilde{\mathbf{S}}^q$, defined as $\mathbf{y}^q_{grid} = [y^q_{\min}, y^q_{\max}]$ where the bounds are defined as: $\exp(H\widetilde{\mathbf{S}}^q \pm 3\sigma(\widetilde{y}^q))$; $H\widetilde{\mathbf{S}}^q$ is the mean income and $\sigma(\widetilde{y}^q)$ is the standard deviation given in equation (5). In other words, these bounds define a three standard deviation confidence interval for income through equation (5) given beliefs $\widetilde{\mathbf{S}}^q$. We take 8 equally spaced points for each income grid. (Using 20 points did not make a noticeable difference in results.) We repeat the same steps for each t.

5. Unlike the other 4 state variables, wealth does not affect and is not affected by the learning process. Thus, we take a fixed wealth grid—that is, one that does not depend on beliefs or income—with 12 points, more densely spaced near the borrowing constraint, (Using 20 points did not make a noticeable difference in results.) At a given age, the final grid is the Cartesian product of this wealth grid and the (4-dimensional) grid $(\mathbf{y}^q_{grid}, \widetilde{\mathbf{S}}^q)$. So the problem is solved on $(12 \times 8 \times Q_t)$ grid points, where Q_t ranges from 240 to 1100 over the life-cycle and averages 830.

Step 1: Solving the Dynamic Problem

1. The dynamic problem is solved using the Bellman equation approach. We solve the problem for each point on the random grid at age t.

2. The non-Cartesian structure of the state space rules out a number of multi-dimensional interpolation methods such as splines, Chebyshev polynomials that typically require Cartesian grids in more than one dimension. Instead, we approximate the value function with a combination of polynomial functions (up to the 4th power) and other functions (such as logs and fractional powers) of the state variables including various interaction terms between them (a total of 162 terms used in the baseline model). After solving the Bellman equation at age t, we regress the values of the value function at the grid points on these functions of the state variables. These coefficients are then used for the interpolations necessary to evaluate the expectation when solving the period $t-1$ problem.

3. After the model is solved, we simulate the decision rules for a large number of individuals. For simplicity we used the same I types drawn above and the N simulated income paths to obtain consumption-savings paths.

implied by the individual's (α^i, β^i), the Bayesian updating results in next period's beliefs that are substantially away from next period's grid for $\widetilde{\mathbf{S}}^q_{t+1}$, because the latter is constructed based on income realizations that are going to be observed in the actual solution. As a result, one needs to extrapolate next period's value function which often yields extremely inaccurate results (despite the fact that these far-off points have low probability). Considering a $\widetilde{y}^h_{it}$ that is much smaller than what is consistent with the type, results in similar problems as well as creating further problems with infeasible borrowing constraints.

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] 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.

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

[5] Becker, Gary S. (1965): "Human Capital : A Theoretical and Empirical Analysis, with Special Reference to Education," reprinted 1994, University of Chicago Press.

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

[7] Blundell Richard, Luigi Pistaferri, and Ian Preston (2002): "Consumption Inequality and Partial Insurance," Working Paper, Stanford University.

[8] Campbell John Y., Joao F. Cocco, Francisco J. Gomes, and Pascal J. Maenhout (2001): "Investing Retirement Wealth: A Life-Cycle Model," in *Risk Aspects of Investment-Based Social Security Reform,* John Campbell and Martin Feldstein eds., University of Chicago Press: Chicago, IL, pp. 439-473.

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

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

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

[13] Cunha, Flavio, James Heckman, and Salvador Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers,* forthcoming.

[14] 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.

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

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

[17] Guvenen, Fatih (2005): "An Empirical Investigation of Labor Income Processes," Working Paper, University of Rochester.

[18] 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.

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

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

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

[22] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2004): "Risky Human Capital," Working Paper, Wharton School, University of Pennsylvania

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

[24] Krueger, Dirk, and Fabrizio Perri (2004): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," Working paper, New York University.

[25] Krueger, Dirk and Jésus Fernández-Villaverde (2004): "Consumption Over the Life-cycle: Facts from the Consumer Expenditure Survey Data," Working Paper, University of Pennsylvania.

[26] Lillard, Lee A., and Robert T Reville (1999): "Life-Cycle Human Capital Investment: New Evidence on an Old Issue," Working Paper, University of Michigan.

[27] 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.

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

[29] 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.

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

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

[32] Navarro, Salvador (2004): "Understanding Schooling: Using Unobserved Choices to Infer Agent's Information in a Dynamic Model of Schooling Choice When Consumption Allocation is Subject to Borrowing Constraints," Working Paper, University of Chicago.

[33] Budria-Rodriguez, Santiago, Javiar Diaz-Gimenez, Vincenzo Quadrini, and Jose Victor Rios-Rull (2002): "Updated Facts on the U.S. Distributions of Earnings, Income and Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol 26, pp. 2-35.

[34] 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.

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

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

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

[38] 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.

[39] Wang, Neng (2004): "Precautionary Savings and Partially Observed Income," *Journal of Monetary Economics,* Vol. 51, pp. 1645-1681

Table 1: Parameter estimates of the labor income process from Guvenen (2005)

	Group	Model	ρ	σ_α^2	σ_β^2	$\sigma_{\alpha\beta}$	σ_η^2	σ_ε^2
(1)	A	RIP	.988	.058	—	—	.015	.061
			(.024)	(.011)			(.007)	(.010)
(2)	A	HIP	.821	.022	.00038	−.0020	.029	.047
			(.030)	(.074)	(.00008)	(.0032)	(.008)	(.007)
(3)	C	RIP	.979	.031	—	—	.0099	.047
			(.055)	(.021)			(.013)	(.020)
(4)	C	HIP	.805	.023	.00049	−.0024	.025	.032
			(.061)	(.112)	(.00014)	(.0039)	(.015)	(.017)
(5)	H	RIP	.972	.053	—	—	.011	.052
			(.023)	(.015)			(.007)	(.008)
(6)	H	HIP	.829	.038	.00020	−.0007	.022	.034
			(.029)	(.081)	(.00009)	(.0012)	(.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 included in the estimation in all rows, but are not reported to save space. The reported variances are averages over the sample period. These parameter estimates are taken from Guvenen (2005)

Table 2: Decomposing Within-Cohort Income Inequality

	(1)	(2)	(3)	(4)
Age	$\sigma_\alpha^2 + \sigma_\varepsilon^2$	$\frac{1-\rho^{2t+1}}{1-\rho^2}\sigma_\eta^2$	$2\sigma_{\alpha\beta}t + \sigma_\beta^2 t^2$	$\frac{(3)}{(1)+(2)+(3)}$
30	.069	.082	.005	.03
35	.069	.088	.030	.16
45	.069	.089	.135	.46
55	.069	.089	.315	.67
65	.069	.089	.568	.79

Table 3: Baseline Parameterization

Annual model		
Parameter		Value
δ	Time discount rate*	0.962
P^f	Price of risk-free bond	0.96
ϕ	Relative risk aversion	2
$\overline{\beta}$	Avg. inc. growth for all households	0.009
$\overline{\beta}^C$	Avg. inc. growth for college educ.	0.012
$\overline{\beta}^H$	Avg. inc. growth for high school educ.	0.007
T	Retirement age	65
T^*	Age of death	95
Φ	Replacement rate	0.34
$\mathbf{P}_{1\|0}$	The variance of prior beliefs	See text

Note: The parameters of the income process are taken from corresponding rows of Table 1. *The time discount rate is adjusted in each experiment to generate a wealth-to-income ratio of 3.5. The value reported in the table is for baseline PHU model with $\lambda = 0$. See text for details

Table 4: Income-Wealth Correlation Over the Life-Cycle

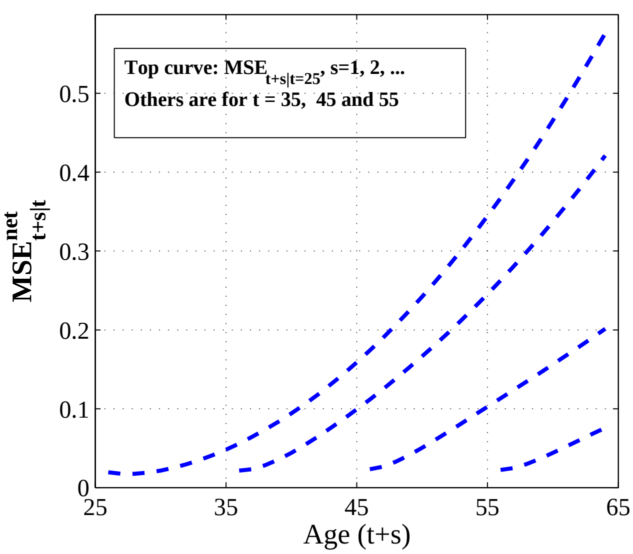
	The average correlation of wealth with:	
	β^i	$\alpha^i + \beta^i t$
PHC ($\lambda = 1.00$)	-0.58	-0.40
PHU ($\lambda = 0.37$)	0.26	0.39

Table 5: Alternative Estimates of the HIP Model

Paper	ρ	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	Stochastic Process	Time effects in variances?
Lillard and	.707	.0305	.00018	.00076	AR(1) +i.i.d	*No*
Weiss (1979)	(.073)	(.0015)	(.00004)	(.0001)		
Baker (1997)	.674	.139	.00039	$-.004$	ARMA(1,2)	*Yes*
	(.050)	(.069)	(.00013)	(.003)		
Haider (2001)	.639	.295	.00041	$-.00827$	ARMA(1,1)	*Yes*
	(.077)	(.137)	(.00012)	(.0036)		

Notes: Lillard and Weiss's data is biannual from the National Science Foundation's Register of Scientific and Technical Personnel covering 1960-70. The reported estimates are from table 7 of their paper, which has the most similar specification to ours. Baker uses PSID data 1967-86, and this result is from Table 4, row 6, which has the best overall fit. Haider's data is also from PSID covering 1967-1992, and the results are from table 4.

Figure 1: The Speed of Resolution of Income Uncertainty Through Bayesian Learning about Profiles



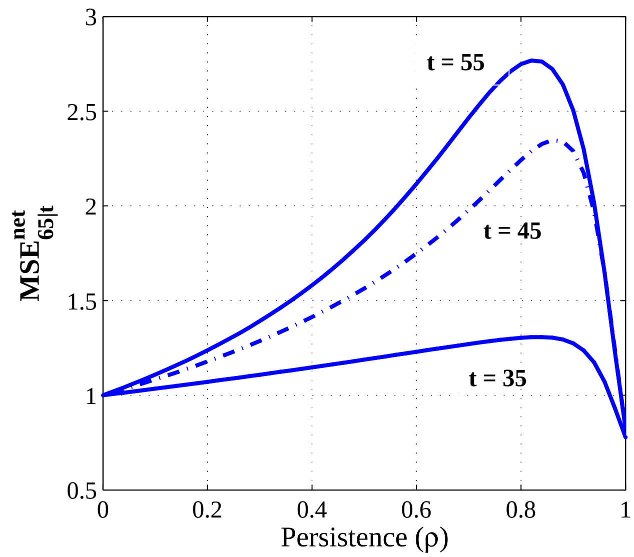


Figure 2: The Change in the Precision of Beliefs about α and β

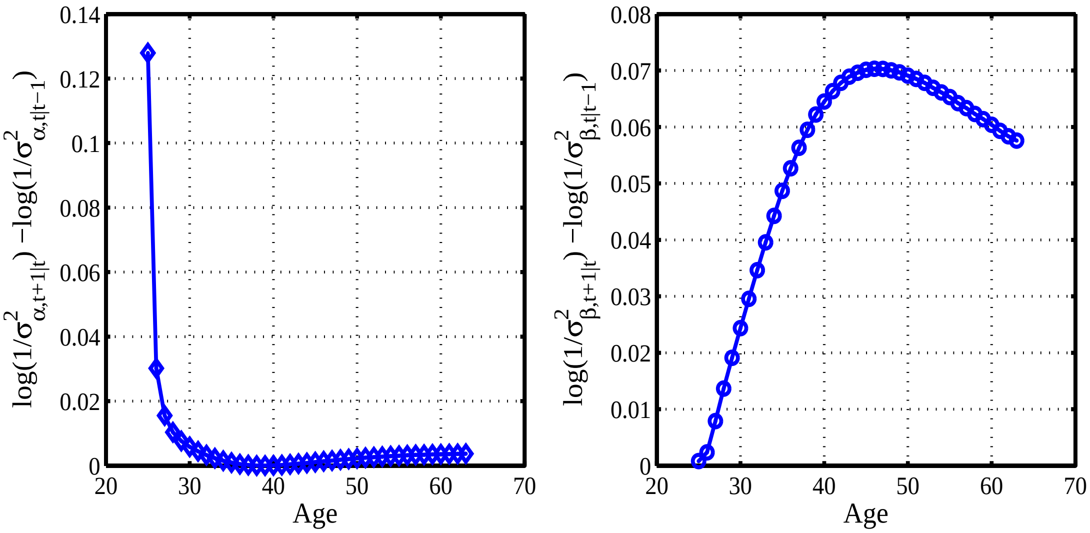


Figure 3: The Change in the Precision of Beliefs about α When $\sigma^2_{\alpha,0}$ is Set to 10 times Its Baseline Value

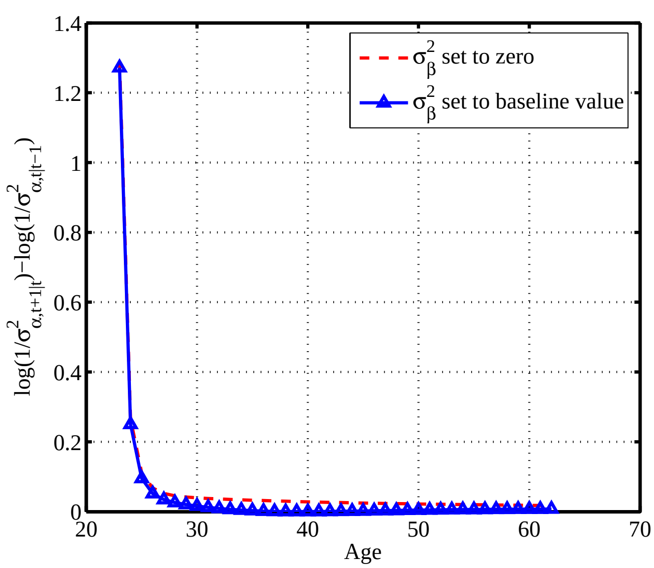


Figure 4: The age-inequality Profile of consumption in the U.S. data and in the PHU model

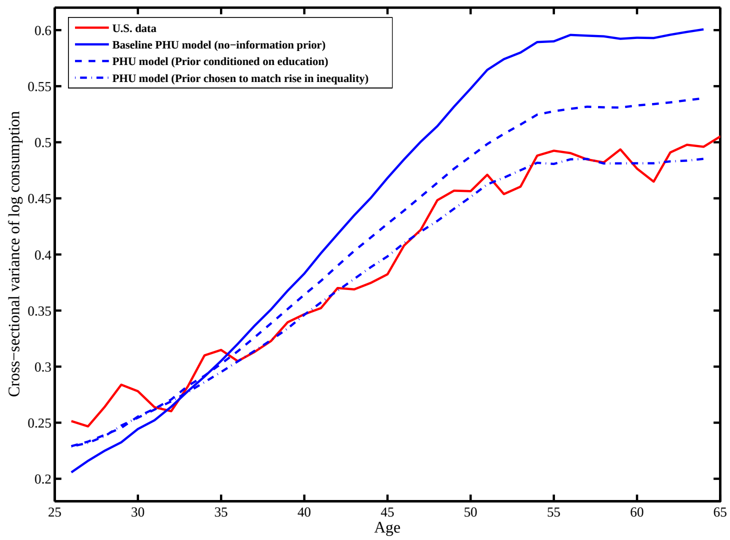


Figure 5: The Age-Inequality Profile of Consumption: Comparisons

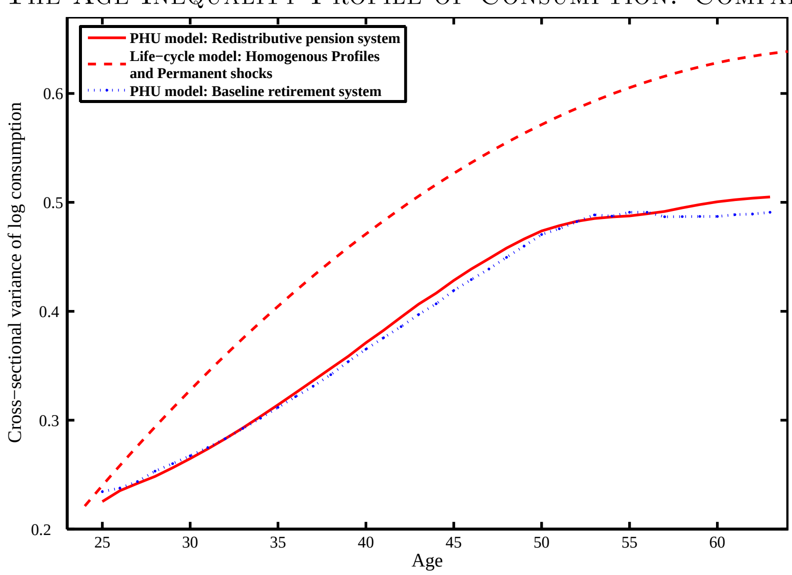


Figure 6: The average income and consumption profiles by education groups in a Lifeycle model with RIP

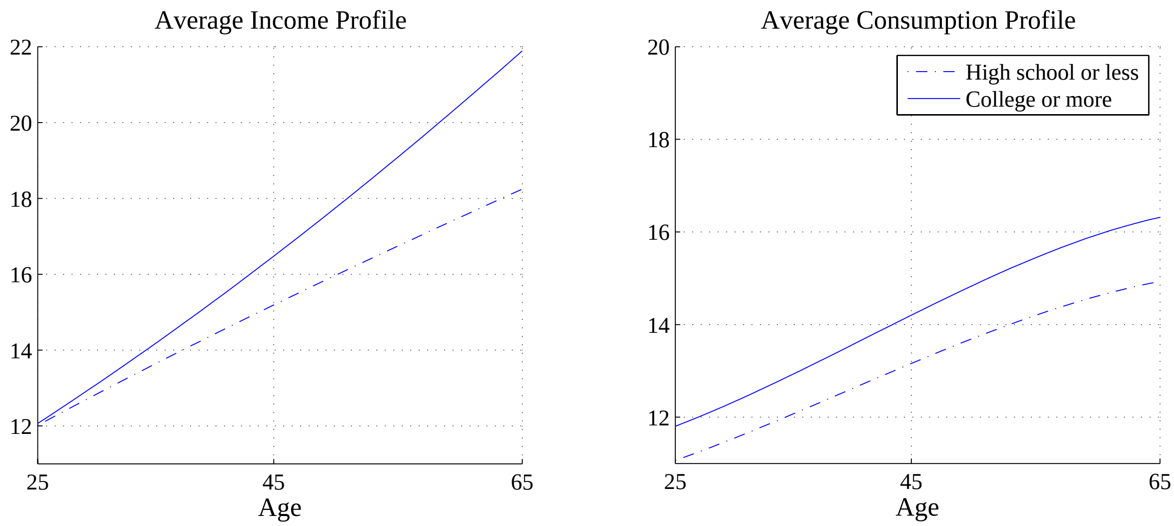


Figure 7: The average income and consumption profiles by education groups in the PHU model