Discussion Paper 145

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

Fatih Guvenen*

University of Texas at Austin

ABSTRACT _____

The current literature offers two views on the nature of the labor 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 (MaCurdy, 1982). According to the alternative view, which we call the "heterogeneous income profiles" (HIP) model, individuals are subject to income shocks with modest persistence while facing individual-specific income profiles (Lillard and Weiss, 1979). In this paper we study the restrictions imposed by the RIP and HIP models on consumption data—in the context of a life-cycle model—to distinguish between these two hypotheses. In the life-cycle model with a HIP process, which has not been studied in the previous literature, 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. The resulting HIP model is consistent with several features of consumption data including (i) the substantial rise in within-cohort consumption inequality, (ii) the non-concave shape of the age-inequality profile, and (iii) the fact that consumption profiles are steeper for higher educated individuals. The RIP model we consider is also consistent with (i), but not with (ii) and (iii). These results bring new evidence from consumption data on the nature of labor income risk.

_____

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. To provide a framework for this discussion, consider the following process for log labor income of individual i with t years of labor market experience:[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. According to the first view in the literature, which we call the "restricted income profiles" (RIP) process, individuals are subject to large and very persistent shocks while facing similar life-cycle income profiles (that is, ρ is close to 1 and σ_β^2 is zero). According to the alternative view, which we call the "heterogeneous income profiles" (HIP) process, individuals are subject to shocks with modest persistence while facing life-cycle profiles that are individual-specific and vary significantly across the population (that is, ρ is significantly less than 1 and σ_β^2 is large).

A vast empirical literature has estimated various versions of the specification in (1) using rich panel data on *labor income* in an attempt to distinguish between these two views. While the results of this literature are arguably best interpreted as more supportive of the HIP model, it is fair to say that this literature has not produced an unequivocal verdict.[2] Thus, to shed further light on this question, we take a different route in this paper: we use the restrictions imposed by the RIP and HIP processes on *consumption* data—in the context of a life cycle model—to bring more evidence to bear on this important question.

Given the weakness of empirical evidence from labor income data in favor of the RIP view, it may seem surprising that the RIP specification is overwhelmingly used as the income process in economic models. Perhaps an important reason for this preference is that the consumption behavior generated in response to the RIP process is consistent with important empirical facts. For example, Deaton and Paxson (1994) have documented the significant rise in within-cohort consumption inequality over

---

[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.

[2] A short list of these studies includes MaCurdy (1982), Abowd and Card (1989), and Topel (1990), which find support for the RIP model; Lillard and Weiss (1979), Hause (1980), and especially the more recent studies such as Baker (1997), Haider (2001), and Guvenen (2005), which find support for the HIP model.

the life cycle (which we reproduce in figure 1 for completeness). As conjectured by these authors and later confirmed by Storesletten et al. (2004a), a life-cycle model is consistent with this observation if idiosyncratic shocks are extremely persistent, as is the case in the RIP process. 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 consumption growth parallels income growth over the life cycle. While this finding may appear to contradict the consumption smoothing motive that underlies the life-cycle framework, several authors have shown that this pattern is consistent with a life-cycle model if income shocks are extremely persistent, again, as in the RIP model (cf. Carroll (1992), and Gourinchas and Parker (2002)).

This congruence between the RIP model's theoretical predictions and consumption data—together with the lack of definitive evidence from labor income data—has made the RIP process the preferred choice for calibrating economic models. Perhaps surprisingly, though, there exists no corresponding study of the consumption behavior when individuals face the HIP process, so the implications of such a model are largely unknown.[3] The goal of this paper is to fill this gap. In particular, we study the implications of the HIP and RIP processes for life-cycle consumption behavior, and provide a systematic comparison of these implications to the U.S. data to assess which income process makes more accurate empirical predictions.

We begin with a detailed investigation of the life-cycle HIP model given that little is known about this framework. This analysis is one of the main contributions of the present paper. Because in the HIP model individuals are ex ante heterogeneous (in their income growth rates), a key question is how much each individual knows about his 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.

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 HIP model learning turns out to be very

---

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

gradual, and its effects on consumption behavior extend throughout the life cycle. The key feature of the model responsible for this result is the presence of learning about the *growth rate* of income. More generally, we show in Section 2.4 that Bayesian learning about a trend parameter (such as β^i in equation (1)) displays some features that are inherently different from, and for our purposes more appealing than, the more standard learning about a level parameter (such as α^i).

In Section 4 we compare the implications of the HIP and RIP models to the U.S. consumption data. We begin by examining the rise in within-cohort consumption inequality—the first empirical fact mentioned above. In the HIP model, as a benchmark we show that when the prior variance of β^i equals the population variance (i.e., when all individuals begin life with the same belief about their income growth prospects), the model generates a rise in consumption inequality that significantly exceeds—by about 50 percent—what is observed in the U.S. data.[4] This finding suggests that one way to estimate the prior variance of β^i is to choose it such that the model exactly matches the rise in inequality in the U.S. data. This procedure yields a prior variance of β^i that equals 40 percent of the population variance of this parameter. The interpretation is that the remaining 60 percent of the variability in income growth rates across the population is forecastable by individuals by the time they enter the labor market. Thus, the HIP model *does* have the potential to generate a substantial rise in consumption inequality, despite the earlier interpretation of this evidence as providing unambiguous support to the RIP model.

The second fact we examine relates to the *shape* of the empirical age-inequality profile. As can be seen in figure 1, this profile has a non-concave shape in the U.S. data, implying that the rise in consumption inequality does not slow down as a cohort gets older (up to about age 55). Several authors have emphasized this fact because the RIP model gives rise to a *concave* shape (Deaton and Paxson (1994), Storesletten, Telmer, and Yaron (2004a)). Instead, we find that the HIP model generates a non-concave profile, which also provides a fairly good fit to its empirical counterpart. Again, learning about the growth rate of income is essential for this result, as we show in Section 4.2.

A third fact, also documented by Carroll and Summers (1991), is that college graduates not only have steeper income profiles than high-school graduates but also have steeper consumption profiles (see figure 2). 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), Guvenen (2005)). As a result, both groups display similar precautionary savings behavior, which in turn results in consumption profiles with similar slopes. Turning to the HIP model, we find that the estimated dispersion of income growth rates among college graduates is more than twice that among high-school graduates. This larger dispersion results in more prior uncertainty, more precautionary savings, and consequently a steeper consumption profile for college graduates.

---

[4] Uncertainty about the intercept of income, α_i, turns out not to play an important role in the model.

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. We conclude from these results that consumption data provide support to the HIP model as a credible contender to the RIP model, and along some dimensions, a more coherent alternative.

This paper is related to several other studies. Huggett, Ventura, and Yaron (2004) have studied a human capital model with heterogeneity in the ability to accumulate human capital. They have also found 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 income growth rate exactly so there is no learning over time. Moreover, they focus on a subset of consumption facts studied in this paper. In addition to Deaton and Paxson (1994) discussed above, the idea of using consumption data to draw inferences about the income process has also been employed by Blundell and Preston (1998) in an attempt to understand whether the changes in income inequality after 1970 were due to an increase in the variance of persistent shocks or transitory shocks. Finally, Pischke (1995) has studied the consumption behavior when individuals cannot distinguish between the aggregate and idiosyncratic components of their labor income and have to solve a signal extraction problem.

The rest of the paper is organized as follows. In the next section, we introduce the RIP and HIP processes and examine the properties of Bayesian learning about profiles in the latter model. In Section 3 we present the life-cycle HIP and RIP models of consumption-savings. 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 processes and discuss the specific parameterizations we use. Second, because individuals are ex ante heterogeneous in the HIP process, 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*, y_t^i, of individual i who is t years old is given by

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

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

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.[5]

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 $\sigma_{\alpha\beta}$.[6]

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, 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. (2004b)).[7] 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 processes 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

---

[5] 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.

[6] 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.

[7] 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.

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 4 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 0.20 has 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 are used in Section 4.3. To our knowledge, the parameter estimates of the HIP model for each education group are 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, $\widetilde{y}_t^i \equiv y_t^i - g$, and use the following equation (derived from equation (2)) to decompose the within-cohort income inequality into its components:

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

The first set of parentheses contains terms that do not depend on age (i.e., the intercept of the age-inequality profile). The second set of parentheses captures the rise in inequality due to the accumulated effect of the autoregressive shock. Finally, the last set of parentheses contains two terms that vary with 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 set of parentheses 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 by 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.

6

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. Therefore, as a benchmark case, we assume that individuals observe only 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:[8]

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

Even though the parameters of the income profile have no dynamics, including them into the state vector 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,

---

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

with $\mathbf{Q}$ and R denoting the covariance matrix of $\boldsymbol{\nu}_t^i$ and the variance of ε_t^i, respectively.[9] 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:[10]

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

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

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

and

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

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

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

---

[9]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.

[10]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 $h_2 = h_1$).

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 parentheses (call it $MSE^{net}_{t+s|t}$) capture 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 3 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.[11] 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 4 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 3, 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 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^2_α is around 0.02, whereas the corresponding point estimate is 0.14 in Baker (1997) and 0.29 in Haider (2001) (see table 4). Figure 5 plots the change in precision of beliefs about α^i when σ^2_α (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

---

[11]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.).

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 3. 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 toward a unit root. In particular, learning is faster when income shocks follow a random walk than for any other value of ρ.[12] Interestingly, the values of ρ where learning is slowest coincide 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 3 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 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.

---

[12]Loosely speaking, this is because when income shocks follow a 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.

2.4.1 Stochastic dynamics for learning uncertainty

So far we have considered the case with homoskedastic Gaussian innovations to the state vector, which implies that $\mathbf{P}_{t|t-1}$ evolves in a deterministic fashion (equation (4)). Nevertheless, it is conceivable that some components of the state vector (such as z_t) could be subject to conditionally heteroskedastic shocks where the innovation variance itself evolves in a stochastic manner.[13] Indeed, Meghir and Pistaferri (2004) report some evidence of an ARCH structure in the variances of labor income shocks, though they only consider RIP processes. In this case, $\mathbf{P}_{t|t-1}$ would also evolve stochastically contrary to our previous analysis. Unfortunately, solving for the consumption-savings decision in the presence of Bayesian learning *and* stochastic volatility would add significant computational challenges into the present framework, which is already quite complicated. Furthermore, for a full comparison of the HIP and RIP models, we would also need to introduce stochastic volatility into the latter model, which would take us well beyond the scope of this paper. Thus, here we briefly examine how learning and the resolution of profile uncertainty is affected by the presence of stochastic volatility, but leave a full analysis of consumption behavior to future work.

Consider the following GARCH(1,1) process for the innovation variance of z_t:

$$\sigma_\eta^2(t) = \omega_0 + \omega_1 \sigma_\eta^2(t-1) + \omega_2 \left(\eta^2(t) - \sigma_\eta^2(t-1)\right),$$

where ω's are positive constants, and the term in the last set of parentheses can be thought of as a zero-mean innovation to volatility. The forecast of future innovation variances have a simple formula:

$$E_t\left(\sigma_\eta^2(t+i)\right) = (\omega_1)^i \left[\sigma_\eta^2(t) - \frac{\omega_0}{1-\omega_1}\right] + \frac{\omega_0}{1-\omega_1}, \text{ for } i = 1, ...$$

To gain a better idea about the effect of stochastic volatility on the MSE, a quantitative example will be useful.[14] We set the persistence of volatility, $\omega_1 = 0.95$, and the size of innovation, $\omega_2 = 0.70$, which generates fairly large and persistent movements in volatility (the standard deviation of $\sigma_\eta^2(t)$ is 1.2 times its average value). Finally, we set $\omega_0 = (1-\omega_1)\sigma_\eta^2$ so that the average volatility remains equal to σ_η^2 as in the benchmark case. Figure 6 plots the results. Because of the stochastic evolution of the posterior variance, there is an entire distribution of $MSE_{t+s|t}$ at each age t depending on the history of innovation volatilities experienced by each individual. To summarize this information, we plot the mean of this distribution at each age (line with circles), as well as the 95 percent confidence bands to quantify the dispersion. For comparison, the dashed line plots the MSE from the benchmark

---

[13] For example, z_t would have stochastic volatility if individuals occasionally switch between occupations that have different levels of idiosyncratic risks.

[14] The GARCH structure modifies the equations for the MSE in two ways. First, in the Kalman recursion (4) which determines $\mathbf{P}_{t|t}$, the matrix $\mathbf{Q}$ needs to be replaced with its forecast, $\mathbf{Q}_{t+1|t}$. This latter differs from $\mathbf{Q}$ only in its lower diagonal element, which now equals $E_t\left(\sigma_\eta^2(t+1)\right)$. Similarly, we replace $\mathbf{Q}$ in equation (7) with its forecast, $\mathbf{Q}_{t+i|t}$.

model (taken directly from the left panel of figure 3).

Several points can be seen from this graph. First, the average MSE with GARCH shocks is somewhat lower than its benchmark counterpart. Moreover, the difference between the two appears to slightly widen with age, indicating a mild acceleration in the speed of learning in the presence of stochastic volatility. In addition, there is a fair amount of variation as indicated by the confidence bands. However, reducing the persistence of the GARCH process, to 0.80 for example, substantially narrows the gap between the two series (not shown). Another point to note is that the *shape* of the average MSE profile is virtually identical to that in the benchmark case. This feature is not specific to the average, but is also true for each individual MSE profile in that distribution. This result is important because this shape determines how individuals perceive their future income risk, and consequently, has a major effect on their consumption decision as we discuss further below. These results suggest that introducing stochastic volatility in z_t may not greatly affect consumption behavior in the HIP model, especially if conditional volatility is not as persistent as we assumed in this example. Nevertheless, a more definitive answer awaits a full investigation of consumption behavior, which is left for future work.

3 Two Life-Cycle Models of Consumption and Savings

In the rest of the paper we embed each income process (HIP and RIP) into a life-cycle model. Then, by comparing the implications of each model to the U.S. data, we will assess which type of income process makes more accurate predictions about life-cycle consumption behavior. The two life-cycle models will be kept as similar to each other as possible to highlight the differential impact of each income process on consumption behavior. We now describe each model in turn.

3.1 The HIP model

Consider an environment where each individual lives for T^* years and works for the first T $(< T^*)$ years of his life, after which he retires. Individuals do not derive utility from leisure and hence supply labor inelastically. During the working life, the income process is given by the HIP process specified in equation (2). During retirement, the individual receives a pension that is designed to mimic the current U.S. Social Security system as described in further detail below. There is a risk-free bond that sells at price P^b (with a corresponding net 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 $\underline{W}_{t+1}$, specified below.

The relevant state variables for this dynamic problem are the asset level, ω_t^i, current log income, y_t^i, and the last period's forecast of the true state in the current period, $\widehat{\mathbf{S}}_{t|t-1}$. In the following equations the superscript i is included in individual-specific variables to distinguish them from

aggregate variables. The dynamic problem of a typical individual can then be written as

$$V_t^i(\omega_t^i, y_t^i, \widehat{\mathbf{S}}_{t|t-1}^i) = \max_{c_t^i, \omega_{t+1}^i} \left\{ U(c_t^i) + \delta E \left[V_{t+1}^i(\omega_{t+1}^i, y_{t+1}^i, \widehat{\mathbf{S}}_{t+1|t}^i) | \widehat{\mathbf{S}}_{t|t-1}^i \right] \right\}$$

$$\text{s.t.}$$

$$c_t^i + P^b \omega_{t+1}^i = \omega_t^i + Y_t^i \tag{9}$$

$$\omega_{t+1}^i \geq \underline{W}_{t+1}, \tag{10}$$

$$eq. \ (3, 4)$$

for $t = 1, ..., T - 1$, where $Y_t^i \equiv e^{y_t^i}$ is the *level* of income, and V_t^i is the value function of a t year-old individual.[15] 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_{t+1}^i given by equation (5).

During retirement, pension income is constant, and since there is no other source of uncertainty or learning, the problem simplifies significantly:

$$V_t^i(\omega_t^i, y^i) = \max_{c_t^i, \omega_{t+1}^i} \left[U(c_t^i) + \delta V_{t+1}^i(\omega_{t+1}^i, y^i) \right] \tag{11}$$

$$\text{s.t } Y^i = \Phi\left(Y_T^i\right), \text{ and } eqs. \ (9, 10)$$

for $t = T, ... T^*$, with $V_{T^*+1} \equiv 0$, and $\Phi\left(Y_T^i\right)$ is a function determining the pension replacement rate.

3.2 The RIP model

The second model is essentially the same as the first one, with the exception that the income process is now given by a RIP process. Because with a RIP process all individuals share the same life-cycle income profile (α, β), there is no learning about individual profiles, which reduces the number of state variables significantly—from five above to two here. Specifically, the dynamic programming problem of a typical worker is

$$J_t^i(\omega_t^i, z_t^i) = \max_{c_t^i, \omega_{t+1}^i} \left\{ U(c_t^i) + \delta E \left[J_{t+1}^i(\omega_{t+1}^i, z_{t+1}^i) | z_t^i \right] \right\}$$

$$\text{s.t. equations } (9, 10)$$

---

[15] 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 $t = 1, ..., T - 1$, where J_t^i is the value function of a t year-old individual. Notice that we assume the worker observes the persistent component of the income process, z_t^i, separately from y_t^i. This is the standard assumption in the existing consumption literature which uses the RIP process, and we follow them for comparability. Finally, because there is no income risk after retirement, the problem of a retiree is the same in this model as in the HIP model given above in (11).

3.3 Baseline parameterization

There is no analytical solution to the dynamic optimization problems stated in the previous section, so we resort to numerical methods. While the computational method used to solve the RIP model is standard and has been employed by several earlier studies, the numerical solution of the HIP model is significantly complicated by the fact that there are five continuous state variables and four of them (excluding ω_t^i) depend on each other as a result of Bayesian learning. In particular, this interdependence 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.

In the remainder of this section we describe the parameterization of each life-cycle model. To make the quantitative exercise meaningful, we choose the same values for parameters that appear in both models (except for one parameter as explained below).

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 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 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 and from the National Income and Product Accounts, 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 4, which is at the lower end of these reported values. Notice that because the amount of precautionary savings depends on the amount of uncertainty, which is different in the RIP and HIP models, this procedure implies (slightly) different values for the time preference rate in each model: $\delta^{HIP} = 0.966$ and $\delta^{RIP} = 0.964$. Furthermore, when we make comparisons across different versions of the model (by varying λ, for example), 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 (in the HIP process), it does not identify their means. The intercept term, α, is simply a scaling parameter, 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 individuals with low and high education, respectively.

Redistributive Social Security.—The U.S. retirement pension system features significant redistribution, thereby providing risk-sharing not only across cohorts but also within each cohort. The extent of this risk-sharing in turn is critical for the rise in consumption inequality over the life cycle. For example, with complete risk-sharing, consumption inequality would be constant over the life cycle in both the HIP and RIP models. Thus, a realistic modeling of the retirement system is essential for a sound quantitative evaluation of the consumption behavior in these models.[16] We adopt the following formulation for the pension system, which captures the salient features of the U.S. Social Security system as described in Storesletten et al. (2004a). Specifically, the retirement replacement rate is a concave function of an individual's income at age T given by

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

where $\overline{Y}_{T}$ is the average income at age T, and $\overline{\Phi}$ is a scaling parameter.[17]

Determining the priors in the HIP model.—Empirical evidence is not particularly helpful for setting $\mathbf{P}_{1 \mid 0}$. 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 his 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 $\beta^{i}=\beta_{k}^{i}+\beta_{u}^{i}$, where β_{k}^{i} and β_{u}^{i} are two random variables, independent of each other, with zero mean and variances of $\sigma_{\beta_{k}}^{2}$ and $\sigma_{\beta_{u}}^{2}$. Clearly then, $\sigma_{\beta}^{2}=\sigma_{\beta_{k}}^{2}+\sigma_{\beta_{u}}^{2}$. The key assumption we make is that individual i observes the realization of β_{k}^{i} but not of β_{u}^{i} (hence the subscripts indicate *k*nown and *u*nknown,

---

[16]Clearly, the retirement system is not the only mechanism providing risk-sharing among individuals. In general, the market structure—the types of assets, in addition to a risk-free bond, available to individuals for consumption smoothing—would also affect consumption behavior. Many such extensions however would significantly complicate the present model and hence are left for future work.

[17]There is one difference between this specification and the one in Storesletten et al. (2004a): 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 $\overline{\Phi}=1 / 1.40 \approx 0.715$ to match the average level of benefits in their specification.

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 $\sigma^2_{\alpha,0} = \sigma^2_\alpha$. The prior covariance matrix is

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

To emphasize that the implications of the HIP model depend on the amount of prior uncertainty facing individuals, we henceforth refer to this general framework as the HIP(λ) 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 the HIP model, however, this condition requires that each individual face a different natural limit, unlike in the RIP model where individuals are ex-ante identical. 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. For comparability, we impose the same borrowing constraint in the RIP model, which almost never binds in our simulations. Finally, 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.

4   Quantitative Results

In this section we compare the life-cycle consumption behavior implied by the HIP(λ) model and the RIP model to the U.S. data. The main goal of this exercise is to assess which income process provides more accurate predictions about life-cycle consumption behavior.

To simulate the HIP model we first draw 1000 (α, β) combinations (each draw corresponding to one type of individual) from a bivariate Normal distribution whose covariance matrix is taken from row 2 of table 1. Then for each (α^i, β^i) type, we simulate 100 income paths, and calculate the corresponding consumption and wealth paths implied by the HIP model. As for the RIP model, we simulate income paths for 10,000 individuals who all have the same (α, β) combination but differ in the realization of their idiosyncratic shocks. In all cases we assume that individuals start life with no assets. The reported statistics are calculated from these simulated data.

4.1   The age-inequality profile of consumption

We first examine the implications of each model for the rise in within-cohort consumption inequality over time. As can be seen in figure 1, in the U.S. data the cross-sectional variance of log consumption increases by about 21 log points over the life cycle, implying that consumption inequality more than doubles during this time.[18] 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 (1994) have discussed several potential explanations and found the existence of persistent (uninsurable) idiosyncratic shocks to be the most promising candidate. Recently, Storesletten et al. (2004a) have shown that a life-cycle model *can* quantitatively match the rise in inequality observed in the data *if* income shocks are extremely persistent. It should be noted, however, that both studies have restricted attention to RIP models, and neither one has allowed for a HIP process for income. Here, we analyze consumption inequality in both the RIP and the HIP models.

---

[18] 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). 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 are not available during 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 estimate the income processes in table 1. See Krueger and Perri (2004) for further details of sample selection and variable construction. To obtain the graph in figure 1, 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 obtain a reasonably large number of observations in each age cell, we define 5-year age bands. For example, to calculate the variance at age h, we use all observations on individuals aged $(h-2)$ to $(h+2)$. The age dummies are scaled so that the average inequality matches that in the sample. Partly due to these differences, the rise in inequality is 21 log points in figure 1 compared to roughly 25 log points in Deaton and Paxson (1994). However, the general pattern of rise in inequality (including the non-concavity, which we comment on below) is very similar in the two papers.

In figure 8, the middle line (with circles) plots the cross-sectional inequality implied by the baseline RIP model, which shows an increase of 26 log points over the life cycle, even exceeding the rise in the data. For completeness, as well as for further comparison later, we also plot the corresponding graph from the benchmark RIP model of Storesletten et al. (2004a). The (top) dashed line in figure 8 shows an even larger increase in inequality—about 40 log points—but has an overall shape that is very similar to the one from our baseline RIP model.[19] Overall, these results reiterate the earlier findings in the literature that a RIP model generates significant fanning out of the consumption distribution as observed in the data.

We next turn to the HIP(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) dashed line in figure 7 plots the age-inequality profile, which also shows a substantial rise in consumption inequality over the life cycle—roughly 32 log points.[20] Thus, the HIP(λ) model *is also* capable of generating a significant rise in consumption inequality, despite the previous interpretation of this empirical evidence, discussed in the Introduction, as supporting the RIP model.

Before proceeding further, it is useful to discuss the mechanism behind increasing inequality in each model. For the sake of this discussion, we make the following simplifying assumptions: (i) the utility function is quadratic, and $\delta = 1 + r^f = 1$; (ii) there is no retirement (i.e., all individuals die at the end of working life); and (iii) in the RIP model income shocks are fully permanent ($\rho = 1.0$). Under these assumptions, in both models consumption will be adjusted every period by the change in perceived lifetime income divided by the number of years left in life:

$$\Delta c_t = \frac{1}{T - t + 1} \left[\sum_{s=0}^{T-t} (E_t - E_{t-1}) Y_{t+s} \right]. \tag{12}$$

The left panel of figure 9 plots the simulated income and consumption paths of an individual in the RIP model (who happens to have a realized income growth that is positive). Using the facts that income follows a random walk (so $E_t(Y_{t+s}) = Y_t$ for all s), we have $\Delta c_t = \Delta Y_t$ for all t. Thus, we obtain the well-known result that in the (certainty-equivalent version of the) RIP model, permanent income movements are fully accommodated. If furthermore initial wealth is zero, consumption equals income every period. Since this is true for every individual in this economy, consumption inequality will rise one-for-one with income inequality over time.

We next turn to the HIP model (with the simplifying assumptions). The right panel of figure

---

[19]Storesletten et al.'s (2004b) benchmark model does not feature a redistributive Social Security system, which largely explains why the rise in inequality is very different in their RIP model compared to ours. However, they do introduce it as an extension, which then results in a smaller rise in inequality similar to ours (of about 25 log points.)

[20]Only a small fraction of this fanning-out is directly attributable to idiosyncratic shocks: if we eliminate profile heterogeneity and consequently learning from the HIP model (without changing the parameter values for persistence, etc.), the rise in inequality would only be 7 log points.

9 plots the income path of an individual whose income grows by a constant amount every period (dashed line).[21] We consider two extreme scenarios to illustrate the importance of the speed of learning for the rise in consumption inequality. First, suppose that the individual exactly knows his true income path from the beginning. Clearly, he will set his consumption once and for all and never adjust it over his lifetime (line with circles). As a result, in an economy populated by individuals with different income growth rates, consumption inequality will be constant as the cohort ages, despite the fact that income inequality is monotonically increasing during the same time.

Consider now the opposite case: suppose that the individual starts life with the belief that his income growth rate is equal to the population average of zero, but then *never* updates his beliefs as new income observations are realized (thus, $E_j(Y_t) = E_s(Y_t)$ for all $s, j < t$). In this case, it is easy to see from the expression in (12) that $\Delta c_t = \frac{Y_t - E_{t-1}(Y_t)}{T-t+1} = \frac{Y_t - Y_0}{T-t+1}$. Because the individual does not update his beliefs, he interprets the deviations of income from his expectations as transitory differences and spreads them over his lifetime. Early in life, the deviation of actual income from expected income (the numerator) is small and the planning horizon (the denominator) is large, so consumption is adjusted by an amount much smaller than the actual increase in income. Over time, the deviations become larger (since Y_t is growing) while the horizon gets shorter, resulting in larger changes in consumption. The solid line in the right panel shows the consumption path in this case. The interesting point to note is that consumption changes later in life significantly exceed even the changes in income. As a result, in an economy populated by such individuals, consumption inequality would not only grow, but in fact grow more than income inequality.

Although in our baseline HIP model with Bayesian learning individuals *do* update their beliefs, as discussed in Section 2.4 optimal learning turns out to be very slow, and the present analysis shows that this slow learning provides a powerful force that amplifies consumption inequality over time.

4.1.1 Measuring the prior uncertainty

The preceding discussion suggests one way to pin down the value of λ from the data. To see this, note that a higher λ means less prior uncertainty, which in turn implies less profile uncertainty over the life cycle, and thus a smaller rise in consumption inequality. Given that the HIP model with $\lambda = 0$ generates more inequality than in the data, we could keep increasing λ to the point where the model matches its empirical counterpart (of 21 log points). This procedure yields $\lambda^* = 0.62$. In figure 7 the middle line (marked with diamonds) plots the implied age-inequality profile. Notice that although λ was chosen to match the *total* rise in inequality, the overall *shape* of the resulting profile also provides a nice fit to its empirical counterpart. We return to this point in the next section.

The estimated value of λ implies that 62 percent of the variability in income growth rates is

---

[21] The graphical intuition is more easily explained assuming a constant "amount of growth" rather than a constant "growth rate." However, the basic argument applies equally well to the latter case.

forecastable by individuals at the time they enter the labor market.[22] While this estimates indicates a great deal of initial predictability in income profiles, the remaining uncertainty is still economically substantial: the standard deviation of the prior belief about β^i is $((1 - 0.62) \times 0.00038)^{1/2} \approx 0.012$, compared to a mean of 0.009, implying, for example, that an individual who is one standard deviation above the mean would earn roughly 2.6 times more income at retirement age than another individual who is one standard deviation below the mean.

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_β^2. As a result, λ^* also contains any forecastability in β^i that is due to differences in education level, making it a useful summary measure relevant for models that do not distinguish among individuals based on education. However, it is also of interest to know the degree of forecastability *conditional* on education, and also allow for the fact that forecastability might vary by education.

To this end, we follow the same procedure as above to estimate a separate λ for each group. We first construct empirical age-inequality profiles from the CEX for each group displayed in figure 10. As can be seen here, the cross-sectional variance of log consumption increases by 26 log points in the college sample (left panel), compared to 17 log points in the high-school sample (right panel). We then solve the HIP model for each education group, using the (HIP) income processes reported on rows 4 and 6 of table 1. Then, for each group we choose a separate λ so that the rise in inequality in each model matches its empirical counterpart.[23] This procedure yields $\lambda_C^* = 0.55$ and $\lambda_H^* = 0.32$, implying that the degree of forecastability of income profiles is increasing with education, partly offsetting the significantly higher profile heterogeneity among the college-educated group. Overall, however, the amount of prior uncertainty (that is, $(1 - \lambda)\sigma_\beta^2$) is increasing with education, which results in a larger rise in inequality in the college sample. The resulting age-inequality profiles are plotted in figure 10 (lines marked with circles). For the college sample, the model does a fairly good job of capturing the shape of the empirical profile, while for the high-school sample the empirical profile exhibits more convexity in the early parts of the life cycle compared to the model.

Finally, the simulated data for each education group can be combined to provide an alternative way to construct the age-inequality profile of consumption for the whole population, which can then be compared to its empirical counterpart previously analyzed in figure 7. In figure 7 the line marked with circles displays the result. Consumption inequality in the combined data (overall sample) now rises by 20 log points—compared to 21 log points in the data—and provides an arguably better fit

---

[22] 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 schooling returns remain largely forecastable.

[23] As noted above, one caveat that should be kept in mind is that the measured λs also capture differences in insurance opportunities between the groups in addition to the degree of forecastability.

to the shape of the empirical profile than the baseline HIP model (marked with diamonds). It is reassuring that even though the estimated HIP processes differ in important ways across education groups (and compared to the whole population), the implied consumption behavior is consistent with the evidence for the overall population.

Overall, these results show that the HIP model easily generates a significant rise in consumption inequality over the life cycle. This is true even when a sizeable fraction of profile heterogeneity is forecastable.

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

A second feature of the age-inequality profile that is apparent in figure 1 is its *non*-concave shape, up to about age 55. We also plot 95 percent confidence bands (obtained by bootstrap) around the graph, which helps to show that this shape is fairly accurately determined. Nevertheless, based on this information alone, it still seems hard to rule out a concave profile that might (perhaps barely) fit through the confidence bands. Thus, to tackle this question more directly, we characterized the sampling distribution of the *curvature* of the age-inequality profiles. We first drew 500 bootstrap samples with replacement and estimated the age-inequality profile from each sample. Then we fitted a second order polynomial to each profile from age 25 to 55. The estimated coefficients on the quadratic term were all positive—ranging from 2.6×10^{-4} to 3.2×10^{-3} with a mean value of 1.44×10^{-3}—confirming that the slightly convex shape of the empirical inequality profile is robust to sampling uncertainty. It is now of interest to see whether each model is consistent with this empirical fact.

Earlier studies have found it hard to reconcile the RIP model with this evidence. For example, using the certainty equivalent version of the permanent income model, Deaton and Paxson (1994) have shown that the inequality profile will be *concave* if the income process has a large persistent component. More recently, Storesletten et al. (2004a) have studied a more general model with a rich set of realistic features and found concavity to be a robust implication of the RIP model. The same result holds true in our baseline RIP model despite several differences in our specifications. This concavity can be seen in figure 8, which plots the age-inequality profile from our baseline RIP model, as well as from Storesletten et al. (2004a) for comparison.

In the HIP(λ) model, instead, inequality rises in a slightly convex/linear fashion up to about age 55. The main driving force behind this result is Bayesian learning about income growth rates. The main intuition could partly be anticipated from the discussion in the previous section (and shown in figure 9), which showed an increase in the size of consumption adjustments as the cohort ages. We now show that this result holds more generally—when individuals update their beliefs optimally in a Bayesian fashion and $r^f \neq 0$. The mechanism can be explained 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

$$\Delta c_t = \frac{1}{\varphi_t} \left[(1 - \gamma) \sum_{s=0}^{T-t} \gamma^s \left(E_t - E_{t-1} \right) Y_{t+s} \right], \tag{13}$$

where $\gamma = 1/\left(1 + r^f\right)$ 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: $Y_t^i = \alpha^i + \beta^i t + \varepsilon_t^i$.[24]

First, consider the case where individuals learn about their intercept α^i but know β^i exactly. It can easily be shown that (13) 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 case where individuals learn about β^i but know α^i with certainty. When an individual updates his beliefs in period t, the revision in expected future income is $\left(E_t - E_{t-1}\right) Y_{t+s} = \left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right)(t + s)$. Substituting this expression into (13) and after performing some algebra, one can show that

$$\Delta c_t = \left[\left(\frac{\gamma}{1 - \gamma} \right) + \frac{t - (T+1)\gamma^{T-t+1}}{1 - \gamma^{T-t+1}} \right] \left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1} \right). \tag{14}$$

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 4—that the speed of learning about β^i increases over time so the absolute value of $\left(\widehat{\beta}_{t|t} - \widehat{\beta}_{t-1|t-1}\right)$ is getting larger on average up to age 50. Therefore, consumption changes will be larger in absolute value as the cohort ages, implying a convex shape for the age-inequality profile. While in the baseline HIP model there is learning about both α^i and β^i, the former happens very quickly and thus 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 third well-documented empirical finding is that consumption growth parallels income growth over the life cycle (Carroll and Summers (1991)). While at first glance this finding may appear to contradict the consumption smoothing motive inherent in permanent income theory, several authors have shown that it is possible to generate this pattern if individuals display precautionary savings behavior (i.e., if they have CRRA preferences, for example) *and* if income shocks are extremely persistent as in the RIP model. 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

---

[24] These assumptions are rather innocuous in this context. See the working paper version for a justification of this assertion.

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)).

Another finding documented by Carroll and Summers, however, poses a challenge to this explanation. In particular, 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. In figure 2 we replicate the same finding using data from our CEX sample: consumption grows by 74 percent between ages 25 and 55 for the college group, compared to 36 percent for the high-school group.[25]

To explain this observation by precautionary savings alone as in the RIP model, it would require the former group to face income shocks that are either more persistent or more volatile than the latter group,[26] neither of which we seem to find in the data. For example, the estimates reported in rows 3 and 5 of table 1 reveal little difference between the RIP processes faced by each group, which is consistent with earlier studies in the literature (Hubbard et al. (1994), Carroll and Samwick (1997)). Using these estimated processes, we solve the RIP model separately for each education group. The right panel in figure 11 displays the consumption profiles of each group normalized to be 1 at age 25. It can be readily seen that consumption growth over the life cycle is virtually the same—43.4 percent versus 42.7 percent—for the two groups, inconsistent with empirical evidence.

We next turn to the HIP model. First, notice that the estimates of the HIP process reveal an important difference between the two groups (rows 4 and 6 of table 1): while the persistence and innovation variance are still similar, the heterogeneity in β^i is dramatically different, with college graduates facing a much wider dispersion of income growth rates ($\sigma_\beta^2 = 0.00049$) than less educated individuals ($\sigma_\beta^2 = 0.00020$). To solve the HIP model for each group, we also use the values of λ that were calibrated in Section 4.1.1 to match the rise in consumption inequality for each group: $\lambda_C^* = 0.55$ and $\lambda_H^* = 0.32$. The left panel of figure 11 plots the resulting average consumption profile, which shows a 53 percent increase for the college group compared to a 29 percent increase for the high-school group. Therefore, although the consumption growth of both groups is understated compared to the data, the HIP model does imply a steeper consumption profile for individuals with higher education. Notice that this difference would have been even larger if it were not for the offsetting effect of higher forecastability for the high-education group. Overall, this analysis shows that a basic version of the HIP model is consistent with differences in consumption profiles between education groups, while the baseline RIP model is not. However, the RIP model can also be made

---

[25] Similarly, Fernández-Villaverde and Krueger (2005) employ non-parametric methods and document that *total* consumption expenditures (as opposed to our use of non-durable expenditures) increase by about 50 percent for the college group and by about 10 percent for the high-school group.

[26] 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.

consistent with this fact by introducing two additional features—systematic differences between education groups in their demographics characteristics as well as in unobservable dimensions such as their time discount factors—as shown by Attanasio et al. (1999).

5   Conclusion

In this paper, we have studied the implications of the HIP and RIP models for life-cycle consumption behavior and compared these implications to the U.S. data to assess which type of income process implies more accurate empirical predictions. While the RIP model has been studied extensively in the previous literature, this has not been the case for the HIP model. One contribution of the present paper, therefore, is to provide a detailed investigation of a life-cycle HIP model. An important finding from this analysis is that the uncertainty about income growth rates is resolved very gradually, and we highlighted two channels that were responsible for this slow learning. Moreover, 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 HIP and RIP models display some similarities as well as differences regarding their consumption implications. For example, both models are consistent with the substantial rise in consumption inequality over the life cycle observed in the U.S. data. However, while the shape of this profile is also consistent with the data in the HIP model, this is not true in the RIP model. Furthermore, the HIP model generates steeper consumption profiles for college-educated individuals compared to lower education individuals as in the data, which again was not the case in the RIP model. As shown by Attanasio et al. (1999), one needs to also assume systematic differences across education groups in preferences and demographics to explain these facts in the RIP model.

Another conclusion that we draw is that if individuals face the HIP process without uncertainty about income growth rates, the resulting consumption behavior is counterfactual: an individual's income and wealth become 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. These results are reported in the working paper version of this article.

An appealing feature of the present framework with slow learning is that it could provide a setup for estimating λ from a broad set of economic actions of individuals over the life cycle. For example, the discussion of equation (14) shows that the dynamic response of consumption to income shocks contains useful information about profile uncertainty. Similarly, one could augment this 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 4 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 reports estimates as high as 0.00082; his lowest estimate is 0.00031. Second, the persistence parameter in these studies is 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, $\widehat{\mathbf{S}}_{t|t-1}^i$, 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 $\left\{\left(\alpha^i, \beta^i\right), i = 1, .., I\right\}$ 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 $\widehat{\mathbf{S}}_{t|t-1}^{ij}$ for $t = 1, .., T$. Thus, for each t, we have $N = 100,000$ points distributed over the 3-dimensional space of beliefs, $(\widehat{\alpha}_{t|t-1}^i, \widehat{\beta}_{t|t-1}^i, \widehat{z}_{t|t-1}^i)$. Instead of choosing independent grids in $\widehat{\alpha}_{t|t-1}^i$, $\widehat{\beta}_{t|t-1}^i$, and $\widehat{z}_{t|t-1}^i$ 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 $\widehat{\mathbf{S}}_{t|t-1}^{ij}$) 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 space 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.[27] However, since (α^i, β^i) is not a state variable it is

---

[27] 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 $\mathbf{y}_{grid}^q = [y_{\min}^q, y_{\max}^q]$, 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 $(\mathbf{y}_{grid}^q, \widetilde{\mathbf{S}}^q)$. 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.

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, 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.

---

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 $\widetilde{\mathbf{S}}_{t+1}^q$, 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 $\widetilde{y}_{it}^h$ 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.

[5] Blundell, Richard, and Ian Preston (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, Vol. 113, pp. 603–640.

[6] Budría-Rodríguez, Santiago, Javiar Díaz-Giménez, Vincenzo Quadrini, and José-Victor Ríos-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.

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

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

[9] Carroll, Christopher D., and Andrew Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, Vol. 40, pp. 41–71.

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

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

[12] Cunha, Flavio, James Heckman, and Salvador Navarro (2005): "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers,* Vol. 57, pp. 191–261.

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

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

[15] Fernández-Villaverde, Jesus and Dirk Krueger (2005): "Consumption Over the Life-cycle: Facts from the Consumer Expenditure Survey Data," *Review of Economics and Statistics,* forthcoming.

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

[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 States: 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. 1013–1029.

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

[21] Huggett, Mark, Gustavo Ventura, and Amir Yaron (2004): "Sources of Lifetime Inequality," Working Paper, Wharton School, University of Pennsylvania.

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

[23] 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.

[24] Lucas, Robert E., Jr. (2003): "Macroeconomic Priorities," *American Economic Review,* Vol. 93, pp. 1–14.

[25] 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.

[26] Meghir, Costas, and Luigi Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, Vol. 72, pp. 1–32.

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

[28] 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.

[29] Pischke, Jörn-Steffen (1995): "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica,* Vol. 63, pp. 805–840.

[30] Storesletten, Kjetil, Chris Telmer, and Amir Yaron (2001): "How Important Are Idiosyncratic Shocks? Evidence from Labor Supply," *American Economic Review,* Vol. 91, pp. 413–417.

[31] Storesletten, Kjetil, Chris Telmer, and Amir Yaron (2004a): "Consumption and Risk-sharing Over the Life-cycle," *Journal of Monetary Economics,* Vol. 51, pp. 609–633.

[32] Storesletten, Kjetil, Chris Telmer, and Amir Yaron (2004b): "Cyclical Dynamics in Idiosyncratic Labor-Market Risk," *Journal of Political Economy*, Vol. 112, pp. 695–717.

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

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

	Group	Model	ρ	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	σ^2_η	σ^2_ε
(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

Age	(1) $\sigma^2_\alpha + \sigma^2_\varepsilon$	(2) $\frac{1-\rho^{2t+1}}{1-\rho^2}\sigma^2_\eta$	(3) $2\sigma_{\alpha\beta}t + \sigma^2_\beta t^2$	(4) $\frac{(3)}{(1)+(2)+(3)}$
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[*]	0.966
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 HIP model with $\lambda = 0$. See the text for details.

Table 4: ALTERNATIVE ESTIMATES OF THE HIP MODEL

Source	ρ	σ_α^2	σ_β^2	$\sigma_{\alpha\beta}$	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 are 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 are also from PSID covering 1967-1992, and the results are from table 4.

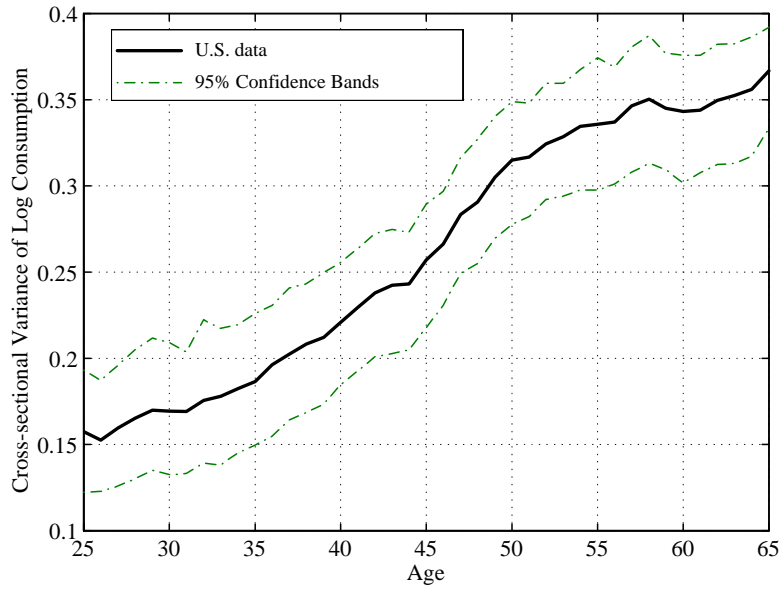Figure 1: Age-Inequality Profile of Consumption in the U.S. Data

Figure 2: Average Life-Cycle Profile of Consumption by Education in the U.S. Data
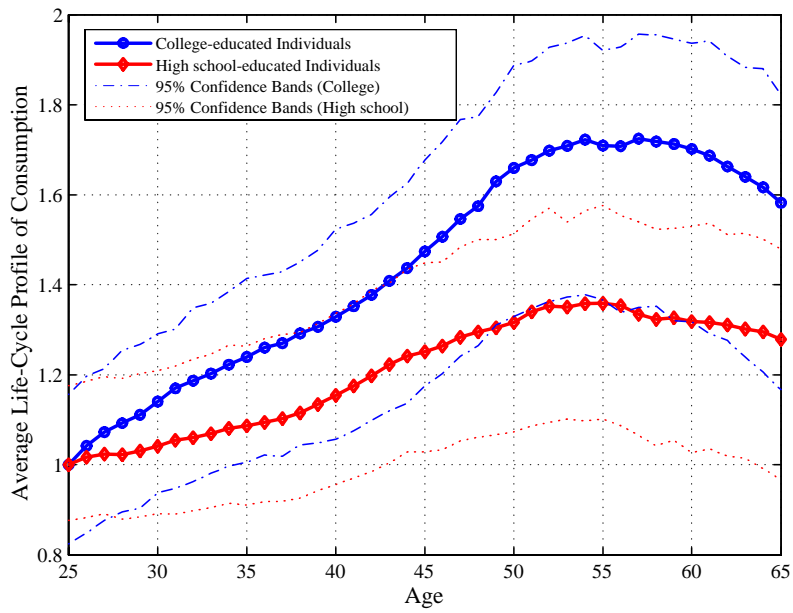
Figure 3: The Speed of Resolution of Income Uncertainty Through Bayesian Learning About Profiles
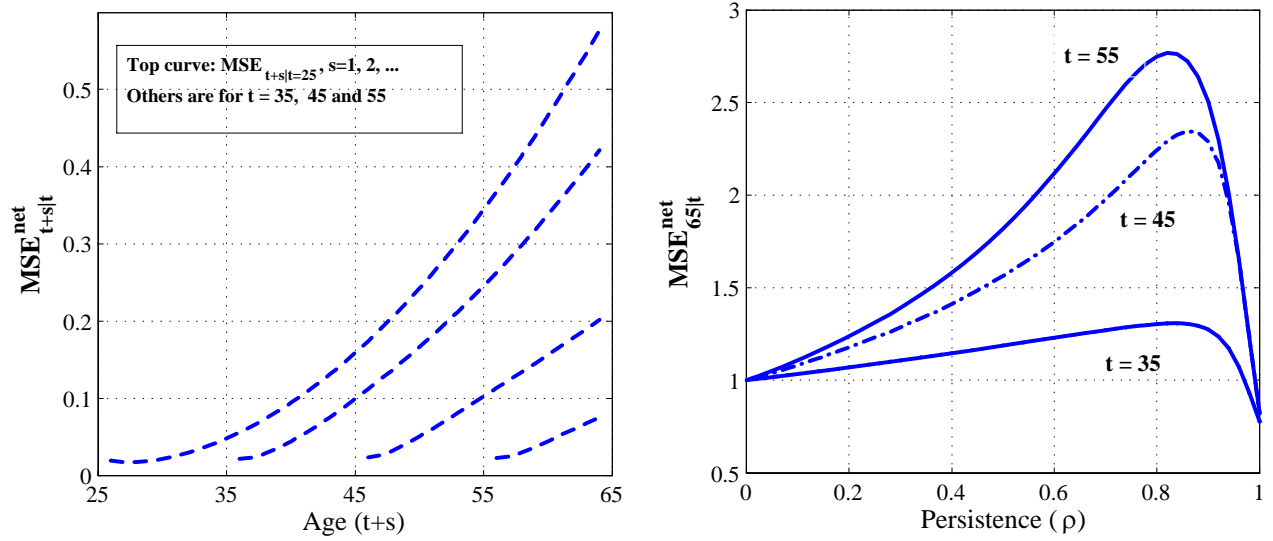

Figure 4: The Change in the Precision of Beliefs About α and β
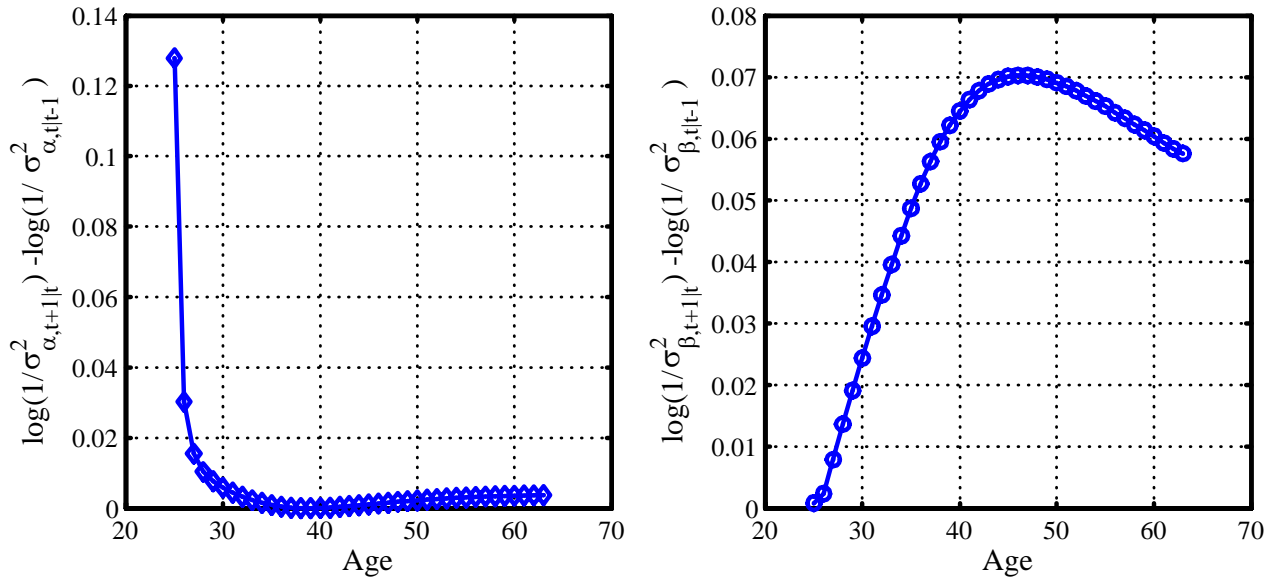
Figure 5: The Change in the Precision of Beliefs About α When $\sigma^2_{\alpha,0}$ Is Set to 10 Times Its Baseline Value
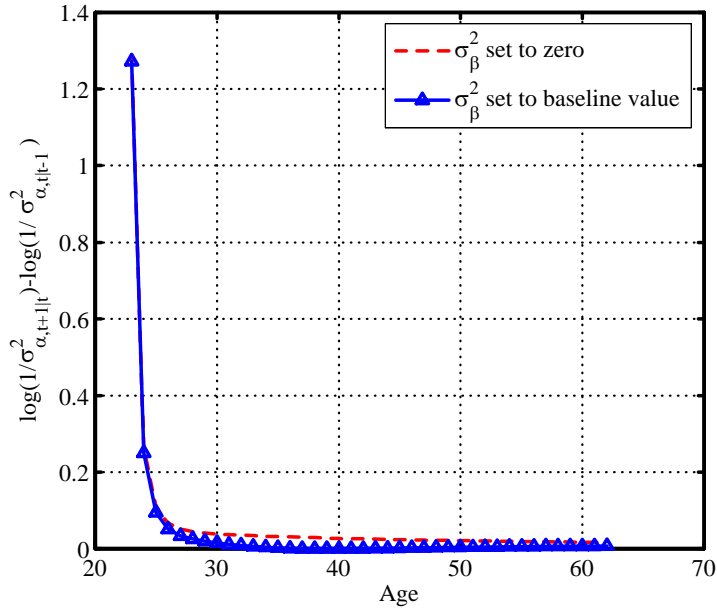

Figure 6: Resolution of Uncertainty When the Volatility of z_t Follows a GARCH Process
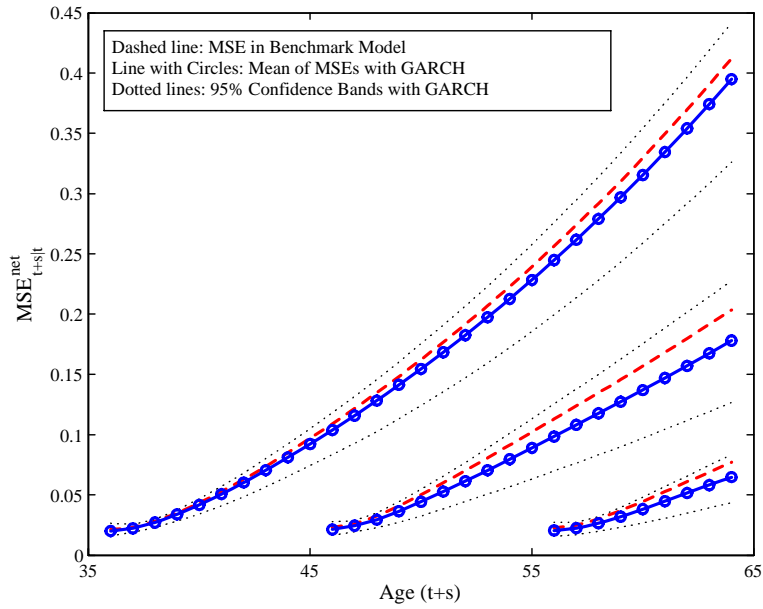
Figure 7: Age-Inequality Profile of Consumption: U.S. Data Versus HIP Model
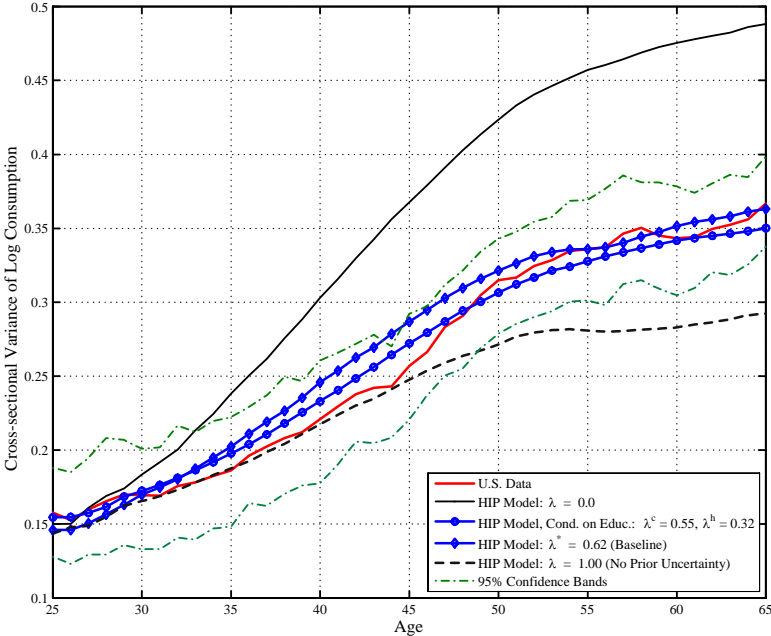
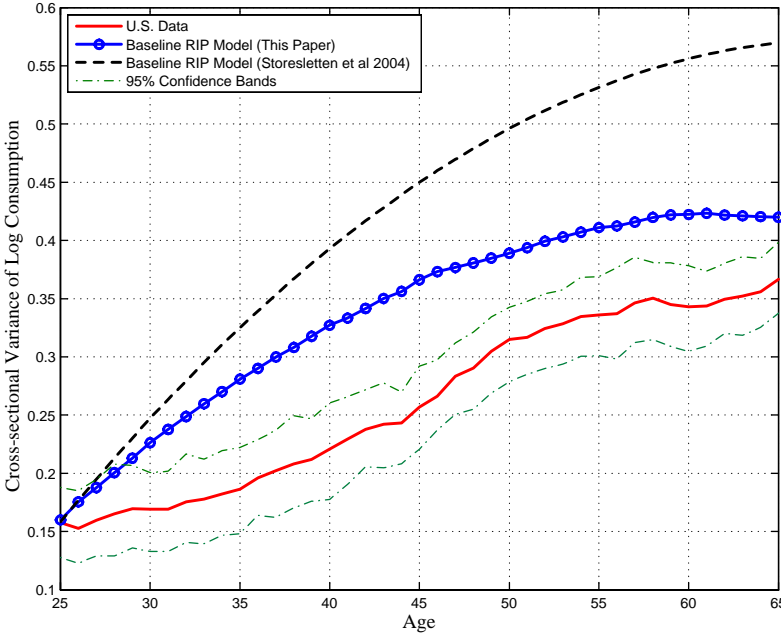Figure 8: Age-Inequality Profile of Consumption: U.S. Data Versus RIP Model

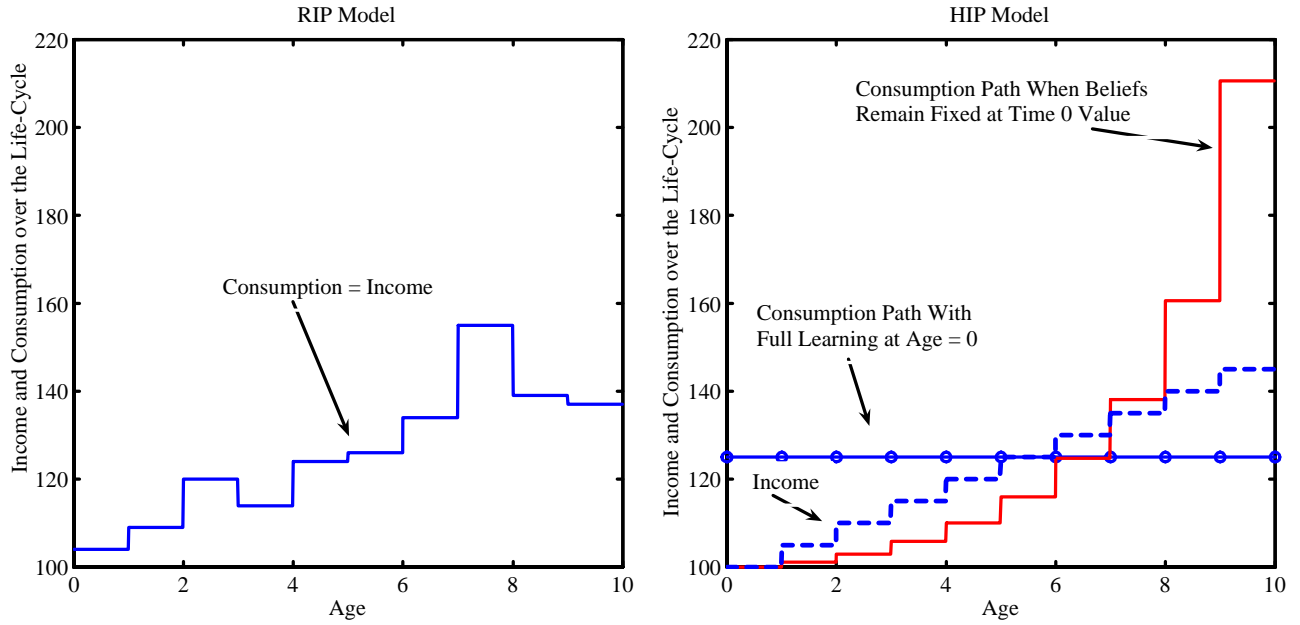Figure 9: Income and Consumption Movements in the RIP and HIP Models

Figure 10: Age-Inequality Profile of Consumption by Education: U.S. Data and Group-Specific HIP Models
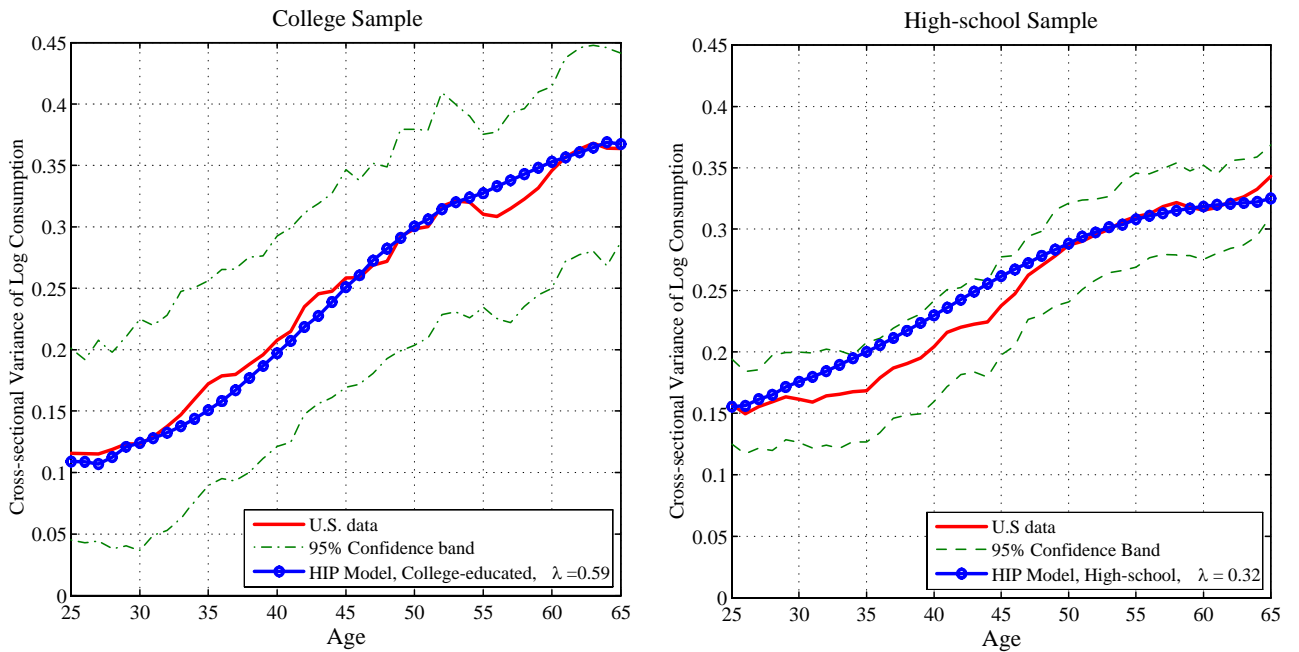
Figure 11: Average Life-Cycle Profile of Consumption by Education: HIP model versus RIP model