

Federal Reserve Bank of Minneapolis
Research Department Staff Report 485

June 2013

Inferring Labor Income Risk and Partial Insurance from Economic Choices*

Fatih Guvenen

University of Minnesota,
Federal Reserve Bank of Minneapolis,
and NBER

Anthony Smith

Yale University and NBER

ABSTRACT

This paper uses the information contained in the joint dynamics of individuals' labor earnings and consumption-choice decisions to quantify both the amount of income risk that individuals face and the extent to which they have access to informal insurance against this risk. We accomplish this task by using indirect inference to estimate a structural consumption-savings model, in which individuals both learn about the nature of their income process and partly insure shocks via informal mechanisms. In this framework, we estimate (i) the degree of partial insurance, (ii) the extent of systematic differences in income growth rates, (iii) the precision with which individuals know their own income growth rates when they begin their working lives, (iv) the persistence of typical labor income shocks, (v) the tightness of borrowing constraints, and (vi) the amount of measurement error in the data. In implementing indirect inference, we find that an auxiliary model that approximates the true structural equations of the model (which are not estimable) works very well, with negligible small sample bias. The main substantive findings are that income shocks are not very persistent, systematic differences in income growth rates are large, individuals have substantial amounts of information about their income growth rates, and about one-half of income shocks are effectively smoothed via partial insurance. Putting these findings together, we argue that the amount of uninsurable lifetime income risk that individuals perceive is substantially smaller than what is typically assumed in calibrated macroeconomic models with incomplete markets.

Keywords: Labor income risk; Idiosyncratic shocks; Indirect Inference Estimation; Heterogeneous income profiles; Persistence

JEL classification: C33, D81, D91, E21

*Guvenen: guvenen@umn.edu; <https://sites.google.com/site/fatihguvenen/>. Smith: tony.smith@yale.edu; <http://www.econ.yale.edu/smith>. For helpful critiques and suggestions, we thank our discussants at various conferences: Victor Aguirregabiria, Richard Blundell, Stephane Bonhomme, Jonathan Parker, Luigi Pistaferri, Sam Schulhofer-Wohl, Kjetil Storesletten, and Harald Uhlig, as well as Daron Acemoglu, Orazio Attanasio, Martin Browning, Raj Chetty, Lars Hansen, Yuichi Kitamura, Jean-Marc Robin, and Victor Rios-Rull. We also thank seminar participants at Chicago FED, IAS-Barcelona, Iowa, Georgetown, HEC Lausanne, Michigan, NYU, Princeton, Oxford, Seoul National, Sogang, Stanford, Tokyo, Toronto, UCL, UIUC, Warwick, and the conference participants at the NBER Economic Fluctuations and Growth meetings in Cambridge, Minnesota Workshop on Macroeconomic Theory, Banque de France/Bundesbank Spring Conference, the SED Conference, Paris Workshop on Heterogeneity and Macrodynamics, IFS/UCL Workshop on Consumption Inequality and Income Dynamics, the 5th International PIER/IGIER Conference at Upenn, Advances in Theory Based Estimation Conference in Honor of Hansen and Singleton, ESSIM Conference, 2008 Research on Money and Markets Conference, JAE Conference on Distributional Dynamics at CEMFI-Madrid, Conference on Structural Models of the Labour Market at IFS, 2009 Cowles Summer Macroeconomics Conference, First Shanghai Macroeconomics Conference, and the 2008 AEA Winter meetings for comments. Matthew Johnson, Marina Tavares, Georgios Stefanides, and Simone Civale provided excellent research assistance. Guvenen acknowledges financial support from the NSF under grant number SES-0649437. All remaining errors are our own. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

The goal of this paper is to use individuals' consumption-savings decisions to learn about the uninsurable labor income risks that they face. Income fluctuations alone, even in the absence of measurement problems, overstate this risk because they do not reveal whether individuals anticipate these fluctuations or can use informal mechanisms, beyond self-insurance through borrowing and lending, to insure against these fluctuations. Instead, we use the joint dynamics of individuals' labor earnings and consumption choices to infer both what they know about the income risks they face and how well they can use informal mechanisms to insure these risks.

Specifically, we build a life-cycle consumption-savings model with CRRA utility, potentially binding borrowing constraints, partial insurance, and a realistic retirement pension system. We assume that the slopes of individuals' labor income profiles (i.e., their income growth rates) vary in the population but that individuals have imperfect information about their own growth rates. Each individual enters the labor market with a prior belief about his own growth rate and then updates his beliefs over time in a Bayesian fashion. A key parameter is the precision of the initial prior belief: this parameter determines the extent to which an individual has advance information about the slope of his future income path.¹ In addition, we assume that some part of the surprise (or shock) to an individuals' labor income can be insured via informal mechanisms that we do not model explicitly. A key parameter here is the fraction of this surprise that can be insured. Finally, we also allow for several types of measurement error, a pervasive feature in individual-level data on income and consumption. Along all three of these dimensions—the amount of advance information about income growth rates, the extent of partial insurance, and the size of measurement error—consumption-choice decisions play a critical role because data on income alone cannot identify any of them.

We use a simulation method, indirect inference, to estimate these key parameters as well as the tightness of borrowing constraints, the discount factor, and the parameters governing labor-income dynamics. Rather than select an arbitrary set of unconditional moments upon which to base estimation (as is usually done in the method of simulated moments), indirect inference focusses instead on the parameters of an auxiliary model that plays the role of a reduced form for the structural model. In particular, we use an auxiliary model that approximates the joint dynamics of income and consumption implied by the structural consumption-savings model. The indirect inference estimator chooses the values of the structural parameters so that the parameters of this auxiliary model, estimated using either the observed data or data simulated from the structural model, are as close as possible. In effect, the indirect inference estimator seeks to find the best-fitting set of auxiliary-model parameters subject to the cross-equation restrictions that the structural model places on these parameters.²

¹The modeling of this learning process builds on Guvenen (2007). As we discuss in the next section, however, the analysis in the present paper differs in several important ways from the one in that paper.

²Thus our estimation methodology follows in the spirit of Sargent (1978) who uses aggregate rather than micro data to estimate a “rational expectations version of Friedman’s time-series consumption model . . . by imposing the pertinent restrictions across the stochastic processes for consumption and income.”

The key findings of our analysis regarding the nature of income risk are that: (i) informal mechanisms insure about one-half of a given income surprise; (ii) systematic differences in income growth rates are large; (iii) individuals have substantial amounts of information about their future income prospects; and (iv) the typical income shock is not very persistent. Taken together, these findings deliver the main substantive conclusion of this paper: The amount of uninsurable future income risk perceived by individuals upon entering the labor market is significantly smaller than what is typically assumed in calibrated macroeconomic models with incomplete markets.

The next section describes a linear-quadratic version of the consumption-savings problem that permits an analytical solution, thereby allowing us to characterize theoretically the information content in consumption choices. Section 3 then lays out the full model used in estimation. Sections 4 and 5 describe the data and the indirect inference methodology. Section 6 presents the results and extensions. Section 7 concludes.

1.1 Literature Review

This paper is related to a growing literature that uses panel data to study the transmission of income shocks to consumption when markets are incomplete. Important examples include Hall and Mishkin (1982) and, more recently, Blundell et al. (2008), Kauffmann and Pistaferri (2009), Krueger and Perri (2009), and Kaplan and Violante (2010). This paper contributes to this literature in the following ways. First, we estimate the amount of partial insurance in a model where the underlying income risk can have a richer structure than what is considered in these papers. In particular, we allow for growth-rate heterogeneity and Bayesian learning, as well as persistent shocks that are not restricted to follow a random walk. This structure allows us to provide a more comprehensive picture of the sources of income uncertainty (arising either from genuine shocks or from learning), as well as the extent of insurance against this background. Second, these papers derive estimable equations first and then impose partial insurance on top of these equations. Instead, in this paper partial insurance is modeled as a transfer in the budget constraint and the implications for the consumption-savings choice are worked out explicitly. Kaplan and Violante (2010) shares some similarities with our paper: this paper incorporates non-permanent shocks and retirement into the model of Blundell et al. (2008) but does not allow for growth-rate heterogeneity with Bayesian learning.³

Although the framework in this paper shares some common elements with that in Guvenen (2007), that work and the present paper have at least three important differences. First, one major goal of the present paper is to distinguish between information about future earnings prospects and partial insurance, whereas Guvenen (2007) restricts insurance opportunities to self-insurance only. Second,

³A related set of papers includes Deaton and Paxson (1994), Blundell and Preston (1998) and, more recently, Primiceri and van Rens (2009). These papers use structural models similar to those above but employ repeated cross sections of consumption and income, rather than a panel. Furthermore, they assume permanent shocks and rule out profile heterogeneity. Primiceri and van Rens (2009) is a partial exception to this statement in that their framework potentially allows for systematic patterns in income growth, although this is conditional on shocks being permanent. They also do not estimate how much individuals know about their own growth rates.

Guvenen (2007) takes all of the parameters of the income process as given (estimated in another paper from income data alone) and uses the consumption data to calibrate the value of one parameter: the amount of prior information regarding one’s own income growth rate. In contrast, this paper brings consumption data to bear on the estimation of the entire vector of structural parameters, which contains all of the parameters of the income process, prior beliefs, preferences, and borrowing constraints, as well as several types of measurement error. Third, that paper focuses exclusively on the life-cycle mean and variance profiles of consumption (using repeated cross-sections); here, we use the joint dynamics of consumption and income using panel data to conduct a formal structural estimation. In Sections 2.2 and 6.3, we show how critical these joint dynamics are for identifying parameters, especially those pertaining to partial insurance. This more thoroughgoing analysis leads us to conclude that the amount of prior information about income growth is quite a bit larger than what Guvenen (2007) gleans from life-cycle profiles. This larger amount of prior information, in combination with a significant amount of partial insurance, leads us to conclude that the amount of uncertainty perceived by a 25-year-old individual about his future income at age 55 is about *one-third* of what was found in that paper.⁴

From a methodological perspective, an important precursor to this paper is Gourinchas and Parker (2002), which estimates a life-cycle consumption-savings model using the method of simulated moments. Their main focus is on whether such a model can explain the hump-shaped consumption profile over the life cycle. These authors estimate the income process parameters from income data first and, in a second step, estimate risk aversion and the time discount factor from consumption data, whereas we estimate all parameters using both data sources jointly. Finally, there is a small but growing literature that uses indirect inference to estimate structural economic models in a variety of fields, including labor economics, finance, macroeconomics, and industrial organization; a non-exhaustive list includes Smith (1993), Magnac et al. (1995), Bansal et al. (2007), Nagypal (2007), Li (2010), Low and Pistaferri (2012), and Collard-Wexler (2013). Among these, our philosophy for selecting the auxiliary model is in the same spirit as Magnac et al. (1995), who use an auxiliary model that mimics the true structural equations of their model. To our knowledge, this paper is the first to use indirect inference to estimate a fully-specified consumption-savings model.

⁴Our paper is also in the same spirit as Carneiro et al. (2003) and Cunha et al. (2005), who use the college choice to measure the amount of information youths have about their future income prospects. The latter paper finds that about 60% of the cross-sectional dispersion of labor income is predictable by age 18. Our estimate using full consumption dynamics over the life cycle—88% by age 25—is consistent with this finding. Both findings suggest that young individuals have substantial information about their future prospects in the labor market.

2 A Framework for Inferring Income Risk

Let the log labor income of individual i with t years of labor market experience be given by

$$y_t^i = \underbrace{g(t, \text{observables}, \dots)}_{\text{common life cycle component}} + \underbrace{[\alpha^i + \beta^i t]}_{\text{profile heterogeneity}} + \underbrace{[z_t^i + \varepsilon_t^i]}_{\text{stochastic component}}, \quad (1)$$

where $z_t^i = \rho z_{t-1}^i + \eta_t^i$, and η_t^i and ε_t^i are zero mean innovations that are i.i.d. over time and across individuals. The terms in the first set of brackets capture the lifecycle variation in labor income that is common to all individuals with given observable characteristics. The second component captures potential *individual-specific* differences in income growth rates (as well as in levels). Such differences would be implied, for example, by a human capital model with heterogeneity in the ability to accumulate skills.⁵ Finally, the terms in the last bracket represent the stochastic variation in income, which is modeled here as the sum of an AR(1) component and a purely transitory shock. This specification encompasses (or differs only in minor ways from) the processes estimated in the literature.

2.1 Bayesian Learning about Income Profiles

We begin by laying out a framework that allows for various possibilities regarding individuals' perceptions of their future income risk. For example, an individual is likely to have more information than the econometrician about his β^i at the time he enters the labor market and will update those beliefs as information is revealed in his income realizations. We model this process by assuming that an individual enters the labor market with some prior belief about his own β^i , which is then refined over time in a Bayesian fashion (following Guvenen (2007)). The prior variance of this belief distribution measures how uncertain individuals are about their own β^i when they enter the labor market and is therefore a key parameter for determining the amount of perceived income risk.

Time 0: Prior Beliefs and Variance (Uncertainty). Imagine that for each individual, nature draws two random variables at time zero, β_k^i and β_u^i , which are independent of each other, with zero mean and variances of $\sigma_{\beta_k}^2$ and $\sigma_{\beta_u}^2$. The income growth rate is given by $\beta^i = \beta_k^i + \beta_u^i$, implying $\sigma_{\beta}^2 = \sigma_{\beta_k}^2 + \sigma_{\beta_u}^2$. The key distinction between the two components is that individual i observes the realization of β_k^i , but not of β_u^i (hence, the subscripts indicate *known* and *unknown*, respectively). Then, the prior mean is $\widehat{\beta}_{1|0}^i = \beta_k^i$, and the prior variance is $\sigma_{\beta,0}^2 = \sigma_{\beta_u}^2$. To express the amount of prior uncertainty in relation to the heterogeneity in income growth rates, it is useful to define

$$\lambda \equiv \frac{\sigma_{\beta,0}}{\sigma_{\beta}},$$

⁵See, for example, the classic paper by Ben-Porath (1967). For more recent examples of such a human capital model, see Guvenen and Kuruscu (2010) and Huggett et al. (2011).

which is simply the fraction of the population dispersion of income growth rates that represents uncertainty on the part of individuals at the time they enter the labor market. Two polar cases deserve special attention. When $\lambda = 1$, individuals do not have any *private prior information* about their income growth rate (i.e., $\sigma_{\beta,0}^2 = \sigma_{\beta}^2$ and $\widehat{\beta}_{1|0}^i = \bar{\beta}$ for all i , where $\bar{\beta}$ is the population average). At the other extreme, when $\lambda = 0$, each individual observes β^i completely and faces no prior uncertainty about its value.⁶

Updating Beliefs over the Life Cycle. We now cast the optimal learning process as a Kalman filtering problem, which yields convenient recursive updating formulas for beliefs. Each individual knows his own α^i , observes his income and the transitory shock, y_t^i and ε_t^i , and must learn about $\mathbf{S}_t^i \equiv (\beta^i, z_t^i)$.⁷ The “state equation” describes the evolution of the vector of state variables that is unobserved by the decision maker:

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

A second (observation) equation expresses the observable variable in the model—in this case, log income net of the fixed effect and transitory shock (\tilde{y}_t^i)—as a linear function of the underlying hidden state:

$$\tilde{y}_t^i \equiv y_t^i - \alpha^i - \varepsilon_t^i = \begin{bmatrix} t & 1 \end{bmatrix} \begin{bmatrix} \beta^i \\ z_t^i \end{bmatrix} = \mathbf{H}_t' \mathbf{S}_t^i.$$

Both innovations have i.i.d. normal distributions and are independent of each other, with \mathbf{Q} and R denoting the covariance matrix of ν_t^i and the variance of ε_t^i , respectively. Each individual’s prior belief over (β^i, z_1^i) is represented by a multivariate normal distribution with mean $\widehat{\mathbf{S}}_{1|0}^i \equiv (\widehat{\beta}_{1|0}^i, \widehat{z}_{1|0}^i)$ and covariance matrix

$$P_{1|0} = \begin{bmatrix} \sigma_{\beta,0}^2 & 0 \\ 0 & \sigma_{z,0}^2 \end{bmatrix},$$

where we use the shorthand notation $\sigma_{\cdot,t}^2$ to denote $\sigma_{\cdot,t+1|t}^2$. After observing $(\tilde{y}_t^i, \tilde{y}_{t-1}^i, \dots, \tilde{y}_1^i)$, the posterior belief about \mathbf{S}_t^i is Normally distributed with a mean vector $\widehat{\mathbf{S}}_t^i$ and covariance matrix \mathbf{P}_t . Define the *perceived innovation* to (log) income as

$$\widehat{\xi}_t^i \equiv \tilde{y}_t^i - \mathbb{E}_{t-1}(\tilde{y}_t^i) = \tilde{y}_t^i - \left(\widehat{\beta}_{t|t-1}^i t + \widehat{z}_{t|t-1}^i \right), \quad (2)$$

⁶Notice that since each household knows its α^i , this already gives some information about its β^i as long as the two parameters are correlated. It is natural to think of λ as already incorporating this information. One way to think about this is that β_k^i captures all the correlation between α^i and β^i , and $\beta_u^i \perp \alpha^i$. We can easily show that there is an upper bound to λ (which depends on $\sigma_{\alpha}^2, \sigma_{\beta}^2$, and $\sigma_{\alpha\beta}$) that captures this minimum information obtained from α^i alone, and this upper bound is 1 when $\text{corr}(\alpha, \beta) = 0$ and is 0 when $\text{corr}(\alpha, \beta) = \pm 1$.

⁷The assumption that α^i is observable is fairly innocuous here because the uncertainty regarding this parameter is resolved very quickly even when the individual has substantial prior uncertainty, as shown in Guvenen (2007). The assumption that ε is observable will turn out not to matter (see note 12) and is made here for expositional reasons.

which does not necessarily have the same sign as the true innovation to income, η_t^i —a point that will play a crucial role below. The recursive Kalman updating formulas are given in Appendix A.

2.2 Understanding Identification: A Stylized Linear-Quadratic Framework

Before delving into the details of the full estimation, it is useful to provide a better understanding of the sources of identification. For example, if individuals indeed differ in their income growth rates (i.e., $\sigma_\beta \gg 0$), how would this fact be revealed in their consumption-choice behavior? Similarly, can we detect the extent of an individual’s prior uncertainty (that is, λ) about his β^i by observing the response of his consumption to income movements? And, finally, what kind of empirical relationship would allow us to measure the degree of partial insurance, θ , in the presence of these other features?

For this purpose, we begin with a very stylized life cycle model of the consumption-savings decision. Specifically, (i) individuals have quadratic utility over consumption, (ii) the time discount factor, δ , is the reciprocal of the gross interest rate, $1 + r$, (iii) there are no borrowing constraints, and (iv) there is no retirement. Finally, we assume a simpler form of the income process in (1):

$$Y_t^i = \alpha^i + \beta^i t + z_t^i, \quad (3)$$

where the income *level* (instead of its logarithm) is linear in the underlying components, and we set $\varepsilon_t^i \equiv 0$.⁸ Under these assumptions, the consumption-savings problem can be written as

$$V_t^i(\omega_t^i, \hat{\beta}_t^i, \hat{z}_t^i) = \max_{C_t^i, a_{t+1}^i} \left\{ -(C_t - C^*)^2 + \frac{1}{1+r} \mathbb{E}_t \left[V_{t+1}^i(\omega_{t+1}^i, \hat{\beta}_{t+1}^i, \hat{z}_{t+1}^i) \right] \right\} \quad (4)$$

$$\text{s.t.} \quad C_t^i + a_{t+1}^i = \omega_t^i, \quad (5)$$

$$\omega_t^i = (1+r) a_t^i + Y_t^{\text{disp},i}, \quad (\text{cash on hand}) \quad (6)$$

and the Kalman recursions (27), (28), (29), and (30) given in Appendix A. Disposable income is $Y_t^{\text{disp},i} = Y_t^i - \theta \hat{\xi}_t^i$, where θ is the partial insurance parameter, which measures the fraction of the perceived income innovation, $\hat{\xi}_t^i$, that is insured. So, for example, when the realization of Y_t^i is lower than what was expected in the previous period, $\hat{\xi}_t^i$ will be negative (by definition), and disposable income will be higher than labor income, thanks to the positive partial insurance term: $-\theta \hat{\xi}_t^i > 0$.

The framework described above is a much simplified version of the full model we estimate in Section 5.3. It is, however, general enough that it encompasses three special cases of interest that we will refer to in the rest of the paper. First, without any further restrictions imposed, the framework has a heterogeneous income profiles process with Bayesian learning about individual income slopes—i.e., the HIP model. A second important benchmark is obtained when $\sigma_\beta \equiv 0$, in which case there is no heterogeneity in profiles and no Bayesian learning. Thus, the framework reduces to the standard

⁸Closed-form solutions such as those below can still be derived in the presence of transitory shocks and retirement. We abstract from them here only for the clarity of exposition.

RIP (restricted income profiles) model, extensively studied in the literature. Finally, a third case of interest is when $\sigma_\beta > 0$ and $\lambda = 0$. In this case, individuals face a HIP process, but each individual knows his β^i at the time he enters the labor market. The only source of uncertainty in this case is the idiosyncratic shocks, as in the RIP model. This is an intermediate case between the HIP and RIP models in cases (1) and (2).

2.2.1 Information in Consumption Growth

For clarity of exposition, in this section we abstract from partial insurance by setting $\theta \equiv 0$. In the next section, we reintroduce it and study its implications. We also show that the presence of partial insurance has no effect on the substantive conclusions of the current section.

For the problem described by equations (3) to (6) with $\theta = 0$, optimal consumption choice satisfies

$$\Delta C_t = \varphi_t \left[\sum_{s=0}^{T-t} \gamma^s (\mathbb{E}_t - \mathbb{E}_{t-1}) Y_{t+s} \right], \quad (7)$$

where $\gamma \equiv 1/(1+r)$ and $\varphi_t \equiv (1-\gamma)/(1-\gamma^{T-t+1})$ is the annuitization factor. After substituting (3) in (7), some tedious but straightforward manipulations yield

$$\Delta C_t = \Phi(t; T, r) \left(\widehat{\beta}_t^i - \widehat{\beta}_{t-1}^i \right) + \Psi(t; T, \rho, r) (\widehat{z}_t^i - \rho \widehat{z}_{t-1}^i), \quad (8)$$

where Φ_t and Ψ_t are some age-dependent positive coefficients.⁹ (In what follows, for clarity, we suppress all the arguments of Φ and Ψ except t , unless it creates confusion.) Finally, substituting (27) and (28) into (8) yields a key structural equation in this framework:

$$\Delta C_t = \Pi_t \times \widehat{\xi}_t^i, \quad (9)$$

where Π_t is the response coefficient (full formula in Appendix A) and $\widehat{\xi}_t^i$ is now reinterpreted as the innovation in the *level* of income—i.e., $Y_t^i - \left(\widehat{\beta}_{t|t-1}^i t + \widehat{z}_{t|t-1}^i \right)$. This equation basically says that consumption changes proportionally to the perceived innovation to income, which, as we shall see, may or may not have the same sign as η_t .

If we eliminate profile heterogeneity by setting $\sigma_\beta \equiv 0$ (and thereby also eliminate learning), the resulting (RIP) model implies

$$\Delta C_t = \Psi_t \times \eta_t^i. \quad (10)$$

The last two equations can be used to understand some of the key differences between the two frameworks.¹⁰ When $\sigma_\beta \equiv 0$, only the current shock, η_t^i , matters for consumption response, whereas in

⁹Full formulas are provided in Appendix A.

¹⁰Before moving further, it is important to stress that equation (8) is obtained by fully solving the consumption-savings model and therefore requires (i) using the Euler equation, (ii) imposing the budget constraint, and (iii) specifying a

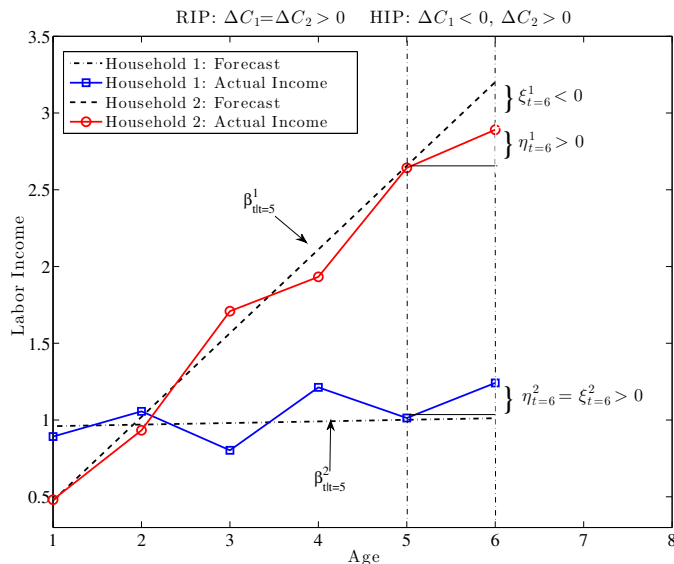


Figure 1: Information about σ_β and λ in Consumption Changes

the HIP model the entire history of shocks matters—through beliefs. As a result, two individuals hit by the same η_t^i may react differently depending on their history. In the rest of this section, we present a series of examples that help explain the intuition behind the identification of two key parameters— σ_β and λ . For this purpose, we specialize to the case where $\rho = 1$, which makes the exposition much simpler (although the main conclusions we reach below hold regardless of the value of ρ). We also assume $\bar{\beta} = 0$, again, without loss of generality and for clarity of exposition.

Example 1 (Consumption Growth Depends Negatively on Past Income Growth). Consider Figure 1, which plots the income paths of two individuals up to period 6. Individual 1 experiences a faster average income growth rate in the first five periods than individual 2, but observes precisely the same rise in income between periods 5 and 6 ($\Delta Y_6^1 = \Delta Y_6^2$). If these income paths are generated by a process with $\sigma_\beta = 0$ (and thus $\beta^1 = \beta^2$), then the consumption choice of both individuals will satisfy equation (10), implying $\Delta C_6^1 = \Delta C_6^2 = \eta_6 > 0$. Instead, if the true income process has $\sigma_\beta > 0$ (HIP process), individual 1 will have formed a belief that his income growth rate is higher than that of individual 2 and will forecast his income to be on the (dashed-blue) trend line. (Obviously, this remains true when $\lambda = 0$ and so when beliefs are perfectly accurate from the beginning.) Therefore, even though his income increases, it is below the expected trend ($\hat{\xi}_6^1 < 0$), which causes him to revise down his beliefs about β^1 and, consequently, reduce his consumption level, from equation (9). In contrast, based on his past income growth (which is nearly zero), individual 2 is positively surprised

stochastic process for income. In this sense, the analysis here is in the spirit of Hall and Mishkin (1982) (and, more recently, Blundell et al. (2008)), who derive the full consumption function (as we do here), rather than Hall (1978) who requires only the Euler equation to hold. Therefore, by imposing stronger restrictions, the current approach allows us to estimate the parameters of the income process in addition to the preference parameters, which is all one can estimate with the Euler equation approach.

to see any increase in his income between periods 5 and 6 ($\hat{\xi}_6^2 > 0$) and will increase his consumption. Thus, we have $\Delta C_6^1 > 0$ and $\Delta C_6^2 < 0$ in response to the same income change for both agents.

The following proposition summarizes these implications in a more rigorous form. First, define $\Delta \bar{C}_t^i \equiv \mathbb{E}(\Delta C_t^i | \beta^i, \Delta Y_t^i)$. In words, $\Delta \bar{C}_t^i$ is the average consumption growth of an individual with slope parameter β^i and income growth ΔY_t^i between $t - 1$ and t , where the expectation is taken over all possible histories up to $t - 1$ (of initial signals and income realizations: $\{Y_1^i, Y_2^i, \dots, Y_{t-1}^i, \hat{\beta}_{1|0}^i\}$). In other words, $\Delta \bar{C}_t^i$ can be thought of as the consumption change of a *typical* individual with β^i and who observed ΔY_t^i .

Proposition 1 [*Information in Consumption Growth*] *In the lifecycle certainty-equivalent model with permanent shocks described above, consumption growth satisfies the following properties:*

- (i) *Controlling for current income growth, consumption growth will, on average, be a decreasing function of an individual's β^i : i.e., $\frac{\partial \Delta \bar{C}_t^i}{\partial \beta^i} < 0$ for all t .*
- (ii) *Although the prediction in (i) is true for all values of λ (including when $\lambda = 0$ and, hence, when there is no learning), the relationship becomes stronger as λ rises: i.e., $\frac{\partial^2 \Delta \bar{C}_t^i}{\partial \beta^i \partial \lambda} < 0$ for all t .*
- (iii) *Similarly (holding everything else constant), the response of consumption growth to income growth becomes stronger as λ increases: $\frac{\partial^2 \Delta \bar{C}_t^i}{\partial \Delta Y_t^i \partial \lambda} > 0$.*

Proof. See Appendix A for all of the omitted proofs and derivations.

To understand the empirical content of the proposition, note that even though β^i is not observed by the econometrician, in any given period, past and future income growth rates will be positively correlated with β^i . Therefore, the empirical relationship predicted by part (i) is that controlling for ΔY_t^i , consumption growth will be a decreasing function of the past income growth rate, which is observable by the econometrician. The second part of the proposition then implies that this negative dependence on the past income growth becomes stronger as individuals receive signals that are less informative at the beginning of life (i.e., a higher λ). Similarly, part (iii) implies that the response coefficient of consumption growth to income growth contains valuable information about the initial prior uncertainty faced by individuals. Loosely speaking, this is because when the initial signal is not very informative, optimal learning will result in a larger updating of beliefs about β^i in response to a given income realization, ΔY_t^i , which will in turn cause a larger change in consumption. Finally, it is easy to see that when $\sigma_\beta \equiv 0$, consumption growth will not depend on an individual's past or future income growth. As we shall see in the next section, we will use these observations to write an auxiliary model that captures the way in which consumption growth depends on past and future income growth rates, as well as how it responds to contemporaneous income growth, to infer the values of σ_β and λ as well as other parameters.

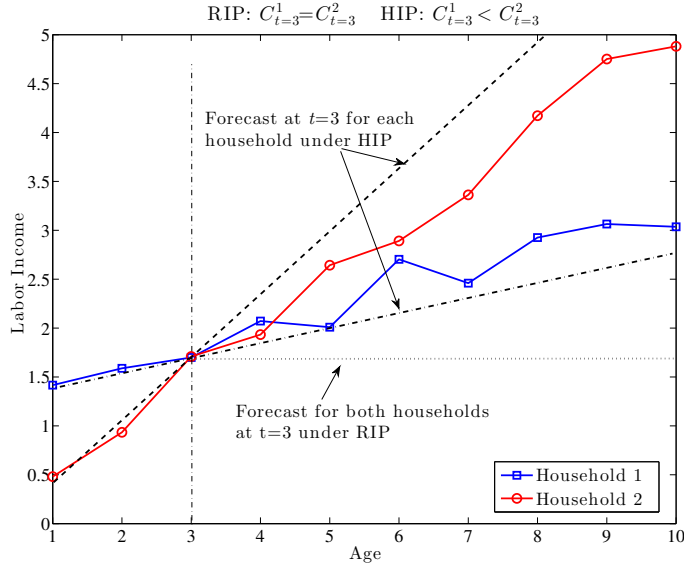


Figure 2: Information about σ_β in Consumption Levels

2.2.2 Information in Consumption Levels

We next turn to the information revealed in the levels of consumption and begin with the following useful lemma. (See Appendix A for the proofs.)

Lemma 1 *The consumption decision rule can be solved in closed form as a linear function of the state vector $(\omega_t^i, \hat{\beta}_t^i, \hat{z}_t^i)$:*

$$C_t^i = \varphi_t \omega_t^i + \gamma \Phi_{t+1} \hat{\beta}_t^i + \gamma \rho \Psi_{t+1} \hat{z}_t^i. \quad (11)$$

This expression clearly shows that, at every point in time, consumption choice reveals valuable information about individuals' perceived future income prospects as reflected in $(\hat{\beta}_t^i, \hat{z}_t^i)$. We now state the key result of this section and then present two examples to illustrate how this information can be used.

Proposition 2 [Information in Consumption Levels] *Controlling for an individual's current income and assets, the consumption level is an increasing function of the individual's beliefs about his income growth rate, $\hat{\beta}_t^i$. This prediction holds true regardless of how much individuals know about their true income growth rate—i.e., for all $\lambda \in [0, 1]$. However, if $\sigma_\beta = 0$, consumption depends only on current income and assets.*

Proof. We present the proof in a way that is helpful for understanding the two examples that follow. Consider two individuals with $Y_t^1 = Y_t^2$, $A_t^1 = A_t^2$, and $\hat{\beta}_t^1 > \hat{\beta}_t^2$ as hypothesized in the proposition. Then we have

$$0 = Y_t^1 - Y_t^2 = (\hat{\beta}_t^1 - \hat{\beta}_t^2)t + (\hat{z}_t^1 - \hat{z}_t^2) \Rightarrow (\hat{\beta}_t^1 - \hat{\beta}_t^2)t = -(\hat{z}_t^1 - \hat{z}_t^2). \quad (12)$$

Taking the difference of the consumption level of each individual as given in equation (11) and using (12), we get

$$\begin{aligned} C_t^1 - C_t^2 &= \gamma\Phi_{t+1}(\hat{\beta}_t^1 - \hat{\beta}_t^2) + \gamma\rho\Psi_{t+1}(\hat{z}_t^1 - \hat{z}_t^2) \\ &= \gamma(\hat{\beta}_t^1 - \hat{\beta}_t^2)[\Phi_{t+1} - \rho t\Psi_{t+1}]. \end{aligned} \quad (13)$$

Since $\hat{\beta}_t^1 - \hat{\beta}_t^2 > 0$ by hypothesis, $C_t^1 - C_t^2 > 0$ if and only if $\Phi_{t+1} - \rho t\Psi_{t+1} > 0$. The proof of Lemma A.1 in Appendix A establishes that $\Phi_{t+1} - t\Psi_{t+1} > 0$, which straightforwardly implies that $\Phi_{t+1} - \rho t\Psi_{t+1} > 0$ as well, since $\rho \leq 1$. Thus, controlling for current income and assets, an individual's consumption is higher if his (perceived) income growth prospect, $\hat{\beta}_t^i$, is higher. In contrast, when $\sigma_\beta = 0$, we have $\beta^1 = \hat{\beta}_t^1 = \beta^2 = \hat{\beta}_t^2 = \bar{\beta}$, which in turn implies from (13) that the two individuals will have the same consumption levels. ■

To better understand the empirical content of the proposition, it is useful to study the following two examples. But first note that at any point in time, $\hat{\beta}_t^i$ depends on two things: (i) an individual's initial belief ($\hat{\beta}_{10}^i$) and (ii) the path of income realizations up to time t . Both examples consider two individuals with $\beta^1 > \beta^2$. In the first example, individuals enter the labor market with the same prior belief but experience different income paths consistent with their own β^i . The second example considers the opposite situation, with different priors but the same income history.

Example 2 (Past Income Growth Affects Current Consumption Level). Figure 2 plots a particular realization of income paths for two individuals. In this example, both individuals experience different growth rates up to period 3, but have $Y_3^1 = Y_3^2$. Now, if the true data-generating process has $\sigma_\beta = 0$ (i.e., $\beta^1 = \beta^2$) and individuals perceive it as such, then both individuals' forecasts of their future income would be the same: $\mathbb{E}_3(Y_{3+s}^1) = Y_3^1 = Y_3^2 = \mathbb{E}_3(Y_{3+s}^2)$, at all horizons $s \geq 0$ (shown in Figure 2 with the horizontal dotted line). Furthermore, if both individuals started life with no wealth, it is easy to see that $C_3^1 = Y_3^1 = Y_3^2 = C_3^2$. In contrast, when $\sigma_\beta > 0$, individuals know that they can have different β^i 's and will use the past income growth to form beliefs about their own β^i . Based on the high past income growth, individual 1 will expect a higher β^i and, therefore, a much higher lifetime income than individual 2. (And if $\lambda = 0$, then each will know his β^i with certainty from the beginning.) Therefore, the first individual will have a higher consumption *level* than individual 2 at the same age, despite having the same income level.

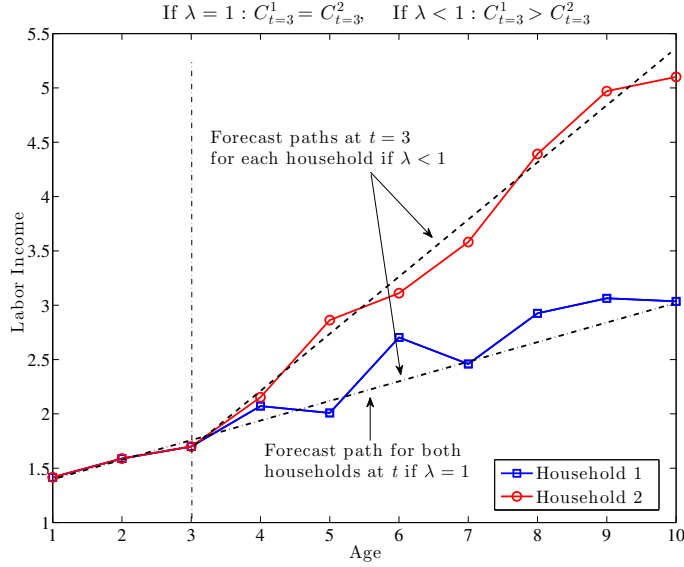


Figure 3: Information about Prior Uncertainty

Example 3 (Dependence of Consumption Level on Future Income Growth Reveals Prior Information). One can turn the same argument around to see how the level of consumption can also be informative about the degree of private information, λ . To show this, we turn to Figure 3, which is a slight variation of Figure 2. Consider two individuals with $\beta^1 > \beta^2$, who nevertheless experience the same income realizations up to period 3. Now if $\lambda = 1$, then both households will have the same beliefs in period 3 and, therefore, will choose the same consumption level. If, on the other hand, households have some prior information (i.e., $\lambda < 1$), the household that starts out with a higher prior belief ($\hat{\beta}_{1|0}^1 > \hat{\beta}_{1|0}^2$) will also have a higher belief at time t ($\hat{\beta}_3^1 > \hat{\beta}_3^2$) and therefore have a higher consumption level. This implication stands in contrast to the RIP model, which predicts no such dependence on past income levels (beyond what is captured by current income and assets). Moreover, since $\hat{\beta}_{1|0}^1$ is positively correlated with the true β^1 , and both households in this example observed the same *past* income paths, the household's prior belief will be correlated with his *future* income growth. Thus, controlling for current income and assets, and the past income path, the correlation between current consumption and future income growth also reveals how much prior information the household has. This is a useful prediction, as it shows how the observable variation in consumption and income can be used to infer the amount of prior information, which is unobservable.

These three examples illustrate how one can use the structural equations—such as (9), (10), and (11) that hold true exactly in a somewhat simplified version of the economic model to be estimated—to choose an auxiliary model. Indirect inference allows one to think in terms of these rich dynamic relationships instead of a set of moments (means, covariances, etc.). Below we shall write a parsimonious auxiliary model that captures these dynamic relationships to identify the key parameters of the income process.

The results of this section illustrate some important advantages of using the information revealed by intertemporal choices, such as consumption-savings, over using panel data on income alone. One difficulty of the latter approach is that identification between different income processes partly depends on the behavior of the *higher-order* autocovariances of income (Guvenen (2009) contains a detailed discussion of this point). In contrast, because of its forward-looking nature, even short-run movements in consumption and the immediate response of consumption to income innovations contain information about the perceived *long-run* behavior of the income process, as can be seen from (9) and (11).

2.2.3 Reintroducing Partial Insurance

We now remove the restriction on θ and summarize the implications of partial insurance in the following proposition.

Proposition 3 [Partial Insurance] *With partial insurance, optimal consumption growth is given by*

$$\Delta C_t = (\Pi_t - \theta\varphi_t) \times \widehat{\xi}_t^i. \quad (14)$$

The parameter θ can be identified from the age profile of the response of consumption to income surprises.

To understand the empirical content of the proposition, first notice that the effect of partial insurance on consumption is separable from the effects of learning (captured by Π_t). Thanks to this separability, all of the results established in the previous section (that is, Propositions 1 and 2 and Lemma 1) continue to hold in the presence of partial insurance with “disposable income” now playing the role of “income.” Second, for a sufficiently low interest rate and/or long enough horizon, the annuitization factor φ_t is nearly constant (especially up to age 55, which is the age range we will be focusing on), implying that the effect of partial insurance is flat over the life cycle. In contrast, the effect of learning (Π_t) is either monotonically increasing or inverse U-shaped depending on the parameterization. Therefore, the age patterns of the two effects are distinctly different in response to a given shock $\widehat{\xi}_t^i$. Third, it is clear that (as long as $\theta > 0$), the total response coefficient is now smaller than without partial insurance (but always positive), so, as expected, the response of consumption to income is muted. In Section 6.3, we examine sample paths for different households to illustrate the effects of partial insurance and how they depend on other variables, such as income level, wealth, age, and so on.

To summarize, the results of this section show that the dynamics of consumption and income contain rich information that can allow us to identify partial insurance as well as various aspects of the income process and households’ prior information about their future income growth prospects.

3 The Full Consumption-Savings Model

We now describe the full model estimated in the empirical analysis. Compared to the stylized framework in the previous section, the most significant change is that here we generalize preferences from quadratic utility, add a retirement period, and allow a richer specification of borrowing constraints.

Specifically, in each period a household faces an age-dependent probability of death, denoted by $p_{t,t+1}^d$, and can live up to at most age T . An individual works for the first R ($< T$) years of his life, after which time he is retired. Preferences over consumption are given by the CRRA specification. As before, individuals can borrow and lend at the constant interest rate r , subject to an age-dependent lower limit as specified below. The relevant state variables for this dynamic problem are cash-on-hand (assets plus labor income), ω_t^i , and the vector of mean beliefs, $\widehat{\mathbf{S}}_t = (\widehat{\beta}_t^i, \widehat{z}_t^i)$. Therefore, the dynamic program is

$$\begin{aligned}
 V_t^i(\omega_t^i, \widehat{\beta}_t^i, \widehat{z}_t^i; \alpha^i) &= \max_{C_t^i, a_{t+1}^i} \left\{ \frac{(C_t^i)^{1-\phi}}{1-\phi} + \delta_{t+1} \mathbb{E}_t \left[V_{t+1}^i(\omega_{t+1}^i, \widehat{\beta}_{t+1}^i, \widehat{z}_{t+1}^i; \alpha^i) \right] \right\} \\
 \text{s.t.} \quad C_t^i + a_{t+1}^i &= \omega_t^i \\
 \omega_t^i &= (1+r)a_t^i + Y_t^{\text{disp},i} \\
 a_{t+1}^i &\geq \underline{a}_t, \quad \text{and Kalman recursions}
 \end{aligned}$$

for $t = 1, \dots, R-1$, where V_t^i is the value function of a t -year-old individual; \underline{a}_t is an age-dependent borrowing limit, which will be specified in a moment. The discount rate embeds the survival probability: $\delta_{t+1} \equiv \bar{\delta}(1 - p_{t,t+1}^d)$, where $\bar{\delta}$ is the pure time discount factor.¹¹ Disposable income, $Y_t^{\text{disp},i}$, is income inclusive of the partial insurance via informal mechanisms, as specified in equation (16) below. The evolution of the vector of beliefs and its covariance matrix are governed by the Kalman recursions (27, 28, 29, 30). Finally, the expectation is taken with respect to the conditional distribution of \tilde{y}_{t+1}^i given in (26).¹²

Partial Insurance. Modeling partial insurance in an economy with learning requires some care. It seems plausible to assume that the informal risk-sharing mechanisms available in the society (which allow partial insurance) are subject to the same informational constraints faced by the households themselves. This means that insurance can only be based on perceived shocks (e.g., $\widehat{\eta}_t, \widehat{\xi}_t^i$) rather than on the true but unobservable sources of uncertainty (e.g., β^i or η). With this assumption, we

¹¹Clearly, some individuals will die with debt as long as $\underline{a}_t < 0$. Since we are not conducting a general equilibrium analysis, it is not necessary to model explicitly how this debt is disposed of. The probability of death is very low up to age 55, so the behavior up to that age is little affected. After retirement, we assume the presence of perfect annuity markets.

¹²In our empirical analysis we have estimated both a version where ε_t is observable and a version where it is unobservable. In both cases, we found that this assumption has a negligible effect on the empirical results. But when ε_t is observable, the expectations operator in this dynamic program needs to be computed as a double integral, which slows down computation. Thus, we report all the results below assuming that ε_t is *not* observable.

define a household's log disposable income to be

$$y_t^{\text{disp},i} \equiv y_t^i - \theta \widehat{\xi}_t^i, \quad (15)$$

where y_t^i is as given in equation (1).¹³ To obtain an alternative expression, we substitute into (15) the expression for $\widehat{\xi}_t^i$ from equation (2) yielding

$$y_t^{\text{disp},i} = y_t^i - \theta (\tilde{y}_t^i - \mathbb{E}_{t-1}(\tilde{y}_t^i)) = (1 - \theta)y_t^i + \theta \mathbb{E}_{t-1}(y_t^i).$$

Now it can be seen that disposable income is a convex combination of actual income and the expected income in period t . When $\theta = 0$, there is no partial insurance. When $\theta = 1$, disposable income equals expected income, so any shock in period t is completely insured via informal mechanisms. Thus, θ provides a useful measure of partial insurance. Finally, the level of disposable income is obtained by exponentiating (15) and adding an income floor:¹⁴

$$Y_t^{\text{disp},i} = \underline{Y} + \exp(y_t^{\text{disp},i}). \quad (16)$$

Borrowing Constraints. As discussed above, the tightness of the borrowing constraints can have a potentially large impact on the estimates of the income process parameters. Therefore, rather than picking a (n arbitrary) value for \underline{a}_t beforehand, we estimate the borrowing limit along with the rest of the structural parameters. Our starting point is the *natural* borrowing limit, which is essentially the *loosest* limit that still guarantees full repayment by the last period (T) even if the household gets the lowest income realization in every period. Here, this limit would be $\underline{a}_t = \sum_{\tau=1}^{T-t} \gamma^\tau \min(Y_\tau) = \underline{Y} \frac{1-\gamma^{T-t}}{1-\gamma}$. Although this is a conceptually clean and useful benchmark, it has the somewhat questionable implication that households face a looser constraint when young rather than when old, which is the opposite of what we seem to observe in real life. To capture this possibility in a simple fashion (without introducing the complications of default and credit rating), we assume that banks use a potentially higher interest rate to discount households' future labor income during working years in calculating their borrowing limit, but simply apply the risk-free rate for discounting retirement income. That is, we define

$$\underline{a}_t \equiv \underline{Y} \left[\sum_{\tau=1}^{R-t} (\psi\gamma)^\tau + \psi^{R-t+1} \sum_{\tau=R-t+1}^{T-t} \gamma^\tau \right],$$

¹³The process in (1) is richer than most of the specifications used to calibrate incomplete markets models, yet it still allows meaningful empirical identification. Although one could postulate even more general processes (for example, allowing for separate permanent and persistent shocks components, or considering household-specific quadratic terms), empirical identification would be problematic given the limited size of PSID samples. Guvenen et al. (2012) consider such richer specifications but have access to a substantially larger and cleaner dataset from administrative records. However, they do not have any consumption data, which is the focus of the present paper.

¹⁴The purpose of the income floor, $\underline{Y} > 0$, is to make sure that income realizations are never too close to zero, which can happen since all the terms in y_t^i are normally distributed and, thus, have no lower bound. The possibility of a zero income state would make borrowing effectively impossible (since the household would have no funds to pay back its debt), which does not seem realistic given the sizable uncollateralized borrowing by households observed in the data.

where $\psi \in [0, 1]$ measures the tightness of the borrowing limit. When $\psi = 0$, no borrowing is allowed against future labor income; when $\psi = 1$, households can borrow up to the natural limit. This specification generates borrowing limits that become looser (tighter) with age when ψ is sufficiently low (close to 1). The tightness parameter ψ will be estimated in the empirical analysis.

Retirement Period. During retirement, households receive annual pension payments from a retirement system that mimics the salient features of the US Social Security Administration’s Old-Age Insurance Benefits System. Since there is no uncertainty (or learning) after retirement, the problem simplifies significantly:

$$V_t^i(\omega_t^i; Y) = \max_{c_t^i, a_{t+1}^i} \left[\frac{(C_t^i)^{1-\phi}}{1-\phi} + \delta V_{t+1}^i(\omega_{t+1}^i; Y) \right] \quad (17)$$

s.t. $Y^i = \Phi(Y_R^i; \bar{Y})$, and eq. (5, 6)

for $t = R, \dots, T$, with $V_{T+1} \equiv 0$.

Social Security System. The pension system in the model—captured by the function Φ —mimics the US Social Security system, with one notable difference. In the actual US pension system, retirement income is tied to households’ average labor income during the working years (denoted by \bar{Y}^i).¹⁵ Adopting this exact structure here, however, would add another state variable— \bar{Y}^i —to the dynamic problem above, increasing the already high computational burden of the estimation. So, instead, we adopt the same functional form used in the US system for the $\Phi(\cdot)$ function, but rather than use \bar{Y}^i , we use instead the *predicted* average income given the worker’s income at the retirement age (Y_R^i). This is accomplished by first running the cross-sectional regression $\bar{Y}^i = k_0 + k_1 Y_R^i$ and then using the predicted average income implied by this regression, which we denote by $\hat{Y}(Y_R^i)$. This structure does not add a state variable but recognizes the empirical relationship between average income and the income at retirement age implied by each stochastic process. Letting \bar{Y} denote the economy-wide average lifetime labor income and defining $\tilde{Y}_R^i \equiv \hat{Y}(Y_R^i)/\bar{Y}$, the pension function is given by

$$\Phi(Y_R^i; \bar{Y}) = \bar{Y} \times \begin{cases} 0.9\tilde{Y}_R^i & \text{if } \tilde{Y}_R^i \leq 0.3 \\ 0.27 + 0.32(\tilde{Y}_R^i - 0.3) & \text{if } 0.3 < \tilde{Y}_R^i \leq 2 \\ 0.81 + 0.15(\tilde{Y}_R^i - 2) & \text{if } 2 < \tilde{Y}_R^i \leq 4.1 \\ 1.13 & \text{if } 4.1 \leq \tilde{Y}_R^i. \end{cases}$$

¹⁵More precisely, the average is taken over the 35 working years with the highest earnings.

4 The Data

This section discusses the data used in the empirical analysis and provides definitions for key variables. The unit of analysis in this paper is a married household—so both income and consumption are measured at the household level.

Constructing a Panel of Imputed Consumption. The PSID has a long panel dimension but covers limited categories of consumption expenditures, whereas the Consumer Expenditure Survey (CE) has detailed expenditures over a short period of time (four quarters). As a result, most previous work has either used food expenditures (e.g., Hall and Mishkin (1982), Altug and Miller (1990), Hayashi et al. (1996)) as a measure of nondurable consumption (available in PSID) or resorted to using repeated cross sections from the CE under additional assumptions.

Blundell et al. (2006) develop a structural method that imputes consumption expenditures for PSID households using information from the CE survey. These authors show that several statistics calculated using the imputed measure compare quite well with their counterparts from the CE data. However, because the CE dataset is available on a continuous basis only after 1980, their method has been tailored to generate imputed consumption for the PSID from 1980 to 1992. In this paper, we modify and extend the method proposed by these authors by also using the information in the (large) 1972–73 CE waves to obtain an imputed panel that covers the period 1968 to 1992. We also conduct a detailed validation study, examining a broader set of statistics to show that the method works well for this longer sample. Appendix C.3 contains the details.

Measure of Household Labor Income. The household labor income data come from the PSID. We restrict attention to households that are in the nationally representative core sample, whose head is between the ages of 25 and 55 (inclusive), and has non-missing data on food expenditures and head and wife’s labor income. In the PSID, households report their total taxable income, which includes labor income, transfers, and financial income of all the members in the household. The measure of labor income that we use subtracts financial income from this measure and therefore includes the labor income of the head and wife as well as several categories of transfer income (unemployment benefits, social security income, pension income, worker’s compensation, welfare payments, child support, financial help from relatives, and so on). We then subtract the labor portion of income taxes paid by each household. A more complete description of the sample selection criteria we use and other details on the PSID (such as the method for estimating taxes) are contained in Appendix C.

Converting the Data to Per-Adult Equivalent Units. We adjust both the imputed consumption and income measures for demographic differences across households, since such differences have no counterpart in our model. This is accomplished by regressing each variable on family size, a race dummy, region dummies, a dummy indicating whether the head is employed, a dummy indi-

cating residence in a large city, and a set of cohort dummies.¹⁶ We then use the residuals of these regressions—which are interpreted as consumption and income per-adult equivalent—in the analysis below.

5 Econometric Methodology: Indirect Inference Estimation

We now describe the method—indirect inference—used to estimate the parameters of the structural model laid out in the previous section. Indirect inference is a simulation-based estimation method whose hallmark is the use of an “auxiliary model” to capture aspects of the data upon which to base the estimation. One of its key advantages over GMM is that this auxiliary model need not correspond to any valid moment condition of the economic model in order to obtain consistent estimates of the structural parameters. Indirect inference therefore permits significant flexibility in choosing an auxiliary model: it can be any statistical model relating the model variables to each other provided that each structural parameter has an independent effect on at least one (reduced-form) parameter of the auxiliary model.¹⁷ Indirect inference can therefore accommodate many realistic features in the consumption-savings models (such as borrowing constraints and risk aversion) that GMM cannot. We elaborate on this point in the next section and then in the following section describe the auxiliary model that we use in estimation.

5.1 A Brief Digression: Indirect Inference versus GMM

The standard method for estimating consumption-savings models since Hall and Mishkin (1982) has been to derive explicit structural expressions that link observable variables (such as consumption and income) to unobservable variables (such as persistent and transitory income shocks). It is useful to contrast this approach with the indirect inference method we employ in this paper. To this end, consider the certainty-equivalent consumption-savings model described above, but in keeping with earlier work (Hall and Mishkin (1982); Blundell et al. (2008)), suppose that the income process is the sum of a permanent and a transitory shock, which implies $\Delta Y_t = \eta_t + \Delta \varepsilon_t$, where $\eta_t, \varepsilon_t \sim \text{i.i.d.}$ Here, it can be shown that $\Delta C_t = \eta_t + \varphi_t \varepsilon_t$. These two equations can be jointly used to estimate the ratio of shock variances ($\sigma_\eta^2/\sigma_\varepsilon^2$), which is a measure of the persistence, or durability, of shocks. To see how, consider this regression:

$$\Delta C_t = \pi \times \Delta Y_t + \text{error}, \quad \text{where} \quad \pi \equiv \frac{1 + \varphi_t(\sigma_\varepsilon^2/\sigma_\eta^2)}{1 + 2(\sigma_\varepsilon^2/\sigma_\eta^2)}. \quad (18)$$

¹⁶Each cohort is defined by 5-year bands based on the birth year of each head of household, e.g., those born between 1951 and 1955, 1956 and 1960, etc.

¹⁷More formally, indirect inference is consistent if the mapping from structural parameters to the parameters of the auxiliary model has full rank near the true structural parameter vector. For more details, see Smith (1993) and Gourieroux et al. (1993).

Thus, $\sigma_\eta^2/\sigma_\varepsilon^2$ can be identified by estimating π from this regression. When income shocks are permanent (i.e., $\sigma_\varepsilon^2 = 0$), we get $\pi = 1$, and consumption moves in lockstep with income from (18). At the other extreme, when $\sigma_\eta^2 = 0$, $\pi = \varphi_t/2 \approx 0$, implying that consumption fluctuations will be much smoother than those in income. This identification strategy is commonly used in the literature, where a set of coefficients (such as π above) are estimated via GMM or a minimum-distance method. Notice, however, that even in this simple example, inference is feasible only because we have two exact expressions that link ΔC_t and ΔY_t to η_t and ε_t , which were used to derive the expression for π . It follows that the consistency of the estimates relies—potentially critically—on the validity of the assumptions we made above (quadratic utility, no borrowing constraints, etc.) to make the derivation of these two equations feasible.

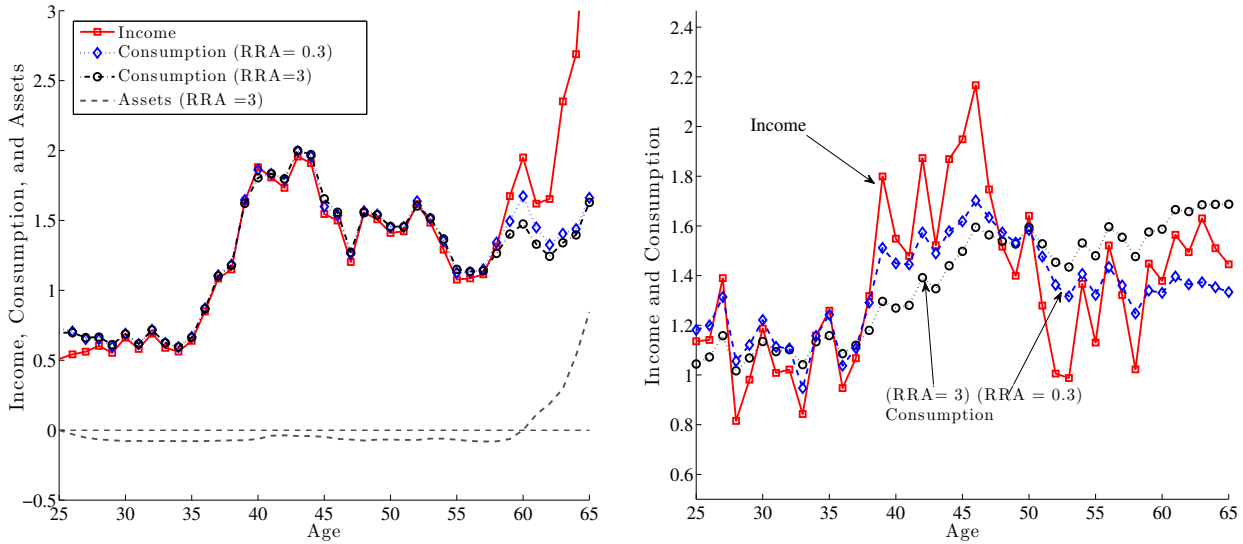
To illustrate some of the potential difficulties with this approach, now consider the example shown in the left panel of Figure 4. The line marked with squares is income, and the dashed line marked with circles is consumption. (Ignore the dashed-dotted line for now.) The main observation that is obvious in this picture is that the individual’s consumption and income move almost one-for-one during most of his life (from 25 to 57). Thus, using (18) for inference would lead one to conclude that income shocks are nearly permanent. Indeed, a GMM estimation using the coefficient above (i.e., $\sigma(\Delta C_t, \Delta Y_t)/\sigma^2(\Delta Y_t)$) as the only moment yields $\sigma_\eta^2/\sigma_\varepsilon^2 = 27.2$ —that is, income shocks are almost completely permanent, when in fact the true ratio used in the simulation was $\sigma_\eta^2/\sigma_\varepsilon^2 = 2$!

To understand the source of this substantial bias, now consider the dashed-dotted line, which plots the asset position (scaled to fit in the figure): the household is right up against the constraint up to the mid-50s, after which time savings for retirement starts to kick in. But borrowing constraints had to be ignored to derive (18), which turns out to be critical for this household. In other words, for this household consumption moves one-for-one with its income because he is not able to borrow any further—not because income shocks are permanent and the household chooses to fully accommodate these shocks. Although this example is clearly an extreme case, it sounds a cautionary note that this common assumption has the potential to bias inference if a non-negligible fraction of households are borrowing constrained.

A second example is shown in the right panel of Figure 4. A particular income path is plotted here along with the consumption path that would be chosen by an individual when his risk aversion is, respectively, 3 (circles) and 0.3 (diamonds).¹⁸ As could be expected, when the individual has higher risk aversion, the consumption path is much smoother than when he is more risk tolerant. Now assume that the econometrician observes the consumption path of the risk tolerant individual, but since risk aversion is not observable, he assumes a value of 3 as a reasonable figure. (When equation (18) above is derived using quadratic utility, an assumption about the value of risk aversion is made automatically.) In this case, observing the strong response of consumption to income, the econometrician would be led to overestimate the persistence of income shocks. Of course, the opposite case would arise if the econometrician assumes a risk aversion lower than the true value. The indirect inference approach we

¹⁸Although assets have not been plotted to save space, the individual is never constrained in either case.

Figure 4: Inferring Persistence of Shocks Using GMM Moment Conditions



use in this paper allows us to relax several of these assumptions, model and estimate the tightness of the borrowing constraints and the time discount factor, and explore the effects of different risk aversion parameters on the estimates of income dynamics parameters.

5.2 A Parsimonious and Feasible Auxiliary Model

Although indirect inference shares a basic similarity to the method of simulated moments (MSM) it differs from in its use of an auxiliary model to generate moment conditions. In particular, indirect inference allows one to think in terms of the dynamic structural relationships that characterize most economic models (such as (9) and (11)) but are difficult to express as simple unconditional moments. Indirect inference is thus similar in some respects to likelihood methods that take into account all of the structural model’s implications for the data. In this section, we describe a set of linear equations suggested by the linear-quadratic approximation to our structural model to serve as an auxiliary model. Because the auxiliary model is linear, it clearly cannot be an exact representation of the non-linear structural model, but it can approximate it more closely than MSM typically can using simple unconditional moments. Moreover, the approach bears some similarities to the early literature on estimation of linear rational expectations models subject to the “cross-equation” restrictions imposed by the structural model on the coefficients of the (linear) reduced-form equations.

In light of Propositions 1 and 2, equations (9) and/or (11) are ideal candidates to form the basis of an auxiliary model. For example, the response coefficient in equation (9) is $\Pi_t(\lambda, \sigma_\beta, \sigma_{\alpha\beta}, \sigma_\eta, r, \rho; R, T)$, which depends on several key variables that we wish to estimate.¹⁹ The presence of the unobserved beliefs $\hat{\beta}_{t|t-1}^i$ and $\hat{z}_{t|t-1}^i$, however, makes it impossible to use these equations directly as an auxiliary model. But, as discussed in Section 2.2, current beliefs depend both on an individual’s prior beliefs

¹⁹The dependence of Π on $\lambda, \sigma_\beta, \sigma_{\alpha\beta}$, and σ_η can be seen from the formulas for A_t and B_t .

before beginning to work and on income realizations over the working life. Furthermore, prior beliefs are likely to be correlated with future income realizations. We can therefore use leads and lags of income as proxies for beliefs, leading to the following equation for consumption that depends only on observables:

$$c_t = \mathbf{a}'\mathbf{X}_{c,t} + \epsilon_t^c = a_0 + a_1y_{t-1} + a_2y_{t-2} + a_3y_{t+1} + a_4y_{t+2} + a_5\bar{y}_{1,t-3} + a_6\bar{y}_{t+3,R} + a_7\Delta y_{1,t-3} + a_8\Delta y_{t+3,R} + a_9c_{t-1} + a_{10}c_{t-2} + a_{11}c_{t+1} + a_{12}c_{t+2} + \epsilon_t^c, \quad (19)$$

where c_t and y_t are the log of consumption and income, respectively; $\Delta y_{\tau_1,\tau_2}$ and \bar{y}_{τ_1,τ_2} are, respectively, the average of the growth rate and the average of the level of log income from time τ_1 to τ_2 ; and \mathbf{a} and $\mathbf{X}_{c,t}$ denote the vectors of coefficients and regressors. The use of logged variables in this regression seems natural given that the utility function is CRRA and income is log-normal. By adding past and future income growth rates as well as past and future income levels, this regression captures the predictions made by the HIP and RIP models. Leads and lags of consumption capture the dynamics of consumption around the current date.

To complete the auxiliary model, we add a second equation with y_t as the dependent variable and use all of the income regressors above as right-hand-side variables:

$$y_t = \mathbf{b}'\mathbf{X}_{y,t} + \epsilon_t^y = b_0 + b_1y_{t-1} + b_2y_{t-2} + b_3y_{t+1} + b_4y_{t+2} + b_5\bar{y}_{1,t-3} + b_6\bar{y}_{t+3,R} + b_7\Delta y_{1,t-3} + b_8\Delta y_{t+3,R} + \epsilon_t^y. \quad (20)$$

We divide the population into two age groups—those between 25 and 38 years of age, and those between 39 and 55 years of age—and allow the coefficients of the auxiliary model to vary across the two groups.²⁰ For each age group, the auxiliary model has 22 regression coefficients (13 in the first equation and 9 in the second), two residual variances, and the correlation between the two residuals for a total of 25 parameters. With two age groups, this yields a total of 50 reduced-form parameters that determine the likelihood of the auxiliary model.

The goal of our estimation procedure, then, is to choose the parameters of the structural model so that the auxiliary model parameters estimated using the observed data are “close” to those estimated using data simulated from the model. Appendix B describes the metric we use to measure the distance between two sets of auxiliary model parameters; we minimize this metric by maximizing a “Gaussian objective” constructed out of this auxiliary model subject to the restrictions that the structural model imposes on its parameters. Using a simple example, Appendix A shows that our approach is asymptotically identical to minimizing a quadratic form in the difference between the two sets of parameters with a specific choice for the weighting matrix.

²⁰Although the auxiliary model would correspond to the structural equations in (9) and (11) more closely if the coefficients were varying freely with age, this would increase the number of parameters in the auxiliary model substantially. We have experimented with having one or three age groups but found the small sample performance to be better with the specification adopted here.

Appendix B also presents the results of a Monte Carlo study to gauge the ability of the auxiliary model described above to identify the structural parameters. This study shows that it works very well, with minimal bias and tight confidence intervals. Alternative auxiliary models with more or less parsimony also perform reasonably well, though not as well as the one we use to obtain our empirical results.

5.3 Empirical Preliminaries

Preset Values. Working life is $R = 41$ years, and the retirement duration is 15 years ($T = 80$). The price of the one-period discount bond is set to 0.95, implying an interest rate of $r = 1/0.95 - 1 \approx 5.26\%$. The *common* lifecycle profile of log income ($g(\cdot)$ in (1)) is captured by feeding into the model the empirical profile computed from our PSID estimation sample. The (potential) lower bound of income, \underline{Y} , is set to be 5% of average income in this economy. As we discuss in further detail below, we have found it very challenging to separately identify the coefficient of relative risk aversion, δ , and the time discount factor, ϕ , despite substantial efforts (expanding the auxiliary model, adding some new moments, and so on). Therefore, in our benchmark case, we proceed by fixing ϕ at 2 and estimating δ . Later, we will conduct detailed sensitivity analyses with respect to the values of the parameters preset in this section (including ϕ , r , and \underline{Y}).

Measurement Error. We add measurement error to simulated consumption and income data:

$$\begin{aligned} y_t^{i,*} &= y_t^i + u_t^{i,y}, \\ c_t^{i,*} &= c_t^i + \bar{w}^{i,c} + u_t^{i,c}, \end{aligned}$$

where $y_t^{i,*}$ and $c_t^{i,*}$ are measured variables of household i , respectively, and $u_t^{i,y}$ and $u_t^{i,c}$ are zero-mean random variables that are independent over time, with standard deviations of σ_y and σ_c .²¹ Notice that we have also added a second term to consumption, $\bar{w}^{i,c}$, which is an individual fixed measurement error with potentially nonzero mean in the cross section, μ_{c_0} , and standard deviation σ_{c_0} . This fixed effect is needed for two reasons. First, and most important, recall that we regress both income and consumption on a set of demographics to convert these variables into per-adult equivalent terms. One effect of this adjustment is to introduce level differences between consumption and income, the magnitudes of which vary across households. This fixed effect captures such differences.²² Second, the model described above abstracts from initial wealth differences across households. These differences in wealth would also drive a household-specific wedge between the levels of income and consumption. The fixed effect is also a simple way to capture these differences in initial wealth levels.

²¹The variable $u_t^{i,c}$ contains the imputation error, which is heteroskedastic owing to the mechanics of the imputation procedure (a point also observed by BPP). In our estimation we will impose stationarity in variances to deal with the computational burden, but further work is needed on the effects of this assumption.

²²See Appendix B.1 for further details.

Missing Observations. In the observed data set, we include only households with at least five observations between the ages of 25 and 55 (of the head), for a total of 2,235 households with an average of 12 observations on each (for a total of 26,441 household-year observations). With more than half of the observations missing compared to a fully balanced panel, one question is, how does one run the regressions in (19) and (20)? For missing values of regressors, we simply use values that are constructed or “filled-in” using a reasonable procedure.²³ However, on the left-hand side of regressions we only use actual (i.e., not filled-in) observations. As will become clear in the Monte Carlo analysis, a strength of the indirect inference method is that the particular filling-in method is not critical for the estimation as long as the same procedure is applied consistently to real and simulated data (as we do). As an extreme example, if we simply fill in all missing values with zeros, the estimates would still be consistent, as we show in the Monte Carlo analysis.

Matching the Wealth-to-Income Ratio. The auxiliary model specified above does not explicitly target the amount of savings and wealth generated by the estimated model. One goal of this paper, however, is to provide estimates of income processes (together with combinations of time discount factor, risk aversion, and borrowing constraints) that can be used for calibrating life cycle models to be used in quantitative macroeconomic analysis. For this purpose, it is important to make sure that the estimated model yields a reasonable amount of savings. Furthermore, households’ wealth holdings contain some useful information that can shed light on (some of) the estimated parameters. Unfortunately, detailed information about asset holdings in the PSID is available only intermittently (and in fact, only in 1984 and 1989 in our sample period), which makes it difficult to use some dynamic equations, such as (19) and (20), as additional auxiliary model regressions. As a compromise, we add one static moment condition—the median wealth-to-income ratio (denoted by WY henceforth)—to the objective function that the indirect inference procedure is minimizing, to be described in a moment. Specifically, we add $10 \times (WY_{PSID} - WY_{Model})^2$ to the objective function in (46). This additional moment also makes pinning down the discount factor, δ , much easier in the estimation.

To calculate the empirical target, WY_{PSID} , we use the wealth supplement of PSID, available in 1984 and 1989. We use households in our estimation sample that were also present in the PSID in 1984 and/or 1989 and use net worth as the measure of wealth (see Appendix C.2 for the precise definition). The median value of WY is 0.99 in 1984 and 1.17 in 1989, averaging 1.08, which we take as our empirical target.²⁴

6 Estimation Results

In this section, we present the estimation results. In Section 6.1, we discuss the structural parameter estimates. In Section 6.2, we compare the implications of the estimated model to the data for lifecycle patterns of income and consumption that has received a lot of attention in the literature. Section

²³Appendix B.1 contains the details.

²⁴When we include all households up to age 65, the corresponding ratios are 1.19 and 1.43.

6.3 studies simulated paths for consumption and income under different estimated models to shed light on how the joint dynamics of consumption and income can be informative about various model parameters.

6.1 Structural Parameters

Table I reports the estimated structural parameters.

Parameters of the Income Process: $\rho, \sigma_\eta, \sigma_\varepsilon, \sigma_\alpha, \sigma_\beta, \text{corr}_{\alpha,\beta}$. The first column reports the results from our benchmark model described in Section 3. First, the AR(1) process has an estimated annual persistence of 0.756 and an innovation standard deviation of 22.7%, both estimated precisely. Below, we present sensitivity analyses that show that the estimates of persistence are quite robust. Therefore, we conclude that the joint dynamics of consumption and income data do not lend support to permanent shocks as a reasonable representation of income shocks. That said, we do not mean to deny the existence of permanent shocks—as even the most casual observation tells us that such shocks exist. Rather, these results indicate that when income shocks are modeled as a univariate persistent process, the *typical* shock received by a *typical* household is better represented as having a moderate persistence.

Turning to the fixed heterogeneity across households, first, the dispersion of the fixed (level) effects in income, σ_α , is 0.288. This figure is consistent with the estimates in the existing literature. The standard deviation of income *growth rates*—a key parameter of interest—is estimated to be 1.76%, which is substantial. For example, by age 55, an individual with a β^i one (two) standard deviation(s) above the mean will earn 1.68 times (2.80 times) the median income. Moreover, both σ_α and σ_β are estimated precisely, with t -statistics exceeding 10. Finally, the correlation between these two parameters ($\text{corr}_{\alpha\beta}$) is small but negative, -0.13 , though it cannot be statistically distinguished from zero. We can, however, easily reject a strong correlation of either sign.

The parameters discussed so far can all be identified with income data alone. Thus, to better understand what consumption data brings, it is instructive to compare these estimates to those obtained by using only the income regression (20), reported in column 4. The main noticeable differences are in the estimates of σ_β and $\text{corr}_{\alpha\beta}$. The remaining parameters are very similar across the two cases, suggesting that they are pinned down very well with income data and that there is no conflicting information in consumption data. The estimate of σ_β is lower, at 1.34%, compared to the benchmark case, but now $\text{corr}_{\alpha\beta}$ is positive and fairly large: 0.56. As a result, the rise in income inequality generated by heterogeneous profiles in this case is, if anything, slightly larger than in the benchmark case (by 4 log points). Finally, notice that the standard errors of these two parameters are significantly higher when estimated with income data alone, suggesting that the two parameters are not identified very precisely with income data alone. Consumption data are especially informative about these two key parameters.

Table I: ESTIMATING THE FULL CONSUMPTION-SAVINGS MODEL

<i>Data:</i>	Income and Consumption			Income
<i>Partial Insurance?</i>	Benchmark	Yes	Self-insurance	
	Insure $\hat{\xi}$	Insure \hat{z}	($\theta \equiv 0$)	
	(1)	(2)	(3)	(4)
<i>Income Processes Parameters (can be identified with income data alone)</i>				
σ_α	0.288 (0.017)	0.286 (0.017)	0.265 (0.022)	0.298 (0.038)
σ_β	1.764 (0.137)	1.881 (0.131)	1.66 (0.118)	1.343 (0.271)
$corr_{\alpha\beta}$	-0.127 (0.102)	-0.140 (0.090)	-0.112 (0.121)	0.558 (0.289)
ρ	0.756 (0.023)	0.755 (0.021)	0.768 (0.025)	0.783 (0.022)
σ_η	0.227 (0.007)	0.427 (0.012)	0.196 (0.005)	0.200 (0.005)
σ_ε	0.100 (0.016)	0.004 (0.018)	0.008 (0.021)	0.147 (0.005)
<i>Economic Model Parameters (need consumption data)</i>				
λ (prior uncertainty)	0.438 (0.045)	0.429 (0.042)	0.345 (0.074)	— —
θ (partial insurance)	0.451 (0.028)	0.552 (0.031)	0.00* —	— —
ψ (borrowing constraint)	0.582 (0.040)	0.859 (0.048)	0.855 (0.083)	— —
$\bar{\delta}$ (subjective time discount factor)	0.953 (0.001)	0.955 (0.001)	0.956 (0.001)	— —
<i>Measurement Error and Transitory Shocks (need consumption data)</i>				
σ_y	0.165 (0.006)	0.146 (0.005)	0.148 (0.007)	— —
σ_c	0.355 (0.007)	0.356 (0.006)	0.356 (0.002)	— —
σ_{c_0}	0.430 (0.011)	0.429 (0.011)	0.427 (0.009)	— —
max % constrained...	17.4%	21.8%	19.5%	
... at age	29	27	27	
$\underline{a}_{25}/\bar{Y}$	0.08	0.13	0.11	—
$\underline{a}_{55}/\bar{Y}$	0.06	0.11	0.09	—

Note: Standard errors (in parentheses) are obtained via parametric bootstrap with 150 repetitions.

Parameters Identified with Consumption Data: $\lambda, \theta, \delta, \psi$. Now we return to column 1 in Table I and discuss the structural estimates pertaining to the economic model. First, notice that all four parameters—and especially λ and θ , which are of key interest—are very precisely estimated with t -statistics exceeding 10. Thus, the joint dynamics of consumption and income contain sufficiently rich information to tightly pin down these parameters.

Turning to the estimates, λ is 0.438, which reveals a modest amount of prior uncertainty regarding individuals' growth rate. To see this, notice that the component of income growth that is predictable by households at time zero represents a substantial fraction of the total dispersion of β^i in the population: $\sigma_{\beta_k}^2 / \sigma_{\beta}^2 = 1 - \lambda^2 = 0.81$. The remaining uncertainty is small but non-negligible, as we show in the next section.

The partial insurance parameter, θ , is estimated to be 0.451, implying that almost one-half of income surprises are smoothed away through informal mechanisms in the data. As we shall see later, this availability of partial insurance further reduces the already moderate amount of income risk implied by the estimates of the income process parameters and λ . As noted earlier, Blundell et al. (2008) also tried to quantify partial insurance, although their framework is rather different from the current one. Although their estimates vary quite a bit across samples and cohorts, their baseline estimate for the whole sample (which is most comparable to ours) is that about 35 percent of permanent shocks are insured along with almost 95% of transitory shocks. In our specification, the income surprise, ξ , includes both transitory and persistent components. So the fact that our estimate of $\theta = 0.45$ is somewhat higher is consistent with their results.

The parameter ψ is estimated to be about 0.582, implying that individuals are able to borrow against only about 60 cents of each dollar of future minimum income at every future date and state. Another useful measure is the maximum debt that a household is allowed to carry as a fraction of average income: $\underline{a}_t / \mathbb{E}(Y^i)$. For a household whose head is 25 years old, this limit is 8% of (the economy-wide) average income and remains quite flat (still 6% at age 55). Finally, in the estimated model, the fraction of households who are constrained peaks at age 29 (17.4% of households) and stays between 6.5% and 13.5% between ages 30 and 40. The fraction constrained falls to about 5% or less beyond age 40.

The estimated time discount factor is 0.953, which implies that $\bar{\delta}(1+r)$ is slightly above unity. However, because average income is growing over the life cycle and individuals further discount the future due to the possibility of death, the average individual is impatient—in the sense of Deaton (1991)—according to these estimates. Although the standard error on $\bar{\delta}$ is extremely small, this is conditional on the fixed value of ϕ : unfortunately, the estimate of δ is quite sensitive to the preset value of risk aversion. In Appendix D.6, we conduct a detailed sensitivity investigation with respect to the value of γ and show that our results are robust over a wide range of its values.

Measurement Errors and Transitory Shocks: $\sigma_{\varepsilon}, \sigma_y, \sigma_c, \sigma_{c_0}, \mu_c$. With consumption data, in principle, we can tell transitory shocks apart from i.i.d. measurement error in income, since consump-

tion should respond to the former but not to the latter. In practice, however, because the response of consumption to transitory shocks is proportional to its annuitized value—which is small—this response is rather weak, and identification is a problem empirically. In this framework, however, borrowing constraints are binding for a non-negligible fraction of households. As a result, these households’ consumption would move one-for-one with transitory shocks, allowing us to distinguish these shocks from pure measurement error. We estimate the standard deviation of the transitory measurement error in income to be about 14.7% annually, whereas true i.i.d. shocks to income have a standard deviation of about 10% annually. Finally, the transitory measurement error in consumption has a standard deviation of 35.5% and includes the noise introduced by the imputation method. Furthermore, the fixed effect in measured consumption has a standard deviation of 43%, and both components are estimated with extremely high precision.

Partial Insurance: An Alternative Specification

The baseline model featured partial insurance against income *surprises*, e.g., $\hat{\xi}$. An important implication is that once a household’s expectations adjust to incorporate an income shock, insurance also ceases to exist. To see what this means, consider the case without learning and assume shocks are permanent: $\lambda \equiv 0$ and $\rho \equiv 1$. Now consider a 10% permanent shock in period t : $\eta_t = -0.10$. Income will drop by 10% at all future dates, but the only “unexpected shock” happens at t . After that, expectations fully adjust and there is no surprise. Consequently, partial insurance entails a compensation of $0.10 \times \theta$ in period t and zero in subsequent periods. It is easy to see that the same effect holds true more generally when $\lambda > 0$ and/or $\rho < 1$. This specification may not be unreasonable especially, for example, in a society where individuals display habit formation preferences, in which case the disutility associated with a consumption drop is largest on impact. This formulation would provide front-loaded insurance precisely to alleviate those short-run pains.

That said, an alternative plausible way to think about partial insurance is that it might insure a given shock at all future dates (e.g., disability insurance). To capture this idea, we modify equation (15) so that now a fraction θ of the entire persistent component, z_t , is insured:

$$y_t^{\text{disp},i} \equiv y_t^i - \theta z_t^i. \quad (21)$$

Column 2 in Table I reports the results from this specification. Starting at the top panel, the first four parameters of the stochastic process, $\sigma_\alpha, \sigma_\beta, \text{corr}_{\alpha\beta}$, and ρ , are very little changed from the benchmark specification (column 1). However, the two innovation variances are quite different now: σ_η is now 0.43, almost twice its baseline value. At the same time, notice that the partial insurance parameter is estimated to be $\theta = 0.55$. So, about half of the total *level* of z_t is not transmitted from income to consumption. Thus, the after-insurance component of z_t is about 0.20—little changed from the baseline. Of course, it is interesting to note that the consumption equation in the auxiliary model

contains so much information that it is able to raise the estimate of σ_η well above what is implied by income data alone.

Turning to λ , its estimated value is virtually unchanged from the baseline, indicating that it is robust to this modification to the way partial insurance is modeled. Overall, these results are quite similar to the benchmark case. But perhaps this should not be surprising: z_t is not a very persistent process, so it reverts to its mean rather quickly. Therefore, its largest effect is upon impact, which is also how partial insurance worked in the benchmark case.

Self-Insurance Model: Shutting Down Partial Insurance

Finally, we examine the case when households have no ability to smooth consumption over and above self-insurance. To this end, we reestimate the model by restricting $\theta \equiv 0$. Column 3 of Table I reports the results. Compared with the baseline specification, three differences are worth noting. One, σ_β is slightly lower (1.66% vs. 1.76%). Two, λ is also lower (0.345 vs. 0.438). Combining these two pieces implies that the prior uncertainty (as measured by $\sigma_{\beta,0}$) is about one-third higher in the baseline model compared with the self-insurance model: $0.438 \times 1.76 \approx 0.77$ versus $0.345 \times 1.66 \approx 0.57$. Perhaps this result should not be too surprising: imposing $\theta \equiv 0$ implies that none of the idiosyncratic risk is insurable (beyond self-insurance), so the estimation procedure lowers the overall amount of risk faced by households to compensate for this lack of partial insurance, by estimating a smaller amount of growth rate uncertainty.

Overall, however, the differences between the two sets of estimates are not very large. These results are intuitively consistent with equation (15), which showed that the response of consumption growth to income surprises was separable from the terms involving beliefs and learning (Π_t). So, while θ is precisely estimated, it does not interact with the other parameters of the model (with the exception of λ), at least for the specifications that we consider in this paper.

6.2 Model-Data Comparison: lifecycle Profiles of Income and Consumption

We now evaluate the fit of the estimated model to the U.S. data along three dimensions that are important benchmarks in the incomplete-markets literature. The first two are the evolution of the within-cohort variance of log income and log consumption over the life cycle. The third one is the average lifecycle profile of log consumption. (The average lifecycle profile of income is matched by construction.) We first discuss the benchmark estimates and then turn to the alternative specifications estimated above.

Figure 5 plots the variance of log income and consumption using our PSID estimation sample, after cohort effects are taken out following the usual procedure in the literature (e.g., Deaton and Paxson (1994)). First, in the left panel, the income variance (red dashed-dotted line marked with squares) increases by about 30 log points from age 25 to 55, consistent with figures reported in other studies

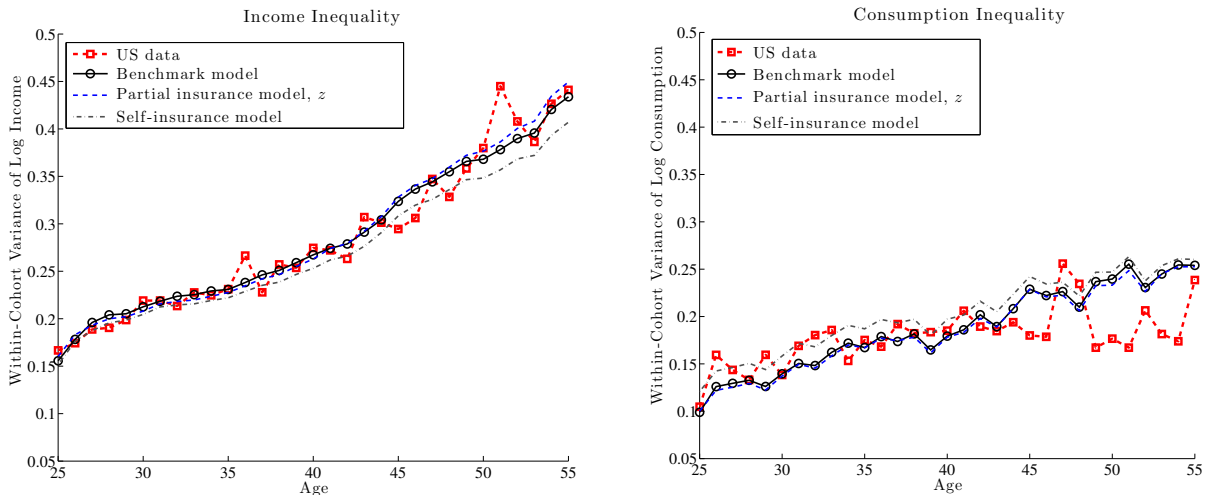


Figure 5: Within-Cohort Income and Consumption Inequality: Data vs Estimated Models

(see, e.g., Storesletten et al. (2004a) and Heathcote et al. (2010)). The estimated benchmark model (black solid line with circles) agrees with the trend in the data quite well, even though these variances were never directly targeted in the estimation. Second, turning to consumption inequality, we see that in the right panel a similar pattern is revealed: the simulated model matches up with the data very well from ages 25 to 45. There is some evidence that consumption inequality grows further after that age in the model, but the data counterpart is quite noisy, making a clear judgement difficult.²⁵ Third, Figure 6 plots the average lifecycle profile of consumption. The estimated model tracks the empirical counterpart from the PSID quite well until about age 49, after which point average consumption in the data starts declining, whereas the model counterpart continues to grow, although at a slower rate.

Overall, the model does a fairly good job of matching these three salient aspects of life cycle income and consumption patterns, despite the fact that these do not appear as explicit moments in the estimation procedure. Furthermore, all three versions of the estimated model appear to fit these profiles, without significant differences among them. In the next section, we will show that looking at household-level dynamics of consumption and income reveal a very different picture and allows us to clearly distinguish between partial- and self-insurance models.

²⁵In figure 5, the rise in the cross-sectional variance of consumption in the US data is about 10 log points, which is small compared with that reported in earlier papers, such as Deaton and Paxson (1994), but is consistent with more recent papers that use CE data with a longer time span, such as Heathcote et al. (2010), Kaplan (2012), and Aguiar and Hurst (forthcoming). In turn, this small rise in consumption inequality found here also explains in part why our structural estimation found a small value of λ , implying a smaller overall amount of uncertainty. Again, despite not having any terms that capture the rise in the variance of consumption explicitly, the auxiliary model attempts to be consistent with the small rise, which in turn requires a small amount of income risk perceived by households.

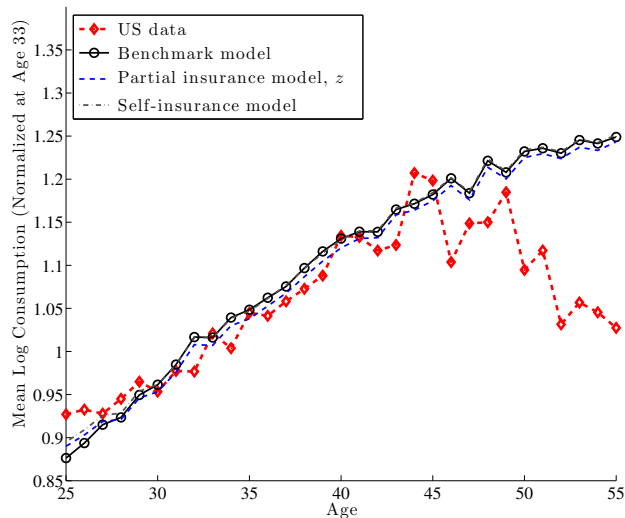


Figure 6: Mean Log Consumption Profile over the Life Cycle: Model vs. US Data

Quantifying Uninsurable Income Uncertainty

As stated in the Introduction, a central goal of this paper is to quantify the magnitude of lifetime income risk *perceived* by households at different points in life, which we are now ready to do. The model we estimated contains three components that affect the *rise* in perceived income risk over time. The first one is the AR(1) income shocks, which is standard. The second one is the uncertainty about β^i . And the third one is the existence of partial insurance. Figure 7 plots the forecast variance of predicted log income at different horizons, from the perspective of a 25-year-old: $\mathbb{E}_{25}((y_{25+t}^{\text{disp}} - \hat{y}_{25+t|25}^{\text{disp}})^2)$ for $t = 1, 2, \dots, 30$. The graph contains four lines. We start from the very top line (black dashed line marked with circles), which shows the maximum amount of uncertainty that is possible in this model—essentially if all the rise in income inequality over the life cycle represented income risk. It is computed by using the benchmark estimates from column 1 of Table I *but* imposing $\theta = 0$ and $\lambda = 1$. This is a useful benchmark, against which we compare each estimated model.

The second line from the top (brown dashed line) plots the amount of perceived income risk in the self-insurance model. As seen here, this profile overlaps very well with the first (top) line in the first five years of the life cycle, implying that all the rise in inequality up to age 30 is considered as risk by households. This is largely because heterogeneity in β^i contributes little to income differences early on, which are instead driven by persistent income shocks. However, notice from the graph that the absolute amount of this risk is quite small. After age 30, however, inequality keeps rising—now almost entirely due to heterogeneous growth rates—whereas perceived risk tapers off because the small initial uncertainty about β^i is largely resolved. Overall, between ages 30 and 55, total income dispersion rises by 25 log points, whereas perceived risk rises by only 3 log points.

Now, we move to the benchmark model (blue line marked with squares), which adds partial

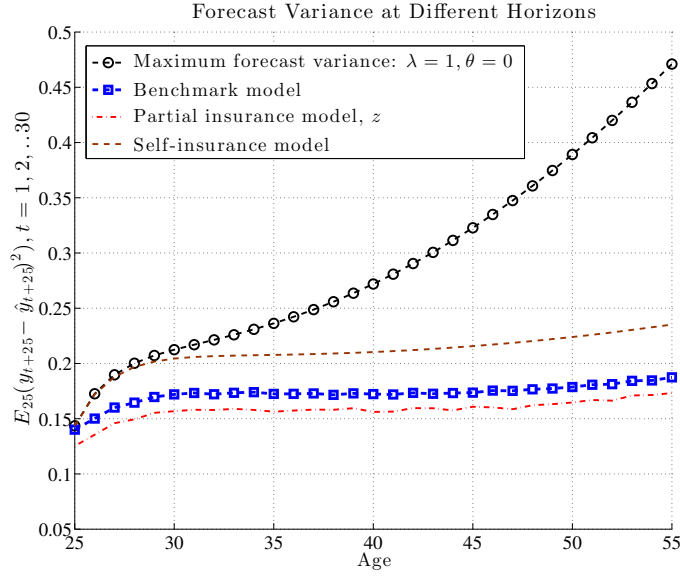


Figure 7: Uninsurable Income Uncertainty

insurance (and has slightly different parameter estimates too). As could be expected, now the overall uncertainty is even lower because part of the risk will be insured through informal channels. This mainly affects the uncertainty in the younger ages though, so the rise after age 35 is similar to the self-insurance model (only slightly lower).

Overall, according to our benchmark estimates, a 25-year-old household perceives that less than one-fifth (5 of the 32 log points) of the increase in income dispersion until age 55 is due to uncertainty, with the rest being predictable and/or insurable by households. Notice that this is not a statement about the *level* of uncertainty but about its *rise* over the life cycle. In contrast, the standard RIP process estimated in the literature together with a self-insurance model would attribute *all* of the 32 log points rise in inequality to risk. The richer model studied here allows us to separate known heterogeneity from risk as well as the part of risk that is insurable, which reveals a much smaller amount of risk than what is currently used to calibrate models in the incomplete markets literature.

At this point, it is also useful to compare the estimates of uncertainty regarding growth rates to the value that Guvenen (2007) gleaned from within-cohort variances and without allowing for partial insurance. The calibrated standard deviation of prior beliefs in that paper is $\sigma_{\beta,0} = 0.012$.²⁶ The corresponding number in this paper is $\sigma_{\beta,0} = 0.0077$. To map these numbers into figures that are easy to interpret, we compute the part of forecast variance that is due to growth rate uncertainty only. In the current paper, this forecast variance for predicting income at age 65 is only 8.4 log points, compared with 23.1 log points using Guvenen (2007)'s numbers. This difference is quite substantial, corroborating our conclusion that accounting for the full joint dynamics of consumption and income significantly reduces our estimate of income uncertainty.

²⁶The notation for λ differs between the two papers, so they are not directly comparable.

6.3 Inspecting the Response of Consumption to Income

One conclusion that we draw from the lifecycle profiles in the previous section is that the three versions of the estimated model have broadly similar implications for lifecycle patterns. In other words, if we were to focus simply on the moments summarized in these three graphs, it would have been difficult to identify and precisely estimate partial insurance along with the other parameters. The reason we are able to do so here is precisely because we use the joint dynamics of consumption and income.

We now take a closer look at the sources of identification. Figure 8 plots six figures in two columns. Each column corresponds to a different household and plots (from top to bottom) the simulated paths of income, annual consumption growth, and wealth over the life cycle. Looking at the left column, we see that household 1 has a fairly high income. The income profile is also hump-shaped and reaches about three times the median income at age 45-50. As seen in the right column, household 2 has a low and flat income profile, earning about half of the median income. The middle panel shows three lines. The dashed line is consumption growth in the self-insurance model. The solid black line displays the *gap* between consumption growth under the benchmark (partial insurance) model and under the self-insurance model: $\Delta C_t^{\theta\xi} - \Delta C_t^{\theta=0}$, where the superscript $\theta\xi$ indicates the benchmark model. The dashed-dotted line displays the same gap, this time for the alternative partial insurance specification (indicated by label superscript θz). Notice that, as expected, both gaps are perfectly negatively correlated with $\Delta C_t^{\theta=0}$, implying that consumption growth under partial insurance is always smoother than under self-insurance. However, the difference is sometimes large and sometimes small. For example, in the left column, household 1 would experience a consumption decline of 35% from age 31 to 32 under self-insurance, but instead experiences a decline of only $35\% - 12\% = 23\%$ under partial insurance. The flip side, of course, is that at age 34, consumption growth is 36% with self-insurance and only 22% under partial insurance.

A key point to observe is that after about age 40, the difference between partial insurance and self-insurance declines significantly—the solid line fluctuates near zero. Looking at the bottom panel makes the reason clear: because this household experiences positive income shocks, its asset holdings rise significantly after about age 40, which makes self-insurance very feasible and effective. In this case, what can be attained in terms of consumption smoothing by partial insurance can also be effectively achieved by self-insurance. Turning to the right panel, though, we see a different picture: the persistently low income of this household does not allow it to accumulate much wealth until later in the working life (about age 55). Consequently, partial insurance allows one to achieve a smoother consumption path than self-insurance at later ages. For example, from ages 50 to 51, consumption declines by 34% under self-insurance but by only 21% with partial insurance.

Inspecting these paths reveals the strength of the panel data analysis of the joint dynamics of consumption and income. Focusing on precisely how a given household's consumption responds to a particular income shock and how this relationship varies by the household's age, past and future average income, income growth rate, and so on, allows us to precisely pin down the parameters of our

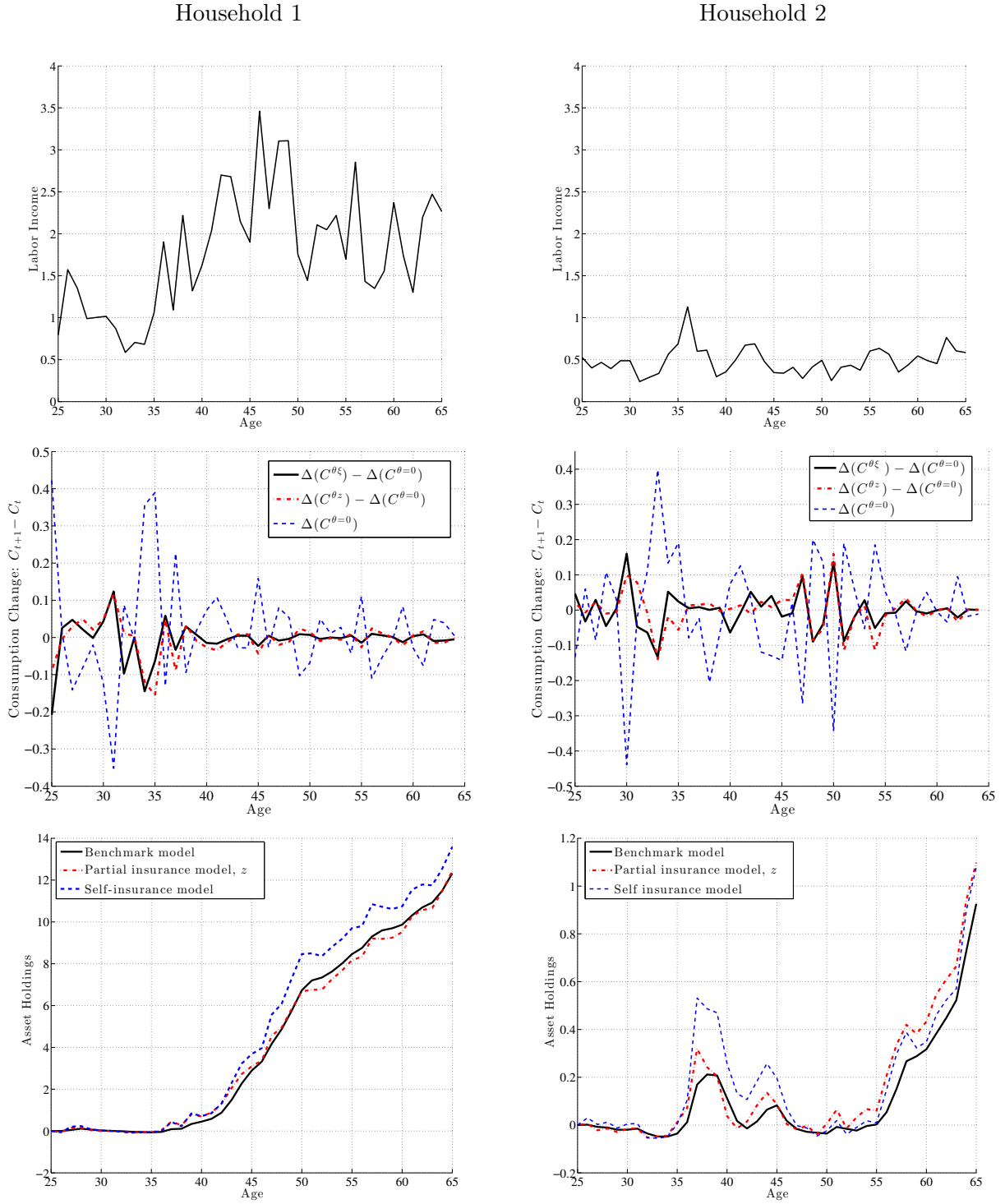


Figure 8: Consumption Response to Income Shocks: Two Sample Paths

model, including the degree of partial insurance. This is an important advantage over some earlier papers that focused only on lifecycle profiles, such as Guvenen (2007) and Storesletten et al. (2004b), among others.

6.4 Inspecting the Auxiliary Model

Before we conclude the analysis of the estimated model, it is also instructive to look at the auxiliary model coefficients. It is useful to know whether the data coefficients are significant as well as how the estimated model does in matching those data coefficients. Table II displays the 50 coefficients of interest. Throughout the table, a boldface font indicates that the coefficient is highly significant (a t -statistic of 5.0 or more).²⁷ The first row reports the data coefficients for the income equation of the young group: all regressors, except for the last two, are very significant. The next row displays their counterparts from the simulated model. For each coefficient, a test is conducted under the null that the model and data coefficients are equal to each other. Rejection with a t -statistic greater than 5 is indicated with the superscript ($\dagger \dagger \dagger$). Similarly, rejections at the 1% and 5% levels are indicated with superscripts ($\dagger \dagger$) and (\dagger), respectively.

In the income equation, the leads and lags of y , as well as the past and future average income, are almost always very significant (with t -statistics exceeding 30). The coefficients from the simulated data match their data counterparts well: for the 11 coefficients in this regression that are very significant (in bold), one cannot reject equality between the model- and data-implied coefficients at a 5% level or higher. Further, for only 3 out of the 18 coefficients in this regression can one reject that the model and data coefficients are statistically different. Overall, the estimated structural model does a fairly good job of matching the auxiliary model coefficients in the income equation.

The consumption regression (19) is more challenging in many ways: the left-hand-side variable is now the output of a complex structural model that filters through the income shocks into the consumption choice. Moreover, consumption data contain not only measurement error (as did income data) but also imputation error. Inspecting the coefficients, we see that the lagged income term is highly significant, and the model matches them quite well. Another set of coefficients that are also very significant is the leads and lags of consumption. Here, we see a fairly clear rejection. As much as the estimation procedure tries to bring these coefficients close together (as evidenced by the fact that the model coefficients are of the right order of magnitude and always have the correct sign), a statistically significant gap still remains.²⁸ Finally, panel C displays the variance-covariance matrix of the regression residuals, whose elements are very precisely estimated. Consequently, the Gaussian objective function puts significant weight on these terms. Out of the six elements in panel C, the

²⁷We compute the standard errors using a nonparametric bootstrap (sampling individuals with replacement).

²⁸One conjecture is that some sort of temporal dependence in the utility function, such as habit formation, could help resolve this problem by increasing the correlation of consumption at lower lags compared with the benchmark model. Nevertheless, exploring alternative utility function specifications is beyond the scope of this paper and is left for future work.

Table II: COEFFICIENTS OF THE AUXILIARY MODEL: BENCHMARK ESTIMATED MODEL VS. US DATA

	$constant$	y_{t-1}	y_{t-2}	y_{t+1}	y_{t+2}	$\bar{y}_{1,t-3}$	$\bar{y}_{t+3,T}$	$\Delta y_{1,t-3}$	$\Delta y_{t+3,T}$	c_{t-1}	c_{t-2}	c_{t+1}	c_{t+2}
PANEL A: INCOME EQUATION													
<i>Young Group:</i>													
(1)	Data	-0.036^a	0.346	0.360	0.077	0.097	0.150	0.037	-0.022				
(2)	Model	-0.027 ^{††b}	0.337	0.356	0.093	0.099	0.131	-0.020 ^{††}	-0.014				
<i>Middle-Age Group:</i>													
(3)	Data	0.006 ^{**}	0.418	0.358	0.111	0.093	0.043 ^{**}	0.031	0.028				
(4)	Model	-0.003 ^{††}	0.407	0.380	0.118	0.104	0.017	0.001	0.048				
PANEL B: CONSUMPTION EQUATION													
<i>Young Group:</i>													
(5)	Data	-0.007	0.108	0.042 [*]	-0.023	-0.005	-0.045 [*]	0.030	-0.002	0.248	0.262	0.178	0.175
(6)	Model	-0.022 ^{††}	0.101	0.092 [†]	-0.024	-0.027	-0.092	0.010	0.011	0.210 ^{††}	0.207 ^{†††}	0.246 ^{†††}	0.229 ^{†††}
<i>Middle-Age Group:</i>													
(7)	Data	-0.004	0.136	0.046 ^{**}	-0.014	-0.040 [*]	0.012	0.030	0.028	0.270	0.260	0.177	0.187
(8)	Model	0.006 ^{††}	0.105	0.061	0.010	-0.052	0.042 ^{†††}	0.029	0.031	0.200 ^{†††}	0.212 ^{††}	0.224 ^{††}	0.255 ^{†††}
PANEL C: RESIDUAL VARIANCES AND CORRELATIONS													
<i>Young Group</i>													
(9)	Data	$\sigma^2(\epsilon_t^y)$	0.222	0.396	$\rho(\epsilon_t^y, \epsilon_t^c)$	0.117	$\sigma^2(\epsilon_t^y)$	0.235	$\sigma^2(\epsilon_t^c)$	0.379	0.114	$\rho(\epsilon_t^y, \epsilon_t^c)$	
(10)	Model		0.226	0.391		0.135 ^{††}		0.232		0.388	0.116		

^aA bold font indicates that the estimated parameter has a t -statistic greater than 5.0; superscripts (**) and (*) indicate that the coefficient is significant at a 1% and 5% level, respectively. The largest t -statistic for parameter estimates is 34.5, and for the residual variances it is 67.81.

^bParameters estimated using simulated data are marked with one, two, or three (†) signs in the superscript if the t -statistic for the hypothesis test of equality between simulated and data coefficients is greater than 1.96, 2.57, and 5.0, respectively. The largest t -statistic for these hypothesis tests is 6.56.

estimated model fails to match only one of them (the correlation of residuals across equations for the young group).

To sum up, the estimated structural model matches several very significant coefficients of the auxiliary model quite well, but also falls short in matching the coefficients on lagged and future consumption, resulting in a statistical rejection.

6.5 Robustness

Despite many appealing features of simulation-based structural estimation methods, a drawback is that the estimates depend *potentially* on all the assumptions made on the structural model. In contrast, GMM has the advantage that any feature of the model that does not appear explicitly in the moment conditions has no relevance for the estimation results. Consequently, it is essential to examine whether the parameter estimates we obtain are sensitive to the key features of the model that have been fixed in advance. In an earlier version of the paper, we have conducted extensive robustness analyses by varying a number of key assumptions made so far. These included (i) an alternative method for filling-in missing observations, (ii) considering a higher income floor \underline{Y} , (iii) a lower interest rate, (iv) fixing (rather than estimating) the borrowing constraints, and (v) using all data available up to age 65. We found that the results were overall quite robust to these changes and none of the substantive conclusions reached in the paper were altered by these changes. For completeness, we included these results in supplemental Appendix D.

7 Discussions and Conclusions

Partial Insurance versus Advance Information: Can we Tell them Apart?

An ongoing debate in the literature attempts to understand if partial insurance can be disentangled from advance information (typically about one-period ahead income realizations) and identified using consumption and income data (Blundell et al. (2008), Kauffmann and Pistaferri (2009), and Kaplan and Violante (2010)). The general conclusion is concisely summarized by Kauffmann and Pistaferri (2009, p. 392): “[D]ata on income and consumption are not sufficient to separately identify advance information that consumers may have about their income from the extent of consumption insurance against income innovations.” It should be noted that this debate takes place squarely in the context of GMM estimation, in which case researchers exclusively rely on moments based on consumption *changes*. In this section, we revisit this question and show that while we confirm the lack of identification between insurance and advance information using consumption change alone, the levels of consumption—as used in our auxiliary model—are clearly able to separate the two. The basic argument can be explained most easily in a stylized example, although it is easy to generalize it.

For this purpose, consider a two-period model with quadratic utility, no time discounting, no

borrowing constraints, and a zero net interest rate:

$$\begin{aligned} & \max_{C_1, C_2} [-(C_1 - C^*)^2 - \mathbb{E}(C_2 - C^*)^2] \\ \text{s.t.} \quad & C_1 + C_2 = Y_1 + Y_2^*. \end{aligned}$$

Further, we assume that income follows a random walk: $Y_2 = Y_1 + \eta$. (These assumptions are made for clarity of exposition and can be easily relaxed.) We model advance information as follows. Suppose that at time 1, the individual receives a signal about his future income, so that his expectation is given by $\mathbb{E}^*(Y_2) = (1 - \alpha)Y_2 + \alpha Y_1$. (The superscript * indicates that the expectation is conditional on information beyond Y_1 .) The parameter α measures the amount of advance information. When $\alpha = 1$, there is no advance information: $\mathbb{E}^*(Y_2) = Y_1$, and when $\alpha = 0$, the signal is fully revealing: $\mathbb{E}^*(Y_2) = Y_2$. This latter case is the same as a model with no uncertainty. For partial insurance, we use the same structure as in Section 3. That is, disposable income is given by

$$Y_2^* = Y_2 - \theta(Y_2 - \mathbb{E}^*(Y_2)) = (1 - \theta)Y_2 + \theta\mathbb{E}^*(Y_2) = Y_2 - \alpha\theta(Y_1 - Y_2).$$

Optimal consumption choices can be shown to be

$$C_1 = \frac{(1 + \alpha)Y_1 + (1 - \alpha)Y_2}{2} \quad \text{and} \quad C_2 = \left[\frac{1}{2} - \alpha\left(\frac{1}{2} - \theta\right) \right] Y_1 + \left[\frac{1}{2} + \alpha\left(\frac{1}{2} - \theta\right) \right] Y_2. \quad (22)$$

The formula for C_1 does not involve θ , so observing C_1 , Y_1 , and Y_2 identifies α ; then C_2 identifies θ straightforwardly. The auxiliary model we used in this paper (19) contains current consumption on the left-hand side and the levels of both past and future income on the right-hand side, which captures precisely the type of information that is needed according to these formula.

Now, let us compute the consumption change:

$$C_2 - C_1 = \alpha(1 - \theta)(Y_2 - Y_1). \quad (23)$$

Notice that the two key parameters, α and θ , appear in (23) multiplicatively and therefore cannot be separately identified. This is precisely analogous to the results found in the literature that the response of consumption growth to income growth identifies a mixture of partial insurance and advance information.²⁹

These results suggest that we can introduce a signal (advance information) into our model that yields information about one-period-ahead income and estimate it along with the partial insurance already included. Conceptually, the framework is perfectly capable of dealing with such an analysis,

²⁹Generalizing this example to a multi-period setting would introduce assets into the level formulas, but indirect inference is easily able to deal with that by using appropriate proxies (such as, e.g., long-run averages of income) for wealth.

especially because indirect inference does not require us to rely solely on consumption growth moments, which, as seen above, cannot distinguish partial insurance from advance information. In this paper, we have not undertaken this extension, because it introduces another state variable, increasing computational complexity by an order of magnitude. However, this is high on our future research agenda.

Conclusions

The joint dynamics of consumption and labor income contain rich information about the economic environment that individuals inhabit. In this paper, we have studied how such information can be extracted from choice data to shed light on different aspects of lifetime income risk. The framework that we analyzed encompasses a number of earlier papers that also attempted to interpret the joint dynamics of consumption and income data through the lens of structural models, such as Gourinchas and Parker (2002), Guvenen (2007), Blundell et al. (2008), and Kaplan and Violante (2010). While our estimate of growth-rate heterogeneity in the population is quite similar to what was assumed in Guvenen (2007), the amount of growth-rate *uncertainty* we estimate here is much smaller than what was found in that paper. Our results on partial insurance are broadly consistent with Blundell et al. (2008)'s findings that up to one-half of persistent shocks are insured through informal channels. Our analysis also addresses Kaplan and Violante (2010)'s critique that the estimated effects of partial insurance could be partly attributable to the non-permanence of income shocks: we left the persistence of income shocks unrestricted but instead estimated it along with the rest of the model parameters.

The Monte Carlo results suggest that the indirect inference method works very well, even in the presence of frequently binding borrowing constraints, missing observations, a concave pension function, and so on, that make the auxiliary model an imperfect approximation to the structural relationships that hold in the model. While in this paper we have focused entirely on consumption-savings choices, the estimation method we use is general enough to accommodate a variety of other static or intertemporal decisions. Economic decisions that involve large fixed costs (and, hence, are made infrequently, such as fertility choice, house purchases, etc.) are likely to be especially forward-looking and therefore useful for inferring the nature and amount of risk. We believe that the indirect inference methodology used in this paper can be fruitfully used in these alternative implementations.

Substantively, we find that (i) income shocks have moderate persistence—much less than a unit root; (ii) income growth rates display significant cross-sectional heterogeneity; (iii) individuals have much better information about their own income growth rates than what can be predicted by the observable variables typically available to the econometrician; and, finally, (iv) there is substantial amounts of partial insurance available to households, over and above what they can achieve on their own through self-insurance. Combining these pieces, *the main conclusion of our analysis* is that the amount of uninsurable lifetime income risk that households perceive is much smaller than what is typically assumed in calibrated macroeconomic models with incomplete markets.

References

- Aguiar, Mark and Erik Hurst**, “Deconstructing Lifecycle Expenditures,” *Journal of Political Economy*, forthcoming.
- Altug, Sumru and Robert A Miller**, “Household Choices in Equilibrium,” *Econometrica*, May 1990, *58* (3), 543–70.
- Bansal, Ravi, A. Ronald Gallant, and George Tauchen**, “Rational Pessimism, Rational Exuberance, and Asset Pricing Models,” *The Review of Economic Studies*, October 2007, *74* (4), 1005–1033.
- Ben-Porath, Yoram**, “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, 1967, *75* (4), 352–365.
- Blundell, Richard and Ian Preston**, “Consumption Inequality And Income Uncertainty,” *The Quarterly Journal of Economics*, May 1998, *113* (2), 603–640.
- , **Luigi Pistaferri, and Ian Preston**, “Imputing consumption in the PSID using food demand estimates from the CEX,” October 2006, (W04/27).
- , – , and – , “Consumption Inequality and Partial Insurance,” *American Economic Review*, December 2008, *98* (5), 1887–1921.
- Carneiro, Pedro, Karsten Hansen, and James Heckman**, “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College,” *International Economic Review*, 2003, *24* (2), 361–421.
- Collard-Wexler, Allan**, “Demand Fluctuations in the Ready-Mix Concrete Industry,” *Econometrica*, May 2013, *81* (3), 1003–1037.
- Cunha, Flavio, James Heckman, and Salvador Navarro**, “Separating Uncertainty from Heterogeneity in Life Cycle Earnings,” *Oxford Economic Papers*, 2005, *57* (2), 191–261.
- Deaton, Angus**, “Saving and Liquidity Constraints,” *Econometrica*, September 1991, *59* (5), 1221–48.
- and **Christina Paxson**, “Intertemporal Choice and Inequality,” *Journal of Political Economy*, June 1994, *102* (3), 437–67.
- Gourieroux, Christian, Alain Monfort, and Eric Renault**, “Indirect Inference,” *Journal of Applied Economics*, 1993, *8* (Supplement), S85–S118.
- Gourinchas, Pierre-Olivier and Jonathan A. Parker**, “Consumption over the Life Cycle,” *Econometrica*, 2002, *70* (1), 47–89.

- Guvenen, Fatih**, “Learning Your Earning: Are Labor Income Shocks Really Very Persistent?,” *American Economic Review*, June 2007, *97* (3), 687–712.
- , “An Empirical Investigation of Labor Income Processes,” *Review of Economic Dynamics*, January 2009, *12* (1), 58–79.
- **and Burhanettin Kuruscu**, “A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000,” in “NBER Macroeconomics Annual 2009, Volume 24,” University of Chicago Press, September 2010, pp. 227–276.
- , **Fatih Karahan, Serdar Ozkan, and Jae Song**, “A User’s Guide to Idiosyncratic Income Risk,” Working Paper, University of Minnesota 2012.
- Hall, Robert E.**, “Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence,” *The Journal of Political Economy*, 1978, *86* (6), 971–987.
- Hall, Robert E and Frederic S Mishkin**, “The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households,” *Econometrica*, March 1982, *50* (2), 461–81.
- Hayashi, Fumio, Joseph Altonji, and Laurence Kotlikoff**, “Risk-Sharing between and within Families,” *Econometrica*, March 1996, *64* (2), 261–94.
- Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante**, “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006,” *Review of Economic Dynamics*, 2010, *13* (1), 15–51.
- Huggett, Mark, Gustavo Ventura, and Amir Yaron**, “Sources of Lifetime Inequality,” *American Economic Review*, December 2011, *101* (7), 2923–54.
- Johnston, David and Stephanie Shipp**, “Trends in Inequality Using Consumption-Expenditures: The US from 1960 to 1993,” *Review of Income and Wealth*, 1997, *43*, 133–152.
- Kaplan, Greg**, “Inequality and the Life Cycle,” *Quantitative Economics*, 2012, *3* (3), 471–525.
- **and Giovanni L. Violante**, “How Much Consumption Insurance Beyond Self-Insurance?,” *American Economic Journal: Macroeconomics*, 2010, *2* (4), 53–87.
- Kauffmann, Katja and Luigi Pistaferri**, “Disentangling Insurance and Information in Intertemporal Consumption Choice,” *American Economic Review P&P*, 2009, *99* (2), 387–392.
- Krueger, Dirk and Fabrizio Perri**, “How Do Households Respond to Income Shocks?,” Technical Report, University of Minnesota 2009.
- Li, Tong**, “Indirect inference in structural econometric models,” *Journal of Econometrics*, July 2010, *157* (1), 120–128.

- Low, Hamish and Luigi Pistaferri**, “Disability Insurance and the Dynamics of the Incentive-Insurance Tradeoff,” Technical Report, Stanford University 2012.
- Magnac, Thierry, Jean-Marc Robin, and Michael Visser**, “Analysing Incomplete Individual Employment Histories Using Indirect Inference,” *Journal of Applied Econometrics*, December 1995, 10, S153–S169.
- Nagypal, Eva**, “Learning by Doing vs. Learning about Match Quality: Can We Tell Them Apart?,” *The Review of Economic Studies*, April 2007, 74 (2), 537–566.
- Primiceri, Giorgio E. and Thijs van Rens**, “Heterogeneous life-cycle profiles, income risk and consumption inequality,” *Journal of Monetary Economics*, 2009, 56 (1), 20 – 39. Carnegie-Rochester Conference Series on Public Policy: The Causes and Consequences of Rising Economic Inequality April 25-26, 2008.
- Slesnick, Daniel T.**, “Aggregate Consumption and Saving in the Postwar United States,” *The Review of Economics and Statistics*, 1992, 74 (4), 585–597.
- Smith, Anthony A.**, “Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions,” *Journal of Applied Econometrics*, 1993, 8, S63–S84.
- Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron**, “Cyclical Dynamics in Idiosyncratic Labor Market Risk,” *Journal of Political Economy*, June 2004, 112 (3), 695–717.
- , **Christopher I. Telmer, and Amir Yaron**, “Consumption and risk sharing over the life cycle,” *Journal of Monetary Economics*, April 2004, 51 (3), 609–633.

NOT FOR PUBLICATION:
SUPPLEMENTAL APPENDIX

A Derivations and Proofs

A.1 Kalman Filtering Equations

The recursive updating formulas are given by

$$\widehat{\mathbf{S}}_t^i = \widehat{\mathbf{S}}_{t|t-1}^i + \mathbf{K}_t \times \widehat{\xi}_t^i \quad (24)$$

$$\mathbf{P}_t = (I - \mathbf{K}_t \mathbf{H}_t') \times \mathbf{P}_{t|t-1}, \quad (25)$$

where $\widehat{\mathbf{S}}_{t|t-1}^i$ and $\mathbf{P}_{t|t-1}$ denote the one-period-ahead *forecasts* of these two variables and obtained from beliefs as $\widehat{\mathbf{S}}_{t|t-1}^i = \mathbf{F} \widehat{\mathbf{S}}_{t-1}^i$, and $\mathbf{P}_{t|t-1} = \mathbf{F} \mathbf{P}_{t-1} \mathbf{F}' + \mathbf{Q}$; and $\mathbf{K}_t \equiv \mathbf{P}_{t|t-1} \mathbf{H}_t' [\mathbf{H}_t' \mathbf{P}_{t|t-1} \mathbf{H}_t]^{-1}$ is the (optimal) Kalman gain. Then, \tilde{y}_t^i has a Normal distribution conditional on an individual's beliefs at time $t - 1$:

$$\tilde{y}_t^i | \widehat{\mathbf{S}}_{t-1}^i \sim \mathcal{N} \left(\mathbf{H}_t' \widehat{\mathbf{S}}_{t-1}^i, \mathbf{H}_t' \mathbf{P}_{t-1} \mathbf{H}_t \right). \quad (26)$$

For the problem at hand, (24) and (25) can be manipulated to obtain some simple expressions. First, (24) simplifies to

$$\widehat{\beta}_t^i - \widehat{\beta}_{t-1}^i = (A_t/X_t) \widehat{\xi}_t^i, \quad (27)$$

$$\widehat{z}_t^i - \rho \widehat{z}_{t-1}^i = (B_t/X_t) \widehat{\xi}_t^i, \quad (28)$$

where $A_t \equiv t\sigma_{\beta,t|t-1}^2 + \sigma_{\beta z,t|t-1}$, $B_t \equiv t\sigma_{\beta z,t|t-1} + \sigma_{z,t|t-1}^2$, and $X_t \equiv \text{var}_{t-1}(y_t^i) = A_t + B_t$. Notice that A_t/X_t and B_t/X_t measure the fraction of the (one-step-ahead) forecast variance that is due to the slope uncertainty and the persistent shock, respectively. Thus, a given $\widehat{\xi}_t^i$ is split according to the perceived share of variance attributed to each component. Second, (25) reduces to

$$\sigma_{\beta,t+1|t}^2 = \sigma_{\beta,t|t-1}^2 - \frac{A_t^2}{X_t} \quad (29)$$

$$\sigma_{z,t+1|t}^2 = \rho^2 \left[\sigma_{z,t|t-1}^2 - \frac{B_t^2}{X_t} \right] + \sigma_\eta^2. \quad (30)$$

As shown by Guvenen (2007), an important feature of Bayesian learning in this framework is that beliefs about β^i change non-monotonically over the life cycle, owing to the inverse U-shape pattern followed by A^2/X . Consequently, the uncertainty regarding β^i can be very slow to resolve. If, instead, the prior uncertainty were to resolve quickly, consumption behavior after the first few years would not be informative about the prior uncertainty faced by individuals ($\widehat{\sigma}_{\beta|0}^2$).

A.2 Proofs of Propositions 1 and 2

We begin by establishing the following lemma that comes in handy in the proofs that follow.

Lemma A.1 *Assume $\sigma_\beta > 0$. Then $\partial \Pi_t / \partial \lambda > 0$.*

Proof. First recall that $\Pi_t = \Phi_t(A_t/X_t) + \Psi_t(B_t/X_t)$, with

$$\Phi(t; T, r) \equiv \left[\frac{\gamma}{(1-\gamma)} + \frac{t - (T+1)\gamma^{T-t+1}}{(1-\gamma^{T-t+1})} \right]$$

and

$$\Psi(t; T, \rho, r) \equiv \frac{1-\gamma}{1-\gamma\rho} \left[\frac{(1-(\gamma\rho)^{T-t+1})}{(1-\gamma^{T-t+1})} \right].$$

The dependence of Π_t on λ comes through A_t, B_t , and X_t , which all contain elements of $\mathbf{P}_{t+1|t}$, which in turn depend on λ . Thus, to establish $\partial\Pi_t/\partial\lambda > 0$, we first iterate on the recursions for updating the posterior covariance matrix (equation (25)). Specifically, start at $t = 0$: $\sigma_{\beta,1|0}^2 = \lambda^2\sigma_\beta^2$, $\sigma_{z,1|0}^2 = \sigma_\eta^2$, and $\sigma_{\beta z,1|0} = 0$. Then we can find $A_1 = \lambda^2\sigma_\beta^2$, $B_1 = \sigma_\eta^2$, and $X_1 = \lambda^2\sigma_\beta^2 + \sigma_\eta^2$. Plugging these expressions into (29) and (30), we obtain

$$\sigma_{\beta,2|1}^2 = \frac{\lambda^2\sigma_\beta^2\sigma_\eta^2}{\lambda^2\sigma_\beta^2 + \sigma_\eta^2},$$

$$\sigma_{z,2|1}^2 = \rho^2 \frac{\lambda^2\sigma_\beta^2\sigma_\eta^2}{\lambda^2\sigma_\beta^2 + \sigma_\eta^2} + \sigma_\eta^2,$$

and

$$\sigma_{\beta z,2|1}^2 = \rho \frac{\lambda^2\sigma_\beta^2\sigma_\eta^2}{\lambda^2\sigma_\beta^2 + \sigma_\eta^2}.$$

After a few iterations like this, the following recursive formulas emerge for the variances and covariances:

$$\sigma_{\beta,t+1|t}^2 = K_t \tag{31}$$

$$\sigma_{z,t+1|t}^2 = t^2\rho^2 K_t + \sigma_\eta^2 \tag{32}$$

$$\sigma_{\beta z,t+1|t} = -t\rho K_t, \tag{33}$$

where $K_t \equiv \frac{\Theta\sigma_\eta^2}{\Theta(\sum_{s=0}^t (s-(s-1)\rho)^2 - 1) + \sigma_\eta^2}$ and $\Theta \equiv \frac{\lambda^2\sigma_\beta^2\sigma_\eta^2}{\lambda^2\sigma_\beta^2 + \sigma_\eta^2}$.

Now, first, it is straightforward to show that $\Pi_t(\lambda = 0) > 0$. To see this, observe that if $\lambda = 0$, then $\Theta = 0$ and $K_t = 0$ for all t . So, $\sigma_{\beta,t|t-1}^2 = \sigma_{\beta z,t|t-1} = 0$ and $\sigma_{z,t|t-1}^2 = \sigma_\eta^2$. It follows that $A_t = 0$, $B_t = \sigma_\eta^2$, and $X_t = \sigma_\eta^2$. Plugging in these values shows that $\Pi_t = \Psi(t; T, r, \rho)$, which is always positive. Second, to show that $\frac{\partial\Pi_t}{\partial\lambda} > 0$, we need to calculate the derivatives of A_t/X_t and B_t/X_t with respect to λ . First, we have

$$\frac{\partial\Theta}{\partial\lambda} = \frac{2\lambda^3\sigma_\beta^2(\sigma_z^2)^2}{(\lambda^2\sigma_\beta^2 + \sigma_z^2)^2} > 0,$$

$$\frac{\partial K_t}{\partial\lambda} = \frac{\frac{\partial\Theta}{\partial\lambda}(\sigma_\eta^2)^2}{(\Theta(\sum_{s=0}^t (s-(s-1)\rho)^2 - \rho^2) + \sigma_\eta^2)^2} > 0.$$

Using the chain rule and $\frac{\partial K_t}{\partial\lambda} > 0$ for all t , we find

$$\frac{\partial(A_t/X_t)}{\partial\lambda} = (t - \rho(t-1))K_{t-1} \frac{\partial K_{t-1}}{\partial\lambda} \sigma_\eta^2 > 0$$

$$\frac{\partial(B_t/X_t)}{\partial\lambda} = -t(t - \rho(t - 1))K_{t-1} \frac{\partial K_{t-1}}{\partial\lambda} \sigma_\eta^2 < 0.$$

We can rewrite the derivative of Π_t with respect to λ as

$$\frac{\partial\Pi_t}{\partial\lambda} = [\Phi_t \frac{\partial(A_t/X_t)}{\partial\lambda} + \Psi_t \frac{\partial(B_t/X_t)}{\partial\lambda}] \quad (34)$$

$$= [\Phi_t - t\Psi_t](t - \rho(t - 1))K_{t-1} \frac{\partial K_{t-1}}{\partial\lambda} \sigma_\eta^2. \quad (35)$$

Note that all terms outside of the square brackets are positive. Thus, $\frac{\partial\Pi_t}{\partial\lambda} > 0$ if and only if $[\Phi_t - t\Psi_t] > 0$. To prove the latter, we proceed in two steps. First, the expression we are interested in is

$$\Phi_t - t\Psi_t = \left[\frac{\gamma}{(1 - \gamma)} + \frac{t - (T + 1)\gamma^{T-t+1}}{(1 - \gamma^{T-t+1})} \right] - t \left[\frac{1 - \gamma}{1 - \gamma^{T-t+1}} \frac{1 - (\gamma\rho)^{T-t+1}}{1 - \gamma\rho} \right]. \quad (36)$$

It is straightforward to see that $\partial(\Phi_t - t\Psi_t)/\partial\rho < 0$, since ρ only appears in the second set of brackets (i.e., Ψ_t), which clearly becomes more negative as ρ rises.³⁰ Therefore, it is sufficient to prove that $\Phi_t - t\Psi_t > 0$ when $\rho = 1$ and the same will hold for all values of $\rho < 1$. This is how we shall proceed. Let the remaining planning horizon of an individual be denoted with $\tau \equiv T - t + 1$. When $\rho = 1$, we have $\Psi_t = 1$ and the expression simplifies to

$$\Phi_t - t\Psi_t = \left[\frac{\gamma}{(1 - \gamma)} + \frac{t - (T + 1)\gamma^{T-t+1}}{(1 - \gamma^{T-t+1})} \right] - t = \frac{\gamma}{(1 - \gamma)} - \frac{\tau\gamma^\tau}{(1 - \gamma^\tau)}. \quad (37)$$

For $\tau = 1$, the expression equals zero. All we need to show is that the derivative of this expression with respect to τ is negative, which will then establish that $\Phi_t - t\Psi_t < 0$ for all $\tau > 1$. This is easy to do:

$$\frac{d}{d\tau} \left(\frac{\tau\gamma^\tau}{(1 - \gamma^\tau)} \right) = \frac{\gamma^\tau(1 + \tau \log \gamma - \gamma^\tau)}{(1 - \gamma^\tau)^2} < 0 \Leftrightarrow 1 + \tau \log \gamma - \gamma^\tau < 0 \Leftrightarrow \gamma^\tau < e^{\gamma^\tau - 1},$$

which is satisfied for all $\tau > 1$ as long as $\gamma < 1$ (i.e., $r > 0$). Since the second term in (37) is decreasing with the horizon, this establishes that $\Phi - t\Psi > 0$ for all $\tau > 1$ (alternatively $t < T$). ■

Proof of Proposition 1:

Part (i): Rewrite (9) as

$$\begin{aligned} \Delta C_t^i &= \Pi_t \times (Y_t^i - \mathbb{E}_{t-1}(Y_t^i)) = \Pi_t \times (\Delta Y_t^i + Y_{t-1}^i - \mathbb{E}_{t-1}(Y_t^i)) \\ &= \Pi_t \times \Delta Y_t^i + \Pi_t \times \left(\beta^i(t - 1) + z_{t-1}^i - (\widehat{\beta}_{t-1}^i t + \widehat{z}_{t-1}^i) \right) \\ &= \Pi_t \times \Delta Y_t^i + \Pi_t \times \left((\beta^i - \widehat{\beta}_{t-1}^i) t + (z_{t-1}^i - \widehat{z}_{t-1}^i) - \beta^i \right) \end{aligned}$$

Taking the expectations of both sides with respect to the history up to time $t - 1$ (of prior beliefs and

³⁰To see this, note that the ratio $\frac{1 - (\gamma\rho)^{T-t+1}}{1 - \gamma\rho}$ can be expanded as $1 + (\gamma\rho) + (\gamma\rho)^2 + \dots + (\gamma\rho)^{T-t}$, which is clearly increasing in ρ .

income realizations, $\{Y_1^i, Y_2^i, \dots, Y_{t-1}^i, \hat{\beta}_{1|0}^i\}$) conditional on $\beta^i, \Delta Y_t^i$:

$$\begin{aligned}\mathbb{E}(\Delta C_t^i | \beta^i, \Delta Y_t^i) &= \Pi_t \times \left(\Delta Y_t^i + \mathbb{E} \left(\beta^i - \hat{\beta}_{t-1}^i | \beta^i, \Delta Y_t^i \right) \times t \right) \\ &\quad + \Pi_t \times \mathbb{E} \left(z_{t-1}^i - \hat{z}_{t-1}^i | \beta^i, \Delta Y_t^i \right) - \Pi_t \times \beta^i \\ \mathbb{E}(\Delta C_t^i | \beta^i, \Delta Y_t^i) &= \Pi_t \times \Delta Y_t^i + \Pi_t \times (\beta^i - \beta^i) \times t \\ &\quad + \Pi_t \times (0) - \Pi_t \times \beta^i.\end{aligned}$$

On the last line, we made use of two facts: $\mathbb{E} \left(z_{t-1}^i - \hat{z}_{t-1}^i | \beta^i, \Delta Y_t^i \right) = 0$ and $\mathbb{E} \left(\hat{\beta}_{t-1}^i | \beta^i, \Delta Y_t^i \right) = \beta^i$, which yields

$$\mathbb{E}(\Delta C_t^i | \beta^i, \Delta Y_t^i) = \Pi_t \times \Delta Y_t^i - \Pi_t \times \beta^i. \quad (38)$$

Therefore, controlling for income growth, on average, consumption growth is decreasing in β^i . Furthermore, since $\partial \mathbb{E}(Y_{t-1}^i - Y_1^i | \beta^i) / \partial \beta^i > 0$, consumption growth is also decreasing in past income growth.

Part (ii): We need to establish three results. First the negative dependence proved above holds even when $\lambda = 0$, that is, when the individual has full information about his/her β^i . Second, the strength of consumption's response to past income growth becomes stronger (i.e, becomes more negative) as λ rises. From the expression for $\Delta \bar{C}_t^i$ given in (38), this is equivalent to showing $\partial \Pi_t / \partial \lambda > 0$, which was proved in Lemma 1. ■

Proof of Lemma 1. From the solution of the model, we know that consumption equals the annuity value of the physical wealth and expected lifetime discounted labor income:

$$C_t = \varphi_t \left[\frac{1}{\gamma} A_t + \sum_{s=0}^{T-t} \gamma^s \mathbb{E}_t(Y_{t+s}) \right]. \quad (39)$$

Taking the expectation of the income process, $Y_t^i = \alpha^i + \beta^i t + z_t^i$, we find $E_t(Y_{t+s}^i) = \alpha^i + \hat{\beta}_t^i(t+s) + \rho^s \hat{z}_t^i$. Plugging (3) in (39) yields

$$C_t^i = \varphi_t \left[\frac{1}{\gamma} A_t^i + \sum_{s=0}^{T-t} \gamma^s (\alpha^i + \hat{\beta}_t^i(t+s) + \rho^s \hat{z}_t^i) \right] \quad (40)$$

$$= \varphi_t \left[\frac{1}{\gamma} A_t^i + \underbrace{(\alpha^i + \hat{\beta}_t^i t + \hat{z}_t^i)}_{Y_t^i} + \sum_{s=1}^{T-t} \gamma^s (\alpha^i + \hat{\beta}_t^i(t+s) + \rho^s \hat{z}_t^i) \right] \quad (41)$$

$$\Rightarrow C_t^i = \varphi_t \left(\underbrace{\omega_t^i}_{\text{Cash on hand}} \right) + \gamma \Phi(t+1; T, r) \hat{\beta}_t^i + \gamma \rho \Psi(t+1; T, r, \rho) \hat{z}_t^i, \quad (42)$$

which is equation (11) in Lemma 1.

A.3 Partial Insurance

Following the same steps in the proof of Lemma 1 above and replacing Y_t with Y_t^{disp} yields the following expression for consumption growth in the presence of partial insurance:

$$\Delta C_t = \hat{\xi}_t \left\{ \frac{A_t}{X_t} [\Phi(t; T, r) - t\varphi_t\theta] + \frac{B_t}{X_t} [\Psi(t; T, r) - \varphi_t\theta] \right\}.$$

This expression can be further simplified by rearranging terms and recognizing that $A_t t + B_t \equiv X_t$. We get

$$\Delta C_t = (\Pi_t - \varphi_t\theta) \times \hat{\xi}_t.$$

A.4 Likelihood Approach vs. Quadratic Objective: An Equivalence

Here we establish the asymptotic equivalence between the “likelihood approach” to indirect inference employed in our estimation and the quadratic objective—also called the Wald approach—that is often used in the literature. We prove the equivalence for a stylized case for clarity, although it will become clear that the proof can easily be extended to allow more general structural models (with a vector of exogenous variables, X_t , as well as more lags and leads of variable Y). Now, consider the structural (i.e., “true”) model:

$$Y_t = f(Y_{t-1}, \beta) + \epsilon_t,$$

where $\epsilon_t \sim \text{iid } \mathcal{N}(0, \sigma^2)$, σ^2 is known, and Y_0 is given. Consider the auxiliary model: $Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \eta_t$, $\eta_t \sim \text{i.i.d. } \mathcal{N}(0, 1)$. The auxiliary model likelihood is $-\sum_{t=1}^T (Y_t - \gamma_0 - \gamma_1 Y_{t-1})^2$. Define

$$\hat{h}_i(\beta) \equiv (\hat{\gamma}_{0,i}(\beta), \hat{\gamma}_{1,i}(\beta)) = \underset{\gamma_0, \gamma_1}{\operatorname{argmin}} \sum_{t=1}^T (Y_t^i(\beta) - \gamma_0 - \gamma_1 Y_{t-1}^i(\beta))^2,$$

where i denotes the i^{th} simulated data set, given β . Now define

$$\hat{h}_M(\beta) \equiv \underset{\gamma_0, \gamma_1}{\operatorname{argmin}} \sum_{i=1}^M \sum_{t=1}^T (Y_t^i(\beta) - \gamma_0 - \gamma_1 Y_{t-1}^i(\beta))^2,$$

as $M \rightarrow \infty$ (holding T fixed), $\hat{h}_M(\beta) \rightarrow h(\beta)$, where

$$h(\beta) \equiv \underset{\gamma_0, \gamma_1}{\operatorname{argmin}} \mathbb{E} \sum_{t=1}^T (Y_t(\beta) - \gamma_0 - \gamma_1 Y_{t-1}(\beta))^2.$$

The approach in this paper is (assuming M is large) to calculate

$$\hat{\beta}_T = \min_{\beta} \sum_{t=1}^T (Y_t - \gamma_0(\beta) - \gamma_1(\beta) Y_{t-1})^2,$$

where $\{Y_t\}_{t=0}^T$ is the observed data. The first-order condition is

$$\begin{aligned}
0 &= \sum_t (Y_t - \gamma_0(\beta) - \gamma_1(\beta) Y_{t-1}) \gamma'_0(\beta) \\
&\quad + \sum_t (Y_t - \gamma_0(\beta) - \gamma_1(\beta) Y_{t-1}) \gamma'_1(\beta) Y_{t-1} \\
&= -\gamma'_0(\beta) \sum Y_t + \gamma_0(\beta) \gamma'_0(\beta) T + \gamma_1(\beta) \gamma'_0(\beta) \sum Y_{t-1} \\
&\quad - \gamma'_1(\beta) \sum Y_t Y_{t-1} + \gamma_0(\beta) \gamma'_1(\beta) \sum Y_{t-1} + \gamma_1(\beta) \gamma'_1(\beta) \sum Y_{t-1}^2,
\end{aligned} \tag{43}$$

where $\gamma'_j(\beta)$ is the derivative of γ_j , $j = 0, 1$. Now, as an alternative, consider minimizing the following quadratic form:

$$\begin{bmatrix} \gamma_0(\beta) - \hat{\gamma}_0 \\ \gamma_1(\beta) - \hat{\gamma}_1 \end{bmatrix}' \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \gamma_0(\beta) - \hat{\gamma}_0 \\ \gamma_1(\beta) - \hat{\gamma}_1 \end{bmatrix},$$

where $\hat{\gamma}_T \equiv \operatorname{argmin}_{\gamma_0, \gamma_1} \sum_{t=1}^T (Y_t - \gamma_0 - \gamma_1 Y_{t-1})^2$. The F.O.C. with respect to β is:

$$\begin{aligned}
&a_{11} (\gamma_0(\beta) - \hat{\gamma}_0) \gamma'_0(\beta) + a_{12} (\gamma_0(\beta) - \hat{\gamma}_0) \gamma'_1(\beta) + a_{12} (\gamma_1(\beta) - \hat{\gamma}_1) \gamma'_0(\beta) + a_{22} (\gamma_1(\beta) - \hat{\gamma}_1) \gamma'_1(\beta) = \\
&= (-a_{11} \hat{\gamma}_0 - a_{12} \hat{\gamma}_1) \gamma'_0(\beta) - (a_{12} \hat{\gamma}_0 + a_{22} \hat{\gamma}_1) \gamma'_1(\beta) + a_{11} \gamma_0(\beta) \gamma'_0(\beta) + a_{12} \gamma_0(\beta) \gamma'_1(\beta) \\
&\quad + a_{12} \gamma'_0(\beta) \gamma_1(\beta) + a_{22} \gamma_1(\beta) \gamma'_1(\beta) = 0.
\end{aligned} \tag{44}$$

We want to make (44) look like condition (43). To do so, first set

$$a_{11} = T, a_{12} = \sum Y_{t-1}, a_{22} = \sum Y_{t-1}^2.$$

Then the last four terms in (44) match four of the six terms in (43). But what about the remaining two terms in each equation? One can show that these terms match up asymptotically, as the observed sample size T grows large. To see this:

$$\begin{aligned}
\operatorname{plim} \left[\gamma'_0(\hat{\beta}_T) \left(\frac{\sum Y_t}{T} - \frac{T \hat{\gamma}_0(\hat{\beta}_T)}{T} - \left(\frac{\sum Y_{t-1}}{T} \right) \hat{\gamma}_1 \right) \right] &= \gamma'_0(\beta_0) \underbrace{(\mathbb{E}Y_t - \gamma_0(\beta_0) - (\mathbb{E}Y_{t-1}) \gamma_1(\beta_0))}_{=0} \\
&= \gamma'_0(\beta_0) \times 0 = 0,
\end{aligned}$$

where β_0 is the “true” value of β (i.e., $\operatorname{plim} \hat{\beta}_T = \beta_0$). The terms in the set of parentheses on the right-hand side are zero asymptotically as $T \rightarrow \infty$ because it is simply the F.O.C. that defines $\gamma_0(\beta_0)$ ($= \operatorname{plim} \hat{\gamma}_0$). Similarly,

$$\begin{aligned}
&\operatorname{plim} \left[\gamma'_1(\hat{\beta}_T) \left(T^{-1} \sum Y_t Y_{t-1} - \left(T^{-1} \sum Y_{t-1} \right) \hat{\gamma}_0 - \left(T^{-1} \sum Y_{t-1}^2 \right) \hat{\gamma}_1 \right) \right] \\
&= \gamma'_1(\beta_0) (\mathbb{E}Y_t Y_{t-1} - (\mathbb{E}Y_{t-1}) \gamma_0(\beta_0) - (\mathbb{E}Y_{t-1}^2) \gamma_1(\beta_0)) = 0,
\end{aligned}$$

again, because the second term is (asymptotically) the f.o.c. that defines $\gamma_1(\beta_0) = \operatorname{plim} \hat{\gamma}_1$. This shows that equations (43) and (44) are asymptotically equivalent, completing the proof.

To summarize, the LR approach—the approach we are currently using—is asymptotically equivalent, in this simplified case, to minimizing the following quadratic form:

$$\begin{bmatrix} \gamma_0(\beta) - \hat{\gamma}_0 \\ \gamma_1(\beta) - \hat{\gamma}_1 \end{bmatrix}' \begin{bmatrix} 1 & \mathbb{E}Y_{t-1} \\ \mathbb{E}Y_{t-1} & \mathbb{E}Y_{t-1}^2 \end{bmatrix} \begin{bmatrix} \gamma_0(\beta) - \hat{\gamma}_0 \\ \gamma_1(\beta) - \hat{\gamma}_1 \end{bmatrix}.$$

Note that the weighting matrix would be the optimal one if the auxiliary model were correctly specified because it is proportional to the inverse of the asymptotic covariance matrix of

$$T^{1/2} \begin{bmatrix} \gamma_0(\beta_0) - \hat{\gamma}_0 & \gamma_1(\beta_0) - \hat{\gamma}_1 \end{bmatrix}',$$

where β_0 is the true value of β .

B Details of the Econometric Procedure

B.1 Implementation: A Gaussian Objective Function

This section describes the details of how we implement the indirect inference estimator. Loosely speaking, the indirect inference estimator is obtained by choosing the values of the structural parameters so that the estimated model and the US data look as similar as possible when viewed through the lens of the auxiliary model. More concretely, define

$$\boldsymbol{\epsilon}_t^{i,Data} \equiv \begin{bmatrix} c_t^{i,Data} - \mathbf{a}' \mathbf{X}_{c,t}^{i,Data}, y_t^{i,Data} - \mathbf{b}' \mathbf{X}_{y,t}^{i,Data} \end{bmatrix}$$

to be the residuals of estimated equations (19) and (20), which is understood to equal zero when data for household i in year t are missing. The superscript *Data* specifies the data source used in the regression, which is either the PSID or the structural model (indicated by SIM). The objective function we use is

$$\mathcal{L}(\mathbf{a}, \mathbf{b}, \Sigma, Data) = |\Sigma|^{-J/2} \exp \left(-\frac{1}{2} \sum_{i=1}^{2235} \sum_{t=1968}^{1993} \boldsymbol{\epsilon}_t^{i,Data} \Sigma^{-1} (\boldsymbol{\epsilon}_t^{i,Data})' \right), \quad (45)$$

where J is the total number of household-year observations used in the regressions (26,411 in the baseline estimation). Although the objective function is in the form of a multivariate Gaussian density, it is not, strictly speaking, the likelihood of the auxiliary model regressions (19) and (20). This is because these equations have as regressors both the past *and* the future values of endogenous variables, which makes it impossible to obtain the proper likelihood by conditioning on past observations (or the future separately). Thus, to avoid a confusion of terminology, and for lack of a better term, we shall refer to \mathcal{L} as a “Gaussian objective function.”

To implement the estimator, we first maximize the objective function in (45) using real data (i.e., from the PSID) to obtain a set of reduced-form parameters, denoted by $(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\Sigma})$. Next, we follow a similar procedure using simulated data. The vector of structural parameters that we estimate is

$$\boldsymbol{\theta} \equiv (\sigma_\alpha, \sigma_\beta, corr_{\alpha\beta}, \rho, \sigma_\eta, \sigma_\varepsilon; \lambda, \sigma_y, \mu_c, \sigma_c, \sigma_{c_0}; \delta, \psi).$$

For a given θ , simulate a data set from the structural model that matches exactly the number of observations and missing data pattern found in the PSID data set, and estimate (19) and (20), which yields $\tilde{\mathbf{a}}_1$, $\tilde{\mathbf{b}}_1$, and $\tilde{\Sigma}_1$. Now, using a fresh sequence of random draws (for all the stochastic elements in the structural model), repeat the same procedure to obtain $\tilde{\mathbf{a}}_2$, $\tilde{\mathbf{b}}_2$, and $\tilde{\Sigma}_2$. Repeat this N^{SIM} times and construct the following averages:

$$\tilde{\mathbf{a}} = \frac{1}{N^{SIM}} \sum_{n=1}^{N^{SIM}} \tilde{\mathbf{a}}_n,$$

and analogously for $\tilde{\mathbf{b}}$ and $\tilde{\Sigma}$. Then, we use these averaged parameter values—estimated from *simulated* data—to evaluate the objective function (45) using the observed (PSID) data:

$$\mathcal{L}(\tilde{\mathbf{a}}(\theta, SIM), \tilde{\mathbf{b}}(\theta, SIM), \tilde{\Sigma}(\theta, SIM), PSID).$$

If the simulated data look *exactly* like the PSID data—in the sense that the estimated auxiliary model parameters for the two data sets are identical—then the two objective values would be identical; otherwise the Gaussian objective function will always be higher when evaluated at $(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\Sigma})$ than $(\tilde{\mathbf{a}}, \tilde{\mathbf{b}}, \tilde{\Sigma})$ because the latter does not maximize (45) with PSID data (but instead with simulated data). Finally, the indirect inference estimator is defined as

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left[\mathcal{L}(\hat{\mathbf{a}}, \hat{\mathbf{b}}, \hat{\Sigma}, PSID) - \mathcal{L}(\tilde{\mathbf{a}}(\theta, SIM), \tilde{\mathbf{b}}(\theta, SIM), \tilde{\Sigma}(\theta, SIM), PSID) \right]. \quad (46)$$

In effect, our indirect inference estimator maximizes the Gaussian objective function associated with the auxiliary model subject to the “cross-equation” restrictions that the structural model imposes on its parameters. An important advantage of this estimator is that it obviates the need to estimate an optimal weighting matrix; obtaining precise estimates of such matrices is often difficult. Instead, our estimator uses an *implicit* weighting matrix that is close to optimal (to the extent that the auxiliary model is close to being correctly specified) and delivers very good small sample results. In particular, in Appendix A.4, we show that this estimator is asymptotically equivalent to one that minimizes a quadratic form in the difference between the auxiliary model parameters calculated using the observed and simulated data, with the weighting matrix being the optimal one if the auxiliary model were actually correctly specified. This weighting matrix is not optimal here (since the auxiliary model is not an exact “reduced form” for the structural model), but our Monte Carlo analysis in the next section demonstrates that we obtain excellent results, with little bias and small standard errors.

Computation of Model Specification Test Statistic. We first generate a simulated data set from the structural model by setting the parameter values to those obtained in the actual benchmark estimation. Call this the “real” data set. We evaluate the Gaussian objective function (\mathcal{L}) given in (45) using this real data set. Then, using a new set of seeds for the random number generators, we simulate a new data set and estimate the parameters that must have generated this newly simulated data. We reevaluate \mathcal{L} using these estimated parameters and the real data simulated in the first step. We repeat this second step a large number of times, which gives us a probability distribution for the test statistic under the null hypothesis that the “real” data are generated from the estimated structural model.

Specifics of the Filling-in Procedure. Basically, at each age that a household has a valid income data point, we find the percentile ranking of this observation in the income distribution (at that age) in our

Table A.1: MONTE CARLO ANALYSIS

	Using Y+C Data			Using Y Data		Using Y Data		
	“True” Value 1 (1)	Estimates Mean Std. Err. (2) (3)		Estimates Mean Std. Err. (4) (5)		“True” Value 2 (6)	Estimates Mean Std. Err. (7) (8)	
<i>Income Processes Parameters</i>								
σ_α	0.288	0.293	0.018	0.285	0.031	0.298	0.301	0.038
σ_β	1.764	1.735	0.136	1.834	0.220	1.343	1.377	0.271
$corr_{\alpha\beta}$	-0.127	-0.106	0.099	-0.161	0.173	0.558	0.531	0.289
ρ	0.764	0.766	0.024	0.754	0.027	0.783	0.780	0.022
σ_η	0.227	0.227	0.008	0.196	0.005	0.199	0.199	0.005
σ_ε	0.100	0.105	0.017	—	—	—		
<i>Economic Model Parameters</i>								
λ	0.438	0.410	0.046					
$\bar{\delta}$	0.953	0.952	0.001					
ψ	0.582	0.610	0.053					
θ	0.451	0.447	0.035					
<i>Measurement Errors</i>								
σ_y	0.165	0.163	0.006	0.147	0.005	0.147 ^a	0.146	0.005
σ_c	0.355	0.356	0.008	—		—		
σ_{c_0}	0.430	0.428	0.013	—		—		

^aIn estimations with income data alone, transitory shocks and measurement error cannot be identified separately. So we assume all i.i.d. shocks are measurement error with a standard deviation equal to $\sqrt{\sigma_y^2 + \sigma_\varepsilon^2}$ using the estimated value from the benchmark.

sample. We then take the average of the percentile rankings for this household over all the ages that it has a valid observation. Then for each missing income observation of this household, we impute the income level corresponding to its average percentile ranking given the income distribution in our sample for that age. We apply the same procedure to fill in missing consumption data. We construct growth rate variables differently: the past growth rate for age t in the auxiliary model is computed by taking the difference between the latest valid observation before t and the first valid observation for the individual in the data set and dividing this difference by the number of years between the two points. The future growth rate at a given age is constructed analogously. If either variable cannot be constructed for a given age, we use the average growth rate of that variable over the life cycle instead.

B.2 A Monte Carlo Study

To investigate the ability of the proposed estimation method to uncover the true structural parameter vector with the specified auxiliary model, we begin by conducting a Monte Carlo study. The results are contained in Table A.1.³¹ In column 1, the “true values” for the parameters are set to our benchmark estimates from

³¹We set $\bar{u}^{i,c} \equiv 0$ in the Monte Carlo analysis, because all households in the simulated data have the same

PSID income and consumption data (column 1 of Table I). For each parameter, the initial values are drawn randomly from a uniform distribution centered around the true value but with a wide support.³² The results discussed here are based on 100 replications, where each Monte Carlo run takes about 20 to 24 hours on a state-of-the-art workstation. Column 2 reports the results when both income and consumption data are used jointly for estimation. Clearly, the estimation method works well: bias is virtually absent for most parameters and is very small for the remaining few. Standard deviations are very small, indicating that all the parameters, with the exception of $corr_{\alpha\beta}$, can be identified in this framework. A useful question to ask is whether there are benefits to using consumption data in the estimation for (the six) parameters that can be identified with income data alone. To investigate this we use the same “true” values as in the previous exercise, but estimate the income process with income data alone and with equation (20) as the only auxiliary model regression (reported in columns 4 and 5). Perhaps, unsurprisingly, the mean estimates have little if any bias, and although the precision of the estimates falls, this is minor for all parameters, except σ_β . However, as we shall see in the next section, when we estimate the income process from “real” (PSID) data, they turn out to be different from those in column of this table (i.e., A.1). Thus, another exercise we conduct is to take as the “true” values as those obtained from the PSID with income data alone (column 2 of Table I). As seen in columns 7 and 8, the estimates are still largely unbiased, but the precision now has fallen significantly for some parameter values, most importantly for σ_β —going from 0.17 to 0.27—and for $corr_{\alpha\beta}$ —going from 0.15 to 0.29. This reduced precision makes it harder to separate whether the rise in income inequality is coming from σ_β or from σ_α through the strong correlation.

Although it is difficult, if not impossible, to prove identification in this general setup, overall these results suggest strongly that local identification near the true parameter vector does indeed hold. These results are encouraging and suggest strongly that the proposed methodology is a feasible and practical method for estimating structural consumption-saving models with widely missing data, binding borrowing constraints, and multiple sources of heterogeneity and randomness.

Finally, before settling down on the auxiliary model used in this paper (equations (19) and (20)), we explored a large number of alternatives. In Table A.2, we report a subset of our results from that work, which is representative of the issues we generally encountered. Columns 2 and 3 report the results when we use the same auxiliary model regressions as in the baseline case, but run them separately for *three* age groups (instead of two) yielding 75 parameters (instead of 50). In Columns 4 to 9, we go in the other direction and examine a sequence of auxiliary models that are successively more parsimonious. First (columns 4 and 5), we use the baseline auxiliary model but put no weight on the *WY* moment. Second (columns 6 and 7), we use the baseline auxiliary model but drop regressors that have *t*-statistics less than two. Finally, in columns 8 and 9, we use the same auxiliary model as in 6 and 7, but use only one set of equations for individuals of all age groups. The overall conclusion from these experiments is that the baseline auxiliary performs better in terms of both bias and precision of the estimates than the four alternatives that we explored. The main differences revolve around three parameters: λ , σ_β , and $corr_{\alpha\beta}$. For these parameters, we find that the alternative auxiliary models tend to generate estimates that exhibit both more bias and less precision. The differences are small in some cases, but quite large in the last experiment (with a single age group and dropping insignificant regressors). We have also experimented with other auxiliary models that seemed a priori plausible and differed more dramatically

demographics and zero initial wealth, making this fixed effect redundant.

³²The exact ranges for the starting values are as follows: $\sigma_\alpha \in [0.2, 0.4]$, $\sigma_\beta \in [1.0, 3.0]$, $corr_{\alpha\beta} \in [-0.60, 0.35]$, $\rho \in [0.55, 0.95]$, $\sigma_\eta \in [0.15, 0.25]$, $\sigma_\varepsilon \in [0.001, 0.20]$, $\lambda \in [0.001, 0.75]$, $\sigma_y \in [0.11, 0.9]$, $\sigma_c \in [0.32, 0.39]$, $\sigma_{c_0} \in [0.38, 0.48]$, $\delta \in [0.92, 0.99]$, and $\psi \in [0.3, 0.99]$.

Table A.2: MONTE CARLO ANALYSIS: ALTERNATIVE AUXILIARY MODELS

Auxiliary Model:	Baseline			No WY moment		Drop regressors with t -stats < 2.0		Drop regressors with t -stats < 2.0		
	Three			Two		Two		One		
Number of Age Groups: ^a	Estimates			Estimates		Estimates		Estimates		
Value 1	Mean	Sd. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.	Mean	Std. Err.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(8)	(9)
<i>Income Processes Parameters</i>										
σ_α	0.284	0.292	0.023	0.284	0.025	0.284	0.024	0.284	0.284	0.024
σ_β	1.852	2.000	0.163	1.814	0.193	1.821	0.207	1.758	1.758	0.397
$corr_{\alpha\beta}$	-0.162	-0.211	0.143	-0.162	0.189	-0.170	0.205	-0.09	-0.09	0.228
ρ	0.754	0.765	0.027	0.756	0.025	0.760	0.029	0.757	0.757	0.038
σ_η	0.196	0.201	0.005	0.196	0.005	0.194	0.005	0.195	0.195	0.005
σ_ε	0.004	0.041	0.031	0.026	0.025	0.036	0.023	0.039	0.039	0.022
λ	0.345	0.320	0.110	0.291	0.110	0.310	0.094	0.272	0.272	0.114
δ	0.950	0.950	0.002	0.949	0.003	0.951	0.002	0.951	0.951	0.002
ψ	0.874	0.949	0.041	0.886	0.082	0.882	0.083	0.797	0.797	0.148
σ_y	0.147	0.146	0.010	0.142	0.008	0.142	0.007	0.141	0.141	0.007
σ_c	0.356	0.371	0.002	0.356	0.003	0.356	0.002	0.356	0.356	0.002
σ_{c_0}	0.428	0.439	0.010	0.421	0.010	0.420	0.010	0.418	0.418	0.010

^aThe baseline estimation uses two age groups in the auxiliary model.

from those reported here and found them to perform very poorly. Therefore, we do not discuss them here.

Two main conclusions emerge from those analyses. One, the estimation method works very well: bias is virtually absent for most parameters and is very small for the remaining few. Standard deviations are very small, indicating that all the parameters, with the exception of $corr_{\alpha\beta}$, can be pinned down fairly precisely. Two, the income process parameters can be estimated using income data alone without any noticeable bias. However, under some plausible parameter combinations, the precision of the estimates of some key variables is significantly higher when estimated from income data alone (e.g., the standard error on σ_β goes from 0.17 to 0.27, and for $corr_{\alpha\beta}$ it goes from 0.15 to 0.29). This reduced precision makes it harder to separate whether the rise in income inequality is coming from σ_β or from σ_α through the strong correlation.

We conclude that although it is difficult, if not impossible, to prove identification in this general setup, this Monte Carlo analysis suggests strongly that local identification near the true parameter vector does indeed hold. These results are encouraging and suggest strongly that the proposed methodology is a feasible and practical method for estimating consumption-saving models with widely missing data, binding borrowing constraints, and multiple sources of heterogeneity and randomness.

C Data Appendix

C.1 CE Data

1972–73 Waves

We create a measure of nondurable consumption expenditures by adding the expenditures on food, alcohol, tobacco, fuel and utilities, telephone, other services, laundry, clothing, transportation, personal goods, recreation, reading, gifts, and other goods. The original size of the 1972–73 CE is 19,975 households. We keep households in our sample if they are headed by a married male who is between 30 and 65 years old and have nonzero food and income reports. In Table A.1 we report the number of households deleted from our sample during each sample selection requirement.

Table A.3: CE Sample Selection

<i>Selection criterion:</i>	Dropped	Remain
Initial sample	—	19975
Male head	4470	15505
Age restriction	5200	10305
Non-zero income and food	709	9596
Married	874	8722
Non-missing education	213	8509

1980–92 Waves

We merge the 1972–73 CE data with the 1980–92 data used in BPP. BPP use a similar sample selection as above. In addition, they exclude households with heads born before 1920 or after 1959. All nominal variables

are expressed in constant 1982–84 dollars. Income is deflated using the CPI. Total food expenditures are deflated using the average food price series provided by the Bureau of Labor Statistics. The inflation rates for food, fuel, alcohol, and transportation were determined by the corresponding price series provided by the BLS. We also drop households that have total real food consumption per adult equivalent less than \$300. Here, adult equivalent is defined as the square root of family size.

C.2 PSID Data

C.2.1 Sample Cleaning

Our measure of total food consumption comes from summing the responses to the questions about food consumed at home and food consumed away from home in each year (except for 1968, where the survey asked only about total food expenditures). This gives us a total food expenditure variable in each survey wave except for 1972, 1987, and 1988, when no food expenditure questions were asked.

In the PSID, the education variable is sometimes missing and sometimes inconsistent. To deal with this problem, we take the highest education level that an individual ever reports and use it as the education variable for each year. Since the minimum age needed to be included in our sample is 25, this procedure does not introduce much bias to our estimated education variable.

A well-known feature of the age variable recorded in the PSID survey is that it does not necessarily increase by 1 from one year to the next. For example, an individual can report being 30 years old in 1970, 30 in 1971, and 32 in 1972. This may be perfectly correct from the respondents' point of view, since the survey date may be before or after the respondent's birthday in any given year. We create a consistent age variable by taking the age reported in the first year that the individual appears as the head of a household and add 1 to this variable in each subsequent year.

The income variable we use is total after-tax non-financial household income. The way we construct this variable varies across years in the PSID because of different questions asked and different variable definitions. From 1968 to 1974 we take total family money, subtract taxable income of the head and wife (which includes both asset and labor income), and add back head and wife annual labor income. The family money variable is defined as total taxable income and transfers of the head, wife, and others in the household. From 1975 to 1983 we take the family money variable and subtract the asset income of the head and the asset income of the wife. From 1975 to 1977 the asset income of the head is defined as the sum of the asset part of business income, the asset part of farming, and the asset part of rental income. From 1978 to 1982 the definition of the asset income of the head is the same, except for the addition of the asset part of gardening. From 1983 to 1991 the definition remains the same except dividend income is also added. For 1992 the definition remains the same except interest income and income from family trusts are added. From 1975 to 1983 the wife's asset income is listed as one variable. From 1984 to 1991 we generate the wife's asset income as the sum of the wife's share of asset income and the wife's other asset income. For 1992 the wife's asset income is the sum of the wife's dividend income, interest income, family trust income, asset part of business income, and other asset income. From 1984 to 1992 to create the nonfinancial income variable we take family money and subtract head asset income, the wife asset income, and asset income of other members of the household.

C.2.2 Sample Selection

We start with a possible sample of 67,282 individuals interviewed between 1968 and 2005. To be in our final sample an individual must satisfy each of eight criteria in at least one year between 1968 and 1992. The number of individuals dropped at each stage in the sample selection is listed in Table A.4.

1. The individual must be from the original main PSID sample (not from the Survey of Economic Opportunities or Latino subsamples).

2. We require that the individual be married and that the individual has not changed partners from the previous year.

3. We require that individuals had no significant changes in family composition. This means that they responded either “no change” or “change in family members other than the head or wife” to the question about family composition changes.

4. The individual must not have missing variables for the head or wife labor income. The education variable for the head must also not be missing (this education variable refers to the one created after the sample cleaning mentioned previously).

5. The individual must not have food or income observations that are outliers. An income outlier is defined as a change in real income relative to the previous year of greater than 500% or less than -80% or total income less than \$1,000. A food expenditure outlier is defined as real total household food expenditure greater than income or food expenditure per effective adult less than \$300. Food expenditure per effective adult is defined as total household food expenditure divided by the square root of the number of members in the family.

6. We require that individuals have non-topcoded observations for the labor income of the head and wife and non-top-coded observations for total non-financial income.

7. The individual must be a male and the head of his household.

8. Household heads must be between 25 and 65 years old. (Only those between 25 and 55 are used in the main estimation in the paper.)

Adjusting for Taxes. From the non-financial income variable we need to subtract taxes paid on non-financial income. The PSID reports estimated total taxes for all households until 1991. For the years 1968–1990, we use the sum of the variables in the PSID that give the estimated federal tax liabilities of the head and wife and of others in the household. For 1975–1978 a variable is available that gives the amount of low income tax credit the household received. For these years the income tax credit is subtracted from the total amount of tax liability. We regress total tax liability on total labor income and its square and on total asset income and its square. We use these estimates to predict the total taxes paid on labor income. For the years 1991 and 1992, we use the NBER TAXSIM software to estimate the total taxes paid by each household on labor income. We assume that the husband and wife file a joint tax return and that the number of dependents claimed equals the number of children in the household. We also use the annual property tax liability variable as an input to the TAXSIM software to account for property taxes being deducted from federal taxable income. Since the public release version of the PSID does not contain state identifiers, we do not use the TAXSIM software to estimate state taxes paid. Finally, we subtract this estimated labor income tax from household income above to obtain the household after-tax labor income measure used in the estimation analysis.

Table A.4: PSID Sample Selection

Criteria	Dropped	Remain
Initial sample	—	67282
Main sample	39906	27376
Continuously married	2805	24571
No major composition change	4	24567
Missing data	1032	23535
Outlier or top-coded	71	23464
Male and head of household	19232	4232
Age restriction	429	3803
Five observations or more	1568	2235

Measure of Net Worth. Our wealth measure includes cash and demand deposits; time and saving deposits, CDs, and money market accounts; stocks, bonds, and mutual fund holdings (including IRAs); cash surrender value of life insurance policies; net equity in unincorporated businesses; and net equity in owner-occupied housing and other real estate. From the sum of these assets we subtract consumer debt (credit card debt, student and auto loans, etc.). The income measure is total household labor and asset income in that year.

C.3 Constructing a Panel of (Imputed) Consumption

The PSID has a long panel dimension but covers limited categories of consumption, whereas the CE survey has detailed expenditures over a short period of time (four quarters). As a result, most previous work has either used food expenditures as a measure of nondurable consumption (available in PSID) or resorted to using repeated cross sections from CE under additional assumptions.

In a recent paper, Blundell et al. (2006) (hereafter, BPP) develop a structural method that imputes consumption expenditures for PSID households using information from the CE survey. The basic approach involves estimating a demand system for food consumption as a function of nondurable expenditures, a wide set of demographic variables, and relative prices as well as the interaction of nondurable expenditures with all of these variables. To deal with the endogeneity of food and nonfood expenditures as well as measurement error in these variables, the estimation is carried out with an instrumental variables regression. The key condition for the imputation procedure to work is that all the variables in the demand system must be available in the CE data set, and all but nondurable expenditures must be available in the PSID. One then estimates this demand system using the CE data, and as long as the demand system is monotonic in nondurable expenditures, one can invert it to obtain a panel of imputed consumption in the PSID. BPP implement this method to obtain imputed consumption in the PSID for the period 1980 to 1992 and show that several statistics calculated using the imputed measure compare quite well with their counterparts from the CE data.

Our Imputation Procedure

In this paper, we modify and extend the method proposed by BPP to cover the period 1968 to 1992. Here we provide a brief overview of our method and a discussion of the quality of the imputation.

First, BPP include time dummies interacted with nondurable expenditures in the demand system to allow for the budget elasticity of food demand to change over time, which they find to be important for the accuracy of the imputation procedure. However, CE data are not available on a continuous basis before 1980, whereas we would like to use the entire length of the PSID (going back to 1968), making the use of time dummies impossible. To circumvent this problem, we replace the time dummies with the food and fuel inflation rate, which is motivated by the observation that the pattern of time dummies estimated by BPP after 1980 is quite similar to the behavior of these inflation variables during the same period.

A second important element in our imputation is the use of CE data before 1980. In particular, CE data are also available in 1972 and 1973, and in fact these cross sections contain a much larger number of households than the waves after 1980.³³ The data in this earlier period also appear to be of superior quality in certain respects compared with those from subsequent waves.³⁴ The use of these earlier data provides, in some sense, an anchor point for the procedure in the 1970s that improves the overall quality of imputation as discussed below. Finally, instead of controlling for life cycle changes in the demand structure using a polynomial in age (as done by BPP), we use a piecewise linear function of age with four segments, which provides more flexibility. This simple change improves the lifecycle profiles of mean consumption and the variance of consumption rather significantly. With these modifications, we obtain an imputed consumption measure that provides a good fit to the corresponding statistics in the CE data. Here, we summarize the most relevant statistics. Further details are contained in Appendix C.3.

We begin with two dimensions of consumption data that are crucial for our estimation exercise. First, the left panel of Figure A.1 plots the *average lifecycle profile* of log consumption implied by the CE data (marked with circles) as well as the counterpart generated by the imputed data (marked with squares).³⁵ To reduce the noise in the data, the figure also plots the corresponding “smoothed” series obtained by a Nadaraya-Watson kernel regression, with a Gaussian kernel. The two graphs overlap remarkably well, especially up to early age 50.³⁶ Second, the right panel plots the *within-cohort variance* of log consumption over the life cycle along with the smoothed series. Both in the CE and with the imputed PSID data, the variance rises between ages 25 and 65, although the total rise is rather small—about 5 log points. The finding of a small rise in within-cohort consumption inequality contrasts with earlier papers that have studied the CE data over the period from 1980 to 1990 (such as Deaton and Paxson (1994) and Storesletten et al. (2004b)) but is consistent with more recent papers that have used samples extending to the late 1990s (cf. Heathcote et al. (2010)). This finding will be important in understanding some of the estimates (especially, λ) that we obtain in the structural estimation below.

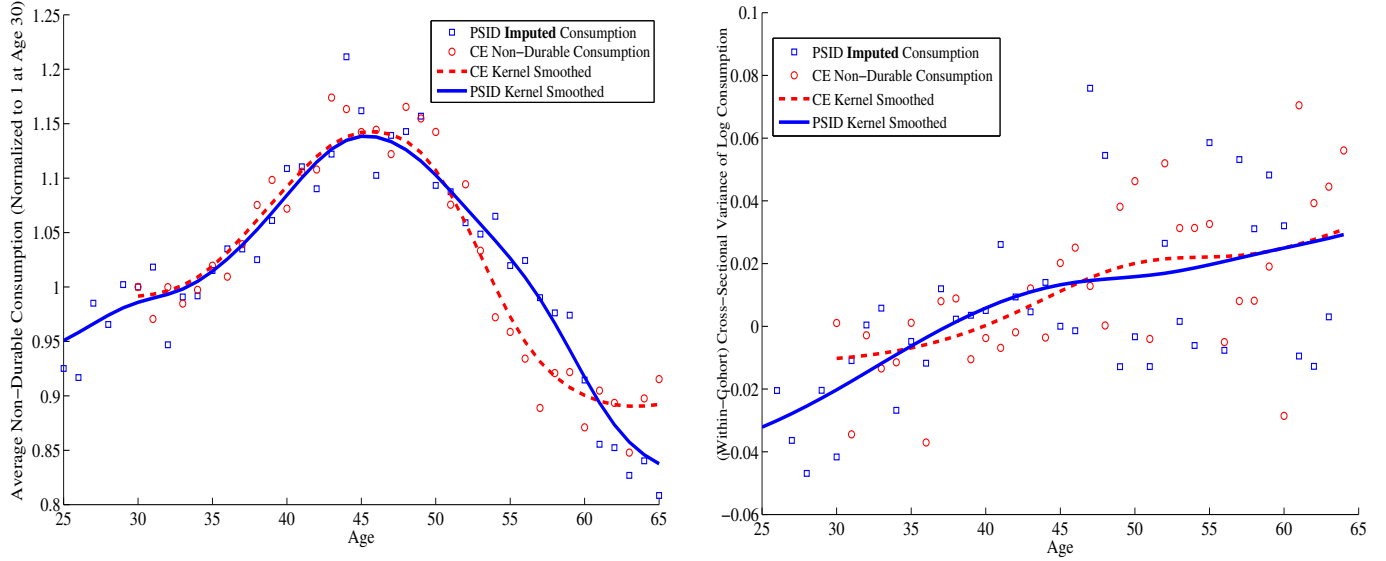
³³The sample size is around 9,500 units in 1972–73 surveys, but ranges from 4,000–6,000 units in the waves after 1980. There are also some differences in the survey design in the earlier CE—such as the non-rotating nature of the sample in the 1972 and 1973 panels—but these differences do not appear consequential for our purposes. See Johnston and Shipp (1997) for a more detailed comparison of different waves of the CE survey over time.

³⁴Slesnick (1992) shows that when one aggregates several subcomponents of consumption expenditures in the CE, they come significantly closer to their counterparts in the National Income and Product Accounts (NIPA) than the CE waves after 1980. For example, in 1973 the fraction of total expenditures measured by the CE is 90% of personal consumption expenditures as measured by NIPA, whereas this fraction is consistently below 80% after 1980 and drops to as low as 75% in 1987. Similarly, the fraction of consumer services in the CE accounts for 93% of the same category in NIPA in 1973 but drops to only 66% in 1989.

³⁵The lifecycle profiles are obtained by controlling for cohort effects as described in Guvenen (2009).

³⁶If we do not use the 1972–73 CE in the imputation procedure, the average profile of imputed consumption would rise by 51% between ages 25 and 45 instead of the 22% rise in the baseline imputation and would therefore vastly overestimate the corresponding rise in the CE data shown in Figure A.1.

Figure A.1: Mean and Variance Profile of Log Consumption over the Life Cycle



Another useful exercise is to test the *out-of-sample* predictive ability of the imputation procedure. To do this, we split the CE sample used in the imputation above into two randomly drawn subsamples (each containing exactly half of the observations in each survey year). We use the first subsample to estimate the food demand system as above, which we then use to impute the nondurable consumption of the second subsample (control group).³⁷ Figure C.3 plots the actual consumption of the control group against the imputed one for each household (for the simulation with the median regression slope). The imputed consumption data form a cloud that aligns very well with the 45-degree line. In fact, a linear regression of imputed consumption on the actual one yields an average slope coefficient of 0.996 and a constant term of 0.25. The average R^2 of the regression is 0.67, implying that the imputed consumption has a correlation of 0.81 with the actual consumption at household level.³⁸ The fact that the slope coefficient is almost equal to 1 is important: a slope above 1 (with a positive intercept) would indicate that the imputation systematically overstates the variance of true consumption, which would in turn overstate the response of consumption to income shocks, thereby resulting in an overestimation of the size of income shocks. The opposite problem would arise if the slope coefficient was below 1.

As a final, and rather strict, test to detect whether systematic patterns exist in the imputation error, we regressed it on household characteristics including dummies for each age group, education dummies, family size, region dummies, number of children dummies, and food and fuel prices. The median R^2 of this regression was 0.002 (and there was at most one variable that was significant at the 5% level in any given simulation), indicating no evidence of systematic imputation errors by demographic groups. Overall, we conclude that the imputation procedure works fairly well and does not result in any systematic over- or under-prediction of actual consumption.

³⁷To control for the randomness of each subsample, we repeat this exercise 200 times.

³⁸The results in the text refer to the average of these 200 replications. Across simulations, the slope coefficient in the regression ranges from 0.978 to 1.020, and the R^2 ranges from 0.644 to 0.691.

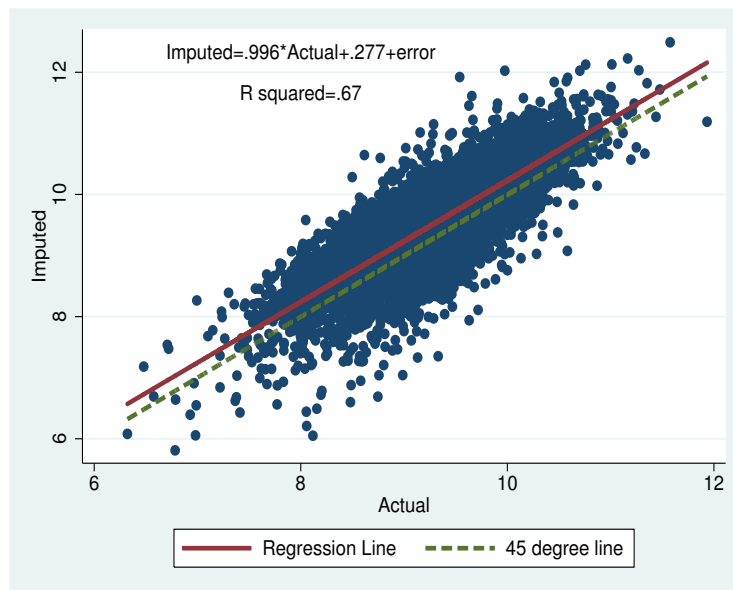


Figure A.2: **Out of Sample Predictive Power of the Imputation Method in the CEX.** This plot is obtained by estimating the IV food demand system on a randomly chosen half of the CEX sample and then imputing the consumption for the other half (control group). The figure plots the actual consumption of the control group versus their imputed consumption. The average regression slope is 0.996, the average constant is 0.24, and the average R^2 is 0.67 over 200 repetitions.

Further Details

This section describes the details of the imputation procedures and reports some further validation tests on the quality of imputation. Specifically, we estimate a food demand system as a function of nondurable expenditures, demographics, and relative prices using an instrumental variables approach. To deal with endogeneity and measurement error, we instrument log nondurable expenditures (as well as their interaction with demographics and prices) with the cohort-year-education specific average of the log of the husband’s hourly wage and the cohort-year-education specific average of the log of the wife’s hourly wage (as well as their interaction with the demographics and prices). The cohort-education-year specific averages of the log of the husband’s and wife’s hourly wage rates are generated as follows. The cohorts are divided into 5-year cells by year of birth, starting with 1910 and ending with 1955. The education cells are divided into high school dropouts, high school graduates, and more than high school education. For each year (1972, 1973, and 1980–1992) and each cohort-education cell we calculate the mean of the log of hourly wages of household heads and wives. The four age dummies used in the interaction terms are less than 37, between 37 and less than 47, between 47 and less than 56, and greater than or equal to 56. There are three inflation dummies: less than 5% inflation, between 5% and less than 11%, and greater than or equal to 11%. There are three children categories used in the interaction terms: one child, two children, and three or more children.

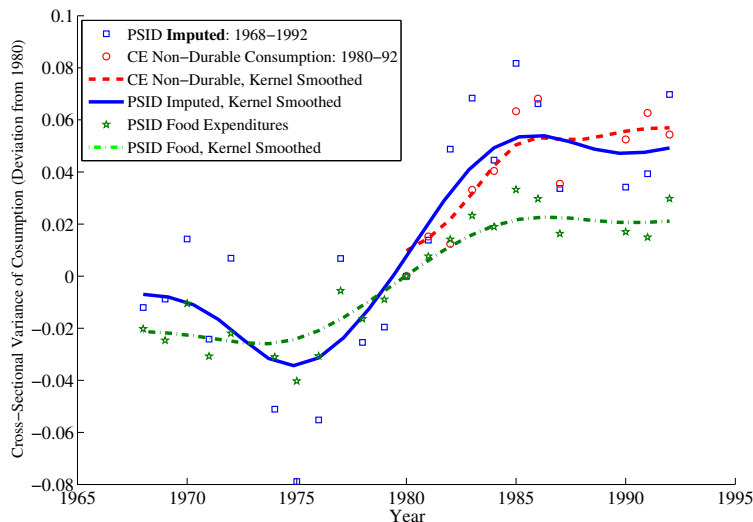
Table A.5 reports the results from the IV estimation of the demand system using the CE data. Several terms that include the log of nondurable expenditures are significant as well as several of the demographic and price variables. Most of the estimated coefficients have the expected sign. We invert this equation to obtain the imputed measure of household nondurable consumption expenditures.

Table A.5: Instrumental Variables Estimation of Demand for Food in the CE

Variable	Estimate	Variable	Estimate
$\ln(c)$	0.798*** (26.80)	$\ln(c) \times I\{11\% \leq \Delta \log p_{fuel}\}$	0.00386* (1.83)
$\ln(c) \times age \times I\{age < 37\}$	0.00036*** (3.38)	$\ln(c) \times (year - 1980)$	-0.00057 (-0.68)
$\ln(c) \times age \times I\{37 \leq age < 47\}$	0.00048*** (5.45)	One child	0.149 (1.16)
$\ln(c) \times age \times I\{47 \leq age < 56\}$	0.00042*** (5.75)	Two children	0.564*** (3.98)
$\ln(c) \times age \times I\{56 \leq age\}$	0.00037*** (6.08)	Three children+	1.203*** (8.23)
$\ln(c) \times$ High school dropout	-0.129*** (-7.57)	High school dropout	1.207*** (7.61)
$\ln(c) \times$ High school graduate	-0.043*** (-2.78)	High school graduate	0.417*** (2.90)
$\ln(c) \times$ One child	-0.014 (-1.01)	Northeast	0.0587*** (10.36)
$\ln(c) \times$ Two children	-0.055*** (-3.68)	Midwest	0.0293*** (5.23)
$\ln(c) \times$ Three children+	-0.123*** (-7.92)	South	-0.0031 (-0.63)
$\ln(c) \times I\{5\% \leq \Delta \log p_{food} < 8\%\}$	0.00096 (1.01)	Family size	0.0509*** (16.20)
$\ln(c) \times I\{8\% \leq \Delta \log p_{food} < 11\%\}$	0.00858*** (4.25)	$\ln p_{food}$	0.581** (2.28)
$\ln(c) \times I\{11\% \leq \Delta \log p_{food}\}$	-0.00091 (-0.39)	$\ln p_{fuel}$	-0.117 (-0.97)
$\ln(c) \times I\{5\% \leq \Delta \log p_{fuel} < 8\%\}$	0.00074 (0.66)	White	0.0824*** (11.38)
$\ln(c) \times I\{8\% \leq \Delta \log p_{fuel} < 11\%\}$	0.00091 (0.53)	Constant	-1.822*** (-2.65)
		Observations	21864

We pool the data from the 1972-73 waves of the CE with the 1980-92 waves. We instrument log food expenditures (and their interactions) with the cohort-education-year specific average of the log husband's and wife's hourly wage rates (and their interactions with age, education, and inflation dummies and a time trend). The t -statistics are contained in parentheses. The lowest value of Shea's partial R^2 for instrument relevance is 0.086, and the p-value of the F-test on the excluded instruments is smaller than 0.001 for all instruments.

Figure A.3: Cross-sectional Variance of Log Consumption in CEX and Imputed PSID Data: 1968–1992



BPP used the evolution of the variance of consumption over time to check the quality of their imputation procedure. For completeness, here we discuss the results of our imputation for the same statistic. Figure C.3 plots the cross-sectional variance of log consumption over time. In the figure, the (red) circles mark the CE data, whereas the (blue) squares show the imputed consumption in the PSID. Similarly, the dashed (red) line and the solid (blue) line show the kernel-smoothed version. The imputed consumption series tracks the CE data fairly well, showing an overall rise in consumption inequality of 6–7 log points between 1980 and 1986, followed by a drop from 1986 to 1987 and not much change after that date. The dashed-dotted line shows that if one simply were to use food expenditures in the PSID instead, the overall pattern would remain largely intact, but the movements would be *quantitatively* muted compared with the data: the rise in consumption inequality would be understated by more than half by 1986 and by as much as two-thirds by 1991.

D Robustness Analysis

We now present results from robustness exercises. These experiments have been conducted with a version of the model in which the probability of death is set to zero until age $T = 80$, and households have access to self-insurance only. These two changes do not make an appreciable difference in the results so we present the robustness analysis using this slightly simpler version of our model.

D.1 Lower Interest Rate

In our benchmark, we interpreted the risk-free asset as corresponding to a broad set of assets available to households, which motivated our relatively high choice of $r = 5.26\%$. Another perspective is that such an asset can be thought of as a government bond, so that a lower return may be more appropriate. To explore the sensitivity of our results, we reestimate the model, setting $r = 3.1\%$ (i.e., $\gamma = 0.97$). As seen in column 1 of Table A.6, this change has virtually *no* effect on the estimates of the income process as well as measurement

Table A.6: SENSITIVITY ANALYSIS: ALTERNATIVE ASSUMPTIONS

	Low interest rate $\gamma = 0.97$ (3)	Alternative filling-in method (4)	Doubling minimum income $\underline{Y} = 0.10$ (5)	Use data up to age 65 (6)	No prior uncertainty $\lambda = 0$ (7)	Maximum prior uncertainty $\lambda = \lambda^{\max}$ (8)
<i>Income Processes Parameters (can be identified with income data alone)</i>						
σ_α	0.284	0.220	0.293	0.228	0.268	0.248
σ_β	1.856	1.916	1.886	1.088	1.756	1.04
$corr_{\alpha\beta}$	-0.164	0.003	-0.166	-0.161	-0.086	0.751
ρ	0.755	0.760	0.759	0.801	0.777	0.806
σ_η	0.196	0.200	0.208	0.200	0.195	0.196
σ_ε	0.005	0.006	0.007	0.003	0.006	0.010
<i>Economic Model Parameters (need consumption data)</i>						
λ	0.380	0.327	0.374	0.520	0.0 (fixed)	0.656 ^a
δ	0.964	0.950	0.951	0.943	0.954	0.951
ψ	0.790	0.921	0.757	0.992	0.761	0.895
<i>Measurement Error and Transitory Shocks (need consumption data)</i>						
σ_y	0.147	0.145	0.156	0.152	0.148	0.151
σ_c	0.356	0.356	0.356	0.356	0.355	0.356
σ_{c_0}	0.429	0.414	0.427	0.432	0.424	0.433
max % constrained..	16.1%	10.2%	12.5%	9.1%	14.1%	13.2%
.. at age	30	31	30	35	30	33
\underline{q}_0 / mean income	0.35	0.44	0.41	0.93	0.21	0.37
\underline{q}_{55} / mean income	0.60	0.59	0.87	0.73	0.44	0.55

^aThis value of λ represents the highest value feasible (i.e., highest prior uncertainty) given that households know their α^i , which is correlated with β^i and thus contains information about it.

error parameters. As for the economic model parameters, λ increases—slightly—from 0.345 to 0.38, whereas δ increases significantly (from 0.95 to 0.964) and borrowing constraints become tighter— ψ falls significantly, from 0.87 to 0.79—both presumably to make the model match the wealth-to-income ratio moment better. Overall, the most important parameters about the income process as well as λ seem very robust to reasonable changes in the interest rate.

D.2 Alternative Filling-in Method

We now examine the robustness of the estimation results to the method chosen for filling in missing data (column 2 of Table A.6). This could be potentially important because more than half of the values in our sample are missing—and therefore filled in—compared to a fully balanced panel. As an alternative procedure, we consider a much simpler filling-in method: for each individual we calculate the lifetime average of either log consumption or log income using available observations. If a consumption or income observation is missing in a given year, we simply replace the missing data with this average. We then use this filled-in data to construct all the missing right-hand-side variables in the regressions. As seen here, with the exception of σ_α , the estimates are largely unchanged from the benchmark case.

D.3 Higher Minimum Income

We now investigate the sensitivity to the choice of minimum income \underline{Y} , by doubling its magnitude to 10% of median income (column 3). The effects on the estimates are very mild, with the only noteworthy changes being a rise in λ , from 0.345 to 0.375, and a fall in ψ from 0.874 to 0.756. However, because \underline{Y} has been doubled, the borrowing constraint is actually looser than before: \underline{a}_0 rises to 41% of average income, compared with 33% in the benchmark case.

D.4 Using All Available Data Up to Age 65

In the estimation so far, we have restricted our sample to ages 55 or younger, for several reasons.³⁹ Still, it is useful to examine how the results would change if the entire sample (up to age 65) were used in estimation. As seen in column 4 of Table A.6, some parameters change very little, whereas other important ones do change. For example, σ_β falls to 1.08 and λ simultaneously rises to 0.51, implying that the amount of prior uncertainty about growth rates (as measured by the prior standard deviation) falls slightly from $1.85 \times 0.345 = 0.638$ to $1.08 \times 0.51 = 0.55$. Now both the variance of log income and consumption graphs fit much better to the data. It seems that the deviations of the variance of log income and consumption figures from their data counterpart were “tolerable” as viewed through the auxiliary model, when data up to age 55 were used in estimation. But the inequality profiles implied by $\sigma_\beta = 1.85\%$ deviate farther from the data after age 55 at an increasing rate. This leads the estimator to reduce σ_β as well as $\lambda\sigma_\beta$, which then results in a better fit for both graphs. It is

³⁹One is that our assumption of linearity for the individual-specific trend is more likely to be accurate for households before this age, as widening income inequality slows down near retirement. Second, labor hours inequality increases near the retirement age, which weakens the link between wage and income inequality. Given that we are abstracting away from labor supply choice here, it seems more appropriate to restrict attention to the earlier period. Finally, the number of individuals in our sample goes down quickly at older ages, increasing the noise and reducing the usefulness of data from this group.

interesting to see that even though these moments have not been used in the estimation explicitly, matching the auxiliary model coefficients somehow ensures that the estimated model does a reasonable job of matching these economically important figures.

D.5 Fixing Borrowing Constraints

We explore the effects of fixing the tightness of the borrowing constraint at some values that have been used in the literature on the estimated parameters. Column 7 of Table A.7 displays the results when the borrowing constraint is chosen to be the natural borrowing limit, which is obtained by setting $\psi = 1$. Similarly, column 8 reports the results of the opposite exercise—of disallowing borrowing—obtained by setting $\underline{a}_t \equiv 0$. The value of risk aversion is set to 2 as in the benchmark case in Table I. Although there is some variation across the two columns and there are some differences from the benchmark case, by and large, these differences are quite minor. The main difference is in the fraction of households that are constrained, which is 38% when no borrowing is allowed, 14% in the benchmark estimation, and 7% with the natural borrowing limit. Thus, there is clearly sufficient information in the auxiliary model to allow us to pin down the value of the borrowing constraint, but its particular value does not seem to affect the other estimates substantially.

D.6 On Risk Aversion, Time Preference Rate, and Borrowing Limit

Although it would be valuable per se if we could separately identify δ and ϕ , this is not the central aim of this paper. But the assumptions we make regarding these parameters (e.g., whether they are fixed, estimated, etc.) still matter, because they could affect the inference regarding income risk—potentially seriously—in turn jeopardizing the main goal of our investigation (as shown in Section 5.1). Therefore, we begin by conducting a series of sensitivity experiments to understand the effects of preference parameters on the overall estimation results. To this end, we first reestimate the benchmark specification, but now fix the risk aversion at, respectively, 1 (column 1 of Table A.7) and 3 (column 2). A quick glance across these two columns reveals two findings. First, δ and ϕ move strongly in opposite directions: δ goes up to 0.9526 when ϕ is reduced to 1, and goes down to 0.9416 when ϕ is increased to 3. Second, and fortunately, the remaining—11—parameters are *virtually unchanged* from the benchmark case, a quite striking finding. We now discuss these two sets of results in turn.

D.6.1 Are δ and ϕ Separately Identified?

The strength of the opposite movement in δ and ϕ (as we vary the latter) is remarkable. In fact, the correlation between the two estimates is worth reporting: -0.97 ! However, a correlation with three data points (one for each of the three values for risk aversion) is, obviously, not as informative as one would like, so we conducted a simple Monte Carlo study where we fixed all the parameter values except δ and ϕ , which are estimated. Across 200 repetitions, the correlation between the estimates of the two parameters was -0.88 . The results suggest that only a particular combination of the two parameters is identified, but that there is insufficient information to disentangle the two. In unreported results, we have tried adding additional regressors into equations (19) and (20), as well as adding new regressions suggested by theory, such as the second order moments of consumption growth or levels (computed in various ways) to capture precautionary savings demand, which could be informative about ϕ . In every case, we found the same strongly negative correlation.

Table A.7: SENSITIVITY OF STRUCTURAL ESTIMATES TO RESTRICTIONS ON ECONOMIC MODEL PARAMETERS

Preset Parameters↓	Role of Preference Parameters								Borrowing Limit
	Low RRA (1)	High RRA (2)	δ fixed (3)	δ fixed (4)	δ estim. (5)	δ fixed (6)	$\psi = 1$ (7)	$\underline{a}_t \equiv 0$ (8)	
ϕ (risk aversion)	1	3	1	1	1	1	2	2	
δ (time discount factor)	estim.	estim	0.94	0.94	estim.	0.953	estim.	estim.	
Weight on WY moment	10.0	10.0	1.0	0.0	0.0	0.0	10.0	10.0	
<i>Income Processes Parameters (can be identified with income data alone):</i>									
σ_α	0.283	0.282	0.339	0.326	0.332	0.281	0.333	0.272	
σ_β	1.838	1.841	3.997	2.165	2.093	1.850	2.129	1.747	
$corr_{\alpha\beta}$	-0.161	-0.161	0.660	-0.243	-0.162	-0.161	-0.139	-0.101	
ρ	0.756	0.756	0.821	0.724	0.738	0.760	0.750	0.765	
σ_η	0.196	0.196	0.238	0.194	0.196	0.195	0.195	0.194	
σ_ε	0.005	0.005	0.007	0.001	0.0124	0.005	0.029	0.005	
<i>Economic Model Parameters (need consumption data)</i>									
λ	0.368	0.330	0.998	0.001	0.035	0.343	0.360	0.283	
δ	0.953	0.942	0.94*	0.94*	0.938	0.953*	0.951	0.949	
ψ	0.877	0.871	0.002	0.997	0.998	0.923	0.99*	$\underline{a}_t \equiv 0$	
Max % constrained	13.9%	10.5%	52.2%	21.7%	14.3%	12.2%	7.2%	38.1%	
Max constr. age	31	33	25	35	35	33	35	25	
Wealth-to-income ratio	1.08	1.08	0.94	-0.03	0.15	1.03	1.08	1.07	
\underline{a}_0 / mean income	0.33	0.32	0.07	0.90	0.90	0.44	0.83	0.0	
\underline{a}_{55} / mean income	0.53	0.53	0.31	0.73	0.73	0.59	0.71	0.0	

Notes: The values of measurement errors and transitory shocks are not reported because they vary by little across specifications. The range for σ_y is from 0.129 to 0.149, for σ_c it is 0.353 to 0.356 and for σ_{c_0} it is from 0.42 to 0.429, with the exception of column 3 for which it is 0.514. *Fixed a priori at the value indicated.

Given how strongly this result manifests itself in our framework, we turn to another paper, Gourinchas and Parker (2002), which estimates the same two parameters also jointly using income and consumption data. There are several important differences between our paper and theirs, leaving open the possibility that there could be different sources of identification in that paper that perhaps could overcome the difficulty we face. These authors report 18 different estimates of δ and ϕ across four classes of experiments: (i) their baseline estimation using robust and optimal weighting matrices (two results), (ii) estimates for five education groups, (iii) estimates for four occupational groups, and finally (iv) estimation results for seven different robustness exercises. In most cases, the standard errors are small, indicating that both parameters are estimated quite precisely. (Furthermore, notice that there is no particular reason for the estimates of δ and ϕ to be correlated across these 18 pairs if each parameter is precisely identified.) Nevertheless, the correlation between the 18 pairs of estimates turns out to be -0.989 . Moreover, this result is not driven by a few outliers. For example, the five estimates for different educational groups have a correlation of -0.978 , whereas the correlation is -0.975 for the different occupation groups and -0.999 across the robustness results. (Throwing out two pairs from the last case that have very low estimates of δ yields the lowest correlation we obtain: -0.953 .)

Based on these findings, we conjecture that the lack of identification between δ and ϕ may be endemic to the estimation of these parameters from consumption-savings models with fixed interest rates. We view these (admittedly negative) results as providing a challenge for future work to find ways to pin these two key parameters down precisely.

D.6.2 Does It Matter for the Estimates of Income Risk That ϕ Is Not Separately Identified?

As noted above, the core issue for this paper is the estimation of income risk. On this front, the news is more encouraging: all the parameters relating to income risk are robust to variations in risk aversion, which is reassuring. One question these results bring up is the following: Can we simply fix a reasonable combination of (ϕ, δ) (say, based on values commonly used in the literature) and estimate only the income process parameters, or is it important to estimate at least one of these two parameters as we have done so far?

To answer this question, we fix $\phi = 1$ and $\delta = 0.94$, and estimate the remaining parameters (reported in column 3 of Table A.7). These estimates are dramatically different from the benchmark values and appear very implausible. For example, σ_β is now 3.997 (which is essentially at the upper bound we imposed for computational reasons) and $corr_{\alpha\beta}$ is 0.66, implying an enormous rise in the variance of log income over the life cycle that is many times what is observed in the data; λ is now 0.687, which is at almost its highest theoretical value and implies that households perceive 2/3 of this *overestimated* rise in income inequality as risk/uncertainty. Moreover, 52% of households now appear borrowing constrained. By any measure, these estimates are quite extreme. To make things worse, we should note that here we are reporting the results of the estimation when the weight on the *WY* moment is reduced from 10.0 in the benchmark case to 1.0. If it were not for this change, the estimates would be even more extreme, with the remaining parameters also getting stuck at their bounds ($\rho = 0.9999$, etc.) Overall, this experiment illustrates how quickly things can go wrong if proper care is not applied. Thus, the specific values of (ϕ, δ) do not seem to matter for the other estimates *only* if we estimate at least one of those parameters.

One reaction could be that perhaps these extreme outcomes are the results of imposing the *WY* moment: because δ is fixed, the model cannot adjust this parameter to easily match the value of *WY* observed in the data and instead substantially increases both the amount of income risk (as roughly measured by $\lambda\sigma_\beta$)

and the tightness of the borrowing constraint ($\psi = 0.02$). Even with these dramatic adjustments the model still undershoots the wealth-to-income ratio: 0.94 in the estimated model vs. 1.08 in the data. Thus, as important as this moment may be for a proper calibration, perhaps we can obtain more plausible estimates if we reestimate the model by dropping that moment. This exercise is carried out in column 4. Indeed, some of the parameter estimates look more reasonable now: σ_β is 2.167% and the correlation is -0.24 , which is not substantially different from the benchmark. Other estimates of the income process are also plausible and close to their benchmark values. Unfortunately, though, the parameters of the economic model now look quite suspect: $\lambda = 0.001$ and $\psi = 0.997$, both having moved from one bound to the other. Furthermore, the value of WY , which has not been imposed as a moment, is now -0.03 as compared to 1.08 in the data! A final thought is that perhaps in addition to eliminating the WY moment, we should also not fix δ and instead estimate it. The results are displayed in column 5, and the results are barely changed from the previous column.

These negative results lead to another important question. It seems that the WY moment is very important for properly estimating some parameters (such as λ and ψ , among others). But previously we spent significant efforts discussing how the auxiliary model regressions were important for pinning down these parameters. Is it possible that this emphasis was misplaced and the identification of several important parameters is coming mainly from the WY moment? It turns out the answer is no. What is happening instead is that a proper value of δ is essential for the estimation exercise and the WY moment simply ensures that δ is pinned down at a reasonable value given the other parameters of the model.⁴⁰ This can be seen as follows. In column 6, we reestimate the same model as in column 5—that is, without the WY moment, *but* we fix δ at its estimated value (0.953) when ϕ was set to 1 (in column 1), when the WY moment was used and δ was estimated. Notice that now we will not use the WY moment but only rely on the auxiliary model regressions. The estimates in column 6 of Table A.7 are very similar to those in column 1, and all appear very reasonable. This confirms our conjecture that the WY moment’s main role is to pin down the appropriate value of δ , and once that is achieved all other parameters are pinned down by the auxiliary model. (The small qualification to this statement is that ψ is 0.927 in column 5 instead of 0.88 in the benchmark case, which suggests that the WY moment perhaps also contains some information about the borrowing constraint. This would not be surprising.)

To summarize, we find that: (i) fixing either δ or ϕ and estimating the other is perfectly fine for properly estimating all the remaining parameters of the structural model, but (ii) fixing both δ and ϕ simultaneously creates severe biases. The main role of the WY moment appears to be to pin down a plausible value of δ that is consistent with the ϕ chosen, but has otherwise very little impact on the estimates of remaining parameters.

⁴⁰Notice that the role of δ as determining the wealth-to-income ratio is slightly different in our model compared to a standard calibration exercise. This is because here the amount of risk is not fixed (as would be the case in a calibration exercise where the income process is calibrated first and then δ is chosen). Instead the amount of risk and patience is jointly estimated.