Inequality and the Lifecycle[1]

Greg Kaplan
New York University
gregkaplan@nyu.edu

February 2007

Abstract
This paper investigates the sources of cross-sectional differences in consumption, labor supply, wealth and welfare over the lifecycle. I document the existence of rich and informative lifecycle patterns in the joint distribution of wages, hours, consumption and wealth. I then estimate a structural model of precautionary savings with endogenous labor supply and uninsurable wage risk in an attempt to assess the ability of the standard incomplete markets model to simultaneously account for the various dimensions of lifecycle inequality. I find that in many dimensions the model provides a coherent explanation. However, the combination of certain features of the data provides an inherent challenge for this class of models. Structural estimates of parameter values are obtained using Monte-Carlo Markov Chain techniques. These are then used to decompose inequality at different points in the lifecycle into differences in preferences, differences in initial wealth endowments, differences in fixed labor productivity and the accumulated effects of shocks realized after entry to the labor market. I find that around 40% of the cross-sectional differences in lifetime welfare are due to fixed skills and around 60% are due to lifecycle productivity shocks. Differences in financial wealth endowments, however, account for almost none of the inequality in lifetime welfare.

JEL Classification Codes: D31, E21,C13, J22

Keywords: Inequality, Lifecycle, Precautionary savings, Structural estimation, Monte-carlo Markov Chain, Labor supply.

[1]This is a revised version of an earlier draft circulated under the title "Idiosyncratic Income Risk and the Patterns of Lifecycle Inequality". Gianluca Violante has provided me with invaluable guidance at each stage of this project, for which I am extremely grateful. I have also benefitted from useful discussions with Richard Blundell, Mark Dean, Matthias Kredler, Costas Meghir and Tom Sargent. Of course, all errors are my own.

1 Introduction

A striking feature of cross-sectional and longitudinal data in the USA is that individuals who, at the age of entry to the labor force, seem similar along a number of observable dimensions, may ultimately have very different experiences over the course of their working lives. Different individuals will consume different quantities of goods and services, will spend different amounts of time working, will choose to concentrate their work effort at different points in their careers and will accumulate different levels of wealth. As a cohort of observably similar individuals ages, they will generate a joint distribution of consumption, labor supply and wealth at each age. Many of the elements of this distribution are observable in commonly used data sets.

In this paper I investigate the patterns of lifecycle inequality, as measured by the evolution of the cross-sectional distribution of consumption, labor supply, wages and wealth as individuals age. There are two broad goals of this research. The first is to quantify the extent to which the standard incomplete markets model with endogenous labor supply can account for these patterns.[2] This model has become a benchmark tool for analyzing policy and welfare in economies with heterogeneous agents and has increasingly been extended to allow for a labor-leisure decision.[3]

Other researchers have studied lifecycle inequality in the presence of incomplete markets, one element at a time.[4] This paper differs from most of the existing literature in that its focus is on evaluating the model on the basis of *all* the relevant dimensions of inequality for which it provides predictions. It is the restrictions that incomplete markets models place on the *joint* behavior of these variables that are most striking in their predictions and most pertinent in identifying parameters. Moreover, *both* consumption and work effort are relevant for individual welfare and almost all redistributive policies designed to impact inequality will induce labor supply responses.

Studying the full distribution of consumption, hours, wages and wealth together, however, comes at an added cost - it necessitates moving from a simple calibration exercise to a structural estimation. This paper is the first to obtain structural parameter estimates in a model with precautionary savings using data on cross-sectional inequality. The estimation strategy that I propose, based on Chernozhukov and Hong (2003), is both computationally and theoretically attractive - it employs Monte-Carlo Markov Chain techniques in a pseudo-Bayesian setting to overcome the curse of dimensionality and many of the common problems associated with

[2] By standard incomplete markets models I refer to the class of models in the spirit of Bewley (1986), Huggett (1993) and Aiyagari (1994).

[3] See for example Marcet, Obiols-Homs and Weil (2002), Heathcote, Storesletten and Violante (2004) and Low (2005). Floden (2006) provides a careful analysis of the effects of endogenous labor supply on precautionary savings.

[4] For example Storesletten, Telmer and Yaron (2004) and Guvenen (2006) have studied consumption inequality; Huggett (1996) has studied wealth inequality; and Low, Meghir and Pistaferri (2006) and Storesletten, Telmer and Yaron (2001) have studied differences in labor supply patterns. Heathcote, Storesletten and Violante (2006) also study lifecycle inequality in consumption and labor supply simultaneously, although their model is one of partial insurance, rather than a bond economy.

simulation-based estimation.

The second goal of the paper involves moving from testing the model's predictions for inequality to using the model to learn about the sources of inequality. The model allows for various sources of *ex-post* differences in outcomes. Understanding which of the competing sources of differences in consumption, hours, wealth and welfare are quantitatively most important is crucial for the design of policies that are intended to impact inequality. Accordingly, I decompose the cross-sectional distributions generated by the model at each point in the life-cycle into the fraction due to differences in preferences, differences in initial conditions (financial wealth and fixed skills) and the cumulative effect of persistent and transitory shocks. I then decompose total cross-sectional variation in welfare at each age into its various sources.

My findings are as follows. The mechanisms at work in the model - precautionary savings and labor supply flexibility in the face of persistent uninsurable wage risk - are consistent with a number of features of the data. The model is able to generate a smaller rise in consumption inequality than wage inequality over the lifecycle, a decreasing covariance between wages and hours and an increasing covariance between consumption and wages. However the presence of endogenous labor supply places important overidentifying restrictions on the joint distribution of hours, wages and consumption, which together provide a challenge for this class of models and reveal a number of puzzles. In particular the shape of the joint moments - how wages correlate with consumption and hours - are at odds with the lifecycle distribution of hours. I examine where and why the data fails to satisfy these restrictions, in the hope of gaining a better understanding of where and how this class of models should be enriched for future work.

The decompositions indicate that around half ($49\% - 59\%$) of the differences in total lifetime utility are due to shocks realized after entry into the labor market and the remainder is due to initial conditions.[5] However amongst initial conditions, almost all is due to human wealth, in the form of fixed skills, rather than financial wealth. An important finding regarding the sources of cross-sectional inequality is that different pictures emerge depending on where in the life-cycle we look. With regards to consumption inequality, I find that the fraction of dispersion accounted for by wage shocks increases from 25% to more than 70% over a career. Dispersion in hours in the model follows a similar pattern - the fraction accounted for by wage shocks increases from around 10% at the time of entry to the labor market, to nearly 70% close to retirement. The main finding with respect to the wealth distribution is that the impact of the wealth distribution of the young generation dies off, but does so fairly slowly. After 20 years of work, initial wealth endowments still account for nearly 40% of cross-sectional wealth inequality.

The remainder of the paper proceeds as follows. In section 2, I document some evidence on the patterns of lifecycle inequality. I then outline the key features of the model in section 3 and discuss the estimation strategy in section 4. I discuss the fit of the benchmark model in section 5 and decompose inequality into its various components in section 6. Section 7 contains some robustness exercises and section 8 concludes.

[5] Differences are due to different choices of weighting matrix in the estimation. See section 6 for details.

2 Empirical Evidence

In this section I document the patterns of lifecycle inequality.[6] Two features of the empirical analysis are worthy of mention. First, the focus is on residual, rather than raw inequality. This refers to cross-sectional differences in wages, hours, consumption and wealth that are not due to cross-sectional differences in educational achievements, race or cohort or time effects. Second, I focus on potential labor market experience, rather than age, as the variable that defines the lifecycle.[7] The terms age and experience are used interchangeably throughout the remainder of the paper.

2.1 Data and Sample Selection

I use data from the 1968 to 1997 waves of the Panel Study of Income Dynamics (PSID) and the 1980 to 1997 waves of the Consumer Expenditure Survey (CEX).[8] The key variables included in the analysis are wages, hours, consumption and wealth. The PSID is the source for data on wages, hours and wealth while the CEX is used for data on consumption and any joint moments that involve consumption. Hours are defined as total annual hours worked. Wages are constructed as total annual labor market earnings, including bonuses, tips and overtime, divided by annual hours. I use total non-durable expenditures as the measure of consumption.[9]

The sample is constructed to be as consistent as possible with the standard incomplete markets model. The model abstracts from a number of potentially important features of household decisions: female labor supply, intra-household decisions, changes in household composition, purchases of durable consumption goods and the extensive labor supply margin. I thus choose the household as the relevant unit of analysis and restrict attention to households with a prime-age male head who has a strong attachment to the labor force. I further restrict the sample to households where there is a single primary earner. Household consumption is equivalized to take account of differences in household needs across the population. I use the census equiva-

[6] Other authors have independently examined some of the facts documented here. Deaton and Paxson (1994) and Slesnick and Ulker (2004) document the fact that cross-sectional dispersion in consumption and earnings increases as a cohort of households age. Heathcote, Storesletten and Violante (2005) study labor supply inequality as well as the joint moments of wages, consumption and hours over the life-cycle. There is some disagreement in the existing literature over the magnitude of changes over the lifecycle and the effect of sample selection, variable definition and empirical methodology on these results. Most of these differences stem from the decision to control for cohort or time effects. See Heathcote, Storesletten and Violante (2005) for a discussion.

[7] Both of these choices are motivated by the structure of the model. Residual inequality is deemed to be the relevant measure of cross-sectional dispersion because the model is stationary and does not allow for an education decision before entry to the labor market. Moreover, in the model, agents are born at the time of entry to the labor market. Hence there is no distinction between two agents of different ages and different education levels with the same number of years of potential experience. Given that it is the cumulation of idiosyncratic productivity shocks that generate patterns of inequality in the model, the relevant lifecycle dimension on which to assess the model is experience rather than age.

[8] The raw CEX data from which I draw my sample comes from Kruger and Perri (2006). Data is available in the CEX until 2003. I truncate the sample at 1997 for consistency with the PSID, for which annual data is only available until 1997.

[9] The CEX is a quarterly panel in which households are followed for four quarters. I construct annual consumption expenditure by aggregating over four quarters in the same manner as in Kruger and Perri (2006).

lence scale, which is based on the number, age and gender of the household members. A full description of the selection criteria and variable definitions can be found in Appendix A.

I also assess the model on the basis of its predictions for wealth inequality over the lifecycle. The most reliable source for data on wealth inequality is the Survey of Consumer Finances (SCF). However in order to avoid the use of a third data source and to maintain a sample that is consistent with the one used for data on consumption, hours and wages, I use wealth data from a supplemental questionnaire in the PSID that was administered to a subsample of the panel in 1984, 1989 and 1994. Moreover the use of PSID allows for the inclusion of evidence on the covariance of wealth with wages and hours over the lifecycle, for which the model provides strong predictions. I define wealth as total net worth of a household. [10]

2.2 The Patterns of Lifecycle Inequality

The second moments of the joint distribution of wages, hours, consumption and wealth are estimated in a two-stage procedure, outlined in Appendix B. Confidence intervals are calculated by block bootstrap to account for serial dependence, heteroscedasticity and additional estimation error induced by the two-stage methodology.

I do not adjust the data for measurement error. Existing studies that do attempt to control for, or estimate, measurement error usually assume that measurement error is classical and constant over age. Note that constant classical measurement error in earnings, hours and consumption would affect only the level of measured variances and correlations, and not their shape over the life-cycle. In this paper, I am concerned with changes in, rather than levels of inequality, and as such I use only the shapes of the documented profiles in the structural estimation phase. Hence the impact of measurement error on the estimation results and decompositions is minimal.[11]

Figures 1 and 2 document the evolution of the cross-sectional distribution of wages, consumption and hours over the lifecycle, together with 95% confidence intervals and an HP-filtered smooth trend.[12] The top left panel in Figure 1 shows that cross-sectional inequality in wages for this sample of employed males increases substantially over the lifecycle. The total increase in the variance of log wages over the working years is around 0.1, with the bulk of this increase taking place in the first 15 years in the labor market. In section 4.2, I model this increase as re-

[10] Although the model features only a single risk-free asset with which households can save, the use of total wealth, which includes elements such as housing and other non-liquid assets, is justified on two grounds. First, it is likely that almost all assets held by households can be liquidated within a year. Since the model period is annual, this is consistent with including seemingly non-liquid assets in the definition of wealth. Second, running down assets to smooth the effects of wages shocks in the model can be thought of in terms of borrowing against collateralizable assets in the real world.

[11] Implicitly, my estimation procedure provides an estimate of classical measurement error in the sample, as the difference between the overall variance of consumption, earnings or hours in the model versus in the data.

[12] The smoothed line is intended merely to help guide the reader's eye in recognizing the patterns in places where the data is relatively noisy. It is not meant to represent an estimate of an underlying smooth relationship between inequality and age. The structural estimation procedure outlined in section 4.3 should be thought of as the method through which I extract the underlying evolution of this distribution over the lifecycle.

sulting from a combination of the cumulation of persistent idiosyncratic shocks and a transitory shock whose variance varies with age. Persistent shocks to wages generate both increasing wage inequality and a downward sloping pattern in the cross-sectional correlation between wages and hours, documented below.

In the top right panel of Figure 1, I show the variance of log consumption over the life-cycle, plotted on the same scale as the variance of log wages. The key feature to note from this graph is that while there is an increase in consumption inequality over the working years (around 0.03 in the variance of logs), the magnitude of the increase is less than half the size of the increase in wage inequality over the same period and is not significantly different from zero. Note that that the age-profile of consumption inequality that I report here is significantly flatter than most existing estimates.[13] This makes it more difficult for a model with missing insurance markets and persistent shocks to provide an explanation for the data. In this sense, the goal posts in this paper have shifted relative to those in Storesletten, Telmer and Yaron (2004).

The evolution of cross-sectional inequality in annual hours worked is shown in the bottom panel of Figure 1. The most interesting feature of inequality in labor supply over the lifecycle is the sharp decrease (by around 50%) over the first 15 years in the labor market. The decrease in the variance of log hours at young ages is a robust feature of the data.[14]

Figure 2 shows the contemporaneous joint second moments of consumption, hours and wages over the lifecycle. The top panels show the correlation of consumption with wages and hours respectively. Note that the correlation of consumption with wages is roughly increasing over the lifecycle, but this increase takes place during the second half of the working life. In contrast, the experience profile of the correlation between consumption and hours is basically flat. Finally, the correlation between wages and hours shows a clear decreasing pattern over the lifecycle. Amongst other things, the slope of the wage-hours correlation is informative as to the relative size of the income and substitution effect in hours responses to individual wage changes.

The definition that I have adopted for consumption excludes all elements of durable consumption expenditures. In Figure 3, I provide evidence that the patterns in the second moments of consumption over the lifecycle are not sensitive to this choice of definition. In particular, I consider two alternate definitions of consumption - one that includes total annual consumption expenditures, including purchases of durables, and an intermediate definition which includes non-durable consumption plus imputed services from housing and vehicles. While the *level* of inequality (not shown) is different across these definitions, the three figures show that all three definitions give essentially the same *lifecycle profiles* for the variance of consumption and the

[13]See, for example, Deaton and Paxson (1994), Storesletten, Telmer and Yaron (2004) and Guvenen (2006). Part of this difference is due to differences in the choice of equivalence scale. The remainder is sue to the choice to control for time rather than cohort effects. Heathcote, Storesletten and Violante (2006) for a discussion of the sources of differences in estimates of the age-profile of consumption inequality.

[14]Heathcote, Storesletten and Violante (2005) document a similar pattern for raw inequality, using age as the lifecycle dimension. Moreover, the decrease in the variance of hours is not being driven by excessive job turnover rates or more frequent short spells of search-related unemployment at young experience levels. The general pattern persists when I restrict the sample to individuals who had not experienced any job changes during the relevant sample year.

correlation of consumption with wages and hours.

The salient features of the distribution of wealth over the life-cycle are shown in Figure 4. For households with positive wealth, I measure wealth inequality in terms of the Gini coefficient, which decreases by around 0.2 over the lifecycle, shown in Panel A. Panel B contains the fraction of the population with zero or negative wealth. This decreases from around 20% at the time of entry to the labor market to less than 5% by retirement. Panels C and D of Figure 3 show the correlation of log wealth with wages and hours respectively, for households with positive total assets. There is a large increase (around 0.4) in the correlation between log wealth and log wages during the first half of the working years, whereas the correlation with hours is flat or slightly decreasing over the same period. Huggett (1996), Quadrini and Rios-Rull (1997), Diaz-Gimenez, Quadrini and Rios-Rull (1997) and Rodriguez, Diaz-Gimenez, Quadrini and Rios-Rull (2002) have all documented features of the wealth distribution in the US using data from the Survey of Consumer Finances (SCF). These authors place most of their emphasis on the overall wealth distribution, rather than the evolution of the distribution over the lifecycle. However, it is reassuring to note that the patterns in Figure 4 appear consistent with SCF data. For example, Table 8 of Rodriguez et al. (2002) reports that the wealth Gini decreases by around 0.2 from age 25 to age 60, which suggests that the magnitude of the decrease in cross-sectional wealth inequality is not sensitive to the source of data on wealth. Moreover the overall distribution of wealth in my PSID sample is comparable with that measure in the SCF. For example, Tables 1 and 2 of Rodriguez et al. (2002) report that the overall wealth Gini is 0.80, the mean-to-median ratio is 4.03 and the correlation of earnings with wealth is 0.47. The corresponding statistics from my sample are 0.76, 3.39 and 0.39.

Taken together, Figures 1, 2 and 3 paint a complex picture of the evolution of cross-sectional inequality in wages, consumption, labor supply and wealth. The task of providing a coherent explanation for these facts is a challenging one, and an important one. Each moment in this distribution has a theoretical counterpart in standard economic models, and individually, almost all have been the focus of existing studies. However, I view the current study as being at the foot of an ambitious research agenda that uses all the relevant information that we have available in order to understand the features of inequality, its relation to lifecycle shocks and choices, and the potential effects of alternative government policies. Understanding the joint properties of these moments is crucial in this respect.

In this section I have documented, amongst other facts, an increasing profile of wage inequality, a much flatter profile of consumption inequality and a decreasing profile of hours inequality, together with a decreasing covariance between wages and hours and an increasing covariance between consumption and wages. Accounting for all these features simultaneously is a challenging goal for any economic theory. The next stage is to ask how far the standard incomplete markets model, which has become the workhorse of quantitative heterogenous agents macroeconomics in recent years, can go in accounting for these facts.

3 Model

The model is an overlapping-generations version of the incomplete markets economies studied by Bewley (1986), Hugget (1993) and Aiyagari (1994), extended to allow for endogenous labor supply. The economy is populated with T overlapping generations, each generation consisting of a continuum of agents, indexed by i. The unconditional probability of surviving to age t, is denoted by $S_t = \prod_{j=0}^{t-1} s_j$, where s_t is the probability of surviving to age t conditional on having survived to age $t-1$. Agents begin working at age $t = 1$ and retire at age T^{ret} if they are still alive. All agents are dead by age T.

Agents have preferences over stochastic streams of a single, non-storable consumption good, c_t, and leisure, l_t, given by

$$E \sum_{t=1}^{T} \beta^t S_t u(c_t, l_t)$$

Before retirement, agents have a time endowment which is normalized to 1 and can be split between working and leisure activities. After retirement, agents value consumption only. In the benchmark model, period utility is assumed to be additively separable in consumption and leisure, each with constant elasticities of intertemporal substitution[15]:

$$u(c, l) = \frac{c^{1-\gamma}}{1-\gamma} + \varphi_i \frac{l^{1-\sigma}}{1-\sigma}$$

In a robustness exercise in section 7 I also consider the effect of non-separability between consumption and leisure in the utility function. φ_i is an individual specific weight on leisure. It follows a log-normal distribution in the population:

$$\log \varphi_i \sim N\left(\mu_\varphi, \sigma_\varphi^2\right)$$

While they are working, agents receive an annual endowment of labor efficiency units, $e_{i,t}$, which evolves according to a known stochastic first-order Markov process. An agent with efficiency $e_{i,t}$ who supplies labor $h_{i,t} \geq 0$ hence contributes $e_{i,t}h_{i,t}$ units of labor to the aggregate labor input used by the representative firm. The after-tax wage rate per unit of effective labor is given by W. The benchmark model for individual labor productivity consists of four components:

$$\begin{aligned}
\log e_{i,t} &= \kappa_t + \alpha_i + z_{i,t} + \varepsilon_{i,t} \\
z_{i,t} &= \rho z_{i,t-1} + \eta_{i,t}
\end{aligned}$$

The first component, κ_t, is a non-stochastic experience profile for the mean of wages which is assumed to be the same for all individuals. It is intended to capture returns to experience over

[15] For $\gamma \neq 1$, these preferences are not consistent with balanced growth. In particular, with $\gamma > 1$, they predict that the fraction of time devoted to labor will fall over time. For my sample of males, the average fraction of time spent working decreased by 2.5 percentage points betwen 1968 and 1997.

the life-cycle and depreciation of relative skills as new cohorts enter the workforce. The second component, $\alpha_i \sim N\left(\frac{-\sigma_\alpha^2}{2}, \sigma_\alpha^2\right)$, is an individual specific fixed effect that is intended to capture difference in unobserved fixed skills that command a return in the labor market that is not captured through the effects of education. Lastly, I allow for two types of productivity shocks. $z_{i,t}$ is an AR(1) component of individual labor market risk with innovations, $\eta_{i,t} \sim N\left(\frac{-\sigma_\eta^2}{2}, \sigma_\eta^2\right)$, that have constant variance. The transitory shock, $\varepsilon_{i,t} \sim N\left(\frac{-\sigma_{\varepsilon,t}^2}{2}, \sigma_{\varepsilon,t}^2\right)$, is distributed independently over time with a variance that is allowed to depend on labor market experience. Persistent shocks are intended to capture the effects of events such as health shocks and career promotions and demotions. On the other hand, examples of transitory shocks may include short spells of unemployment, opportunities for overtime, bonuses, commissions and other transitory components of compensation.

Agents can trade quantities of a single risk-free asset, $k_{i,t}$. A financial intermediary pools savings in every period and returns the pooled savings, with interest accumulated at a gross after-tax rate, R, proportionally to those agents who are still alive at the beginning of the following period. In this way, annuity markets are perfect, so that there is no independent effect of mortality risk on savings decisions. During the period, the intermediary rents the pooled savings to the representative firm who uses it as a capital input in production.

In the benchmark model, I impose a no-borrowing constraint which restricts agents' wealth to be non-negative in all states of the world. However I calibrate wealth levels in the economy in terms of total net worth, including non-liquid assets such as housing. This is in contrast to focusing on just financial wealth. It is thus appropriate to interpret the borrowing constraint as allowing agents to borrow up to the amount for which they can provide assets as collateral, from a non-modelled banking sector. In section 7, I show that all the qualitative results are robust to the polar opposite assumption of unrestricted borrowing, subject to the constraint that agents who live to age T die with non-negative assets with probability 1.

Agents are born with an initial wealth endowment, k_0, which is assumed to follow a lognormal distribution in the population. Panel A of Figure 3 showed that inequality in wealth is highest at the time of entry to the labor market and decreases substantially over the lifecycle. This fact is not consistent with an assumption that all agent in the model are born with zero wealth. Moreover, the mean wealth of young households is positive in the data - the average wealth-to-earnings ratio of individuals with less than 5 years experience is 0.89. To account for these facts I restrict the mean of the initial wealth distribution to match the wealth-earnings ratio of young households in the data and estimate the variance of the distribution along with the other structural parameters.

There is also a government sector which collects proportional taxes on labor earnings and interest earnings at rates τ_L and τ_K respectively. These are used to fund a progressive pension system that pays each agent a constant fraction of the persistent and permanent components of their final wage. The replacement rate is based on US old age insurance and is taken from Storesletten, Telmer and Yaron (2004). An exogenous amount G, determined residually, is spent on non-valued government expenditure.

On the production side of the economy there is a single representative firm with Cobb-Douglas production function, $Y = K^{\theta}N^{1-\theta}$, where K is the aggregate amount of capital used as input in production and N is the aggregate amount of domestic labor supplied, measured in efficiency units. Capital depreciates at a constant rate δ. There is no aggregate uncertainty.

I focus on an open-economy equilibrium in which the gross after tax interest rate, R, is determined exogenously by world factors. The choice of optimal factor inputs in production then determine the wage rate per efficiency unit W. The remaining elements and definition of a stationary equilibrium are standard - individuals optimize, the labor market clears and the induced probability measure of agents over the state space is constant over time. The aggregate capital stock includes domestic savings and capital rented on world markets for use in production.[16] The model is solved numerically using an extension of the endogenous gridpoints method of Carroll (2005) to allow for endogenous labor supply. Full details can be found in Appendix C.

4 Estimation

Virtually all of the existing literature that has used structural models to investigate cross-sectional distributions over the lifecycle have either set parameters exogenously or followed a calibration approach. That is, parameters are chosen based on aggregate moments, averaged over agents of different ages, rather than on the basis of the lifecycle patterns that are the subject of investigation.[17] Two notable exceptions are Gourinchas and Parker (2000) and Imai and Keane (2004). These papers explore the age profiles of the first moments of consumption and labor supply respectively, in models with a dynamic savings decision. Gourinchas and Parker (2000) estimate structural parameters using simulated method of moments while Imai and Keane (2004) are able to implement a simulated maximum likelihood estimation strategy. This paper is complementary to these in that it moves the focus from the first to the second

[16] Storesletten, Telmer and Yaron (2004) have shown in a model with exogenous labor supply that the single most important factor in determining the sustainable amount of risk sharing in this class of economies is the level of domestic capital. This is intuitive - the higher the amount of wealth that is available to agents to accumulate for precautionary reasons, the more successful they are at smoothing consumption in the face of idiosyncratic uncertainty. In fact it turns out that for a given set of parameters, the interest rate is only important for the patterns of lifecycle inequality to the extent that it affects the amount of domestic capital in equilibrium. It is hence important in an exercise such as this to ensure that the aggregate level of domestic savings is commensurate with the level we actually observe in the US economy. The most common approach to generating a realistic level of domestic capital, and the one that I follow here, is to use the discount factor, β, to match the capital-income ratio in the model to the corresponding figure in the data. Provided that this is done, whether the interest rate is such that the domestic capital market clears has no impact on cross-sectional inequality over the lifecycle. Note that the computational cost of estimating the open economy model is exactly the same as if β were estimated along with the other structural parameters, and a closed economy was assumed. Both versions require iteration using aggregate variables (either over β or over R) for each guess of a parameter vector.

[17] Examples of similar models to this one that follow a calibration strategy include Storesletten, Telmer and Yaron (2004), Low(2005), Pijoan Mas (2006) and Fernandez-Villaverde and Kruger (2005)

moments of the distribution of consumption and hours.[18]

By far the biggest barriers to the estimation of Bewley-Hugget-Aiyagari incomplete markets models are the computational challenges that must be overcome. In general, even if the criterion function that underlies the estimation strategy can be evaluated quickly enough for estimation to be feasible, the simulation based approach generates criterion functions that are generally non-smooth, locally flat and contain many local minima. This makes classical estimation difficult and sensitive to initial parameter guesses. I overcome these problems by using the quasi-Bayesian Laplace Type Estimator (LTE) of Chernozhukov and Hong (2003) which allows Monte-Carlo Markov Chain (MCMC) techniques to be applied in non-likelihood settings. The estimator is based on the same simulated GMM objective function as is standard in the literature on simulation-based estimation.[19] However, rather than attempting to minimize the criterion function directly, the criterion function is used to define a Markov Chain, which converges to a limiting distribution that can be used to derive point estimates and confidence intervals for parameters. A full description of the estimation method, together with a derivation of formulas for asymptotic standard errors can be found in Appendix E. The estimation technique provides a number of benefits compared to existing methodologies and the success of the strategy here should be encouraging to other researchers attempting structural estimation of computationally demanding models. In addition to being robust to the presence of multiple local minima and a non-smooth objective function, it allows for priors to be specified for subsets of parameters and MCMC techniques to be used, even though specification of a full likelihood function is not possible.

Why should structural estimation be preferred to calibration in this type of exercise? First, some of the parameters, particularly those related to the distribution of preference heterogeneity are not easily identified outside the model and would be difficult to calibrate externally. Other important parameters, such as the Frisch elasticity of labor supply, have commanded a wide range of estimates in the existing literature, with estimated values being sensitive to the aspect of behavior and particular model being studied. In fact even within the model described in Section 3, different dimensions of the data are consistent with different parameter values.[20] Moreover, the evolution of the distribution of consumption, hours and wages over the lifecycle is just as, if not more, informative about parameter values (in particular, the two elasticities of substitution, γ and σ) than are average moments. It seems appropriate to choose parameters using the same model that is under investigation, on the basis of the dimensions of the data that we wish to understand.

Second, structural estimation allows confidence intervals to be constructed for parameter

[18] Heathcote, Storesletten and Violante (2006) also estimate a structural model using data on inequality over the lifecycle. However their economy is one of partial insurance, in contrast to the bond economy being studied here. As such their work can be viewed as complimentary to this study.

[19] See, for example, Pakes and Pollard (1989).

[20] For example, the flat profile of the covariance between consumption and wages at young ages and the small cross-sectional correlation between hours and wages are suggestive of a low labor supply elasticity, whereas the relatively high variance of individual hours changes at young ages suggests much higher elasticity levels.

estimates. These can then be used to help guide the choice of robustness exercises, and can be informative as to the relative sensitivity of different parameters for observed behavior.

A third reason for preferring a structural estimation stems from our goal of seeking to understand the features of the lifecycle distribution of consumption, hours and wages simultaneously. By specifying an overidentified set of the model's restrictions, many of which are in conflict with each other *vis-a-vis* the data, we can use the failure to satisfy all restrictions simultaneously as a diagnostic tool for choosing where to enrich and expand the model. Estimation allows us to choose the feature of the data that we are interested in explaining, in this case the lifecycle patterns of cross-sectional distributions and assess whether the model can provide a coherent explanation. Attention can then be focused on the mechanisms at work in the model, rather than the choice of parameter values.

Below I separate the model's parameters into two subsets. The first set are those that are not directly relevant for, nor easily identifiable from, changes in the cross-sectional distribution of consumption, wages and hours over the lifecycle. These are technology, government and demographic parameters, and are fixed outside the model. The second set of parameters are those that govern preferences, the individual productivity process and the initial wealth distribution. These are estimated in two stages using second moments of the distribution of wages, hours and consumption. In the first stage, the parameters of the productivity process are estimated directly from the auto-covariance structure of wages over the life-cycle. In the second stage, the remaining parameters are estimated from the lifecycle profile of the moments in Figures 1 and 2 using a Laplace-type estimator with a GMM criterion.

4.1 Fixed Parameters

The model period is assumed to be annual. I define age as years of potential labor market experience and set the first period in the model to correspond to the third year of potential experience in the data. I abstract from modeling the transition from education to the workforce as this may differ substantially across education groups and time periods. Moreover, the sample sizes in the CEX and PSID are small at experience levels less than three and generate unreliable estimates of inequality. Surviving agents retire after 38 years of work and are dead 78 years after entering the labor force (this corresponds to retirement at age 60 and death at age 100 for an individual entering the labor force at age 22). Mortality rates are chosen to match male US mortality rates from the 2000 census.[21]

I follow the majority of the existing literature in setting the capital share in production, θ, to 0.33, and the depreciation rate, δ, to 6%. The world interest rate is set exogenously at 4.5% after tax. I follow Domeij and Heathcote (2004) in setting the labor tax rate at 40% and the capital tax rate at 27%. A summary of all the exogenously fixed parameters is shown in Table 1.

[21] Mortality rates are taken from Table 3 of Kochanek et al. (2004). They are adjusted so that all cohort members are dead by age 100.

4.2 Wage Process

The benchmark wage process has been specified to capture salient features of the autocorrelation function for wages. This shows a sharp drop after the first lag, which suggests the presence of a purely transitory component, and a roughly exponential decline from the second lag onwards, which suggests the presence of an auto-regressive component. Furthermore, the autocorrelation of wages is non-zero even at long lags, suggesting an individual fixed effect. I estimate the parameters in two stages. In the first stage I remove time effects, education effects and experience effects (together with interactions) from the level of log wages. The estimated sequence of experience effects, $\{\kappa_t\}_{t=1}^{T^{ret}}$, are used as the deterministic experience profile in the model, re-scaled so that the average productivity in the population is 1. Panel A of Figure 5 shows that the estimated experience profile for wages is concave and increasing up to experience level 30, at which point it flattens out and then decreases slightly.

I use the autocovariance function of the residuals from the first stage estimation as the basis for estimation of the variances of the three components and the auto-regressive parameter in the second stage. I calculate the autocovariance of the residuals in each experience/lag cell and use a minimum distance estimation algorithm that minimizes the distance between the empirical and theoretical autocovariance matrix at each age.[22] I only include moments for which there are at least 50 observations. In total I match 663 moments that include up to 26 lags. I use a diagonal weighting matrix that weights each moment by the number of observations that were used in its calculation.[23] The results of the estimation are displayed in Table 2. Standard errors are computed using a block bootstrap with 250 repetitions to account for the use of estimated residuals from the first stage and the fact that asymptotic standard errors may be heavily biased in small samples.

Panel B of Figure 5 shows the implication of these estimates for the composition of the variance of wages at each point in the lifecycle. The variance of transitory shocks is decreasing over the first five years in the labor market and is roughly flat thereafter. Both transitory shocks and fixed skills account for approximately the same fraction of cross-sectional wage variation.[24] The relative contribution of persistent shocks increases over the lifecycle to account for about 50% of total variation after 20 years. The estimate for the auto-regressive parameter is 0.94, which is toward the low end of estimates in the existing literature.

An alternative view of the increase in the variance of wages over the lifecycle is that it is the result of heterogeneous (but deterministic) age-wage profiles across individuals, generated, for example, by cross-sectional differences in the pattern of human capital investments. For the purpose of comparison, I also include results from the estimation the benchmark model,

[22] In appendix D I show that the parameters of the wage model are identified from the set of moments used in estimation.

[23] Altonji and Segal (1996) show that the finite sample bias of the optimal GMM weighting can be large in this type of model and suggest the use of an identity weighting matrix.

[24] If there is classical measurement error in wages then this will be incorporated into the estimates of the variance of transitory shocks.

extended to allow for this type of heterogeneity[25]. Allowing for profile heterogeneity results in a lower estimate for the persistence of wage shocks. The estimated variance of such heterogeneity is small - about an order of magnitude lower than that found in earnings by Baker (1997), Haider (2001) and Guvenen (2006). Moreover, the estimated covariance between the slopes and intercepts of individual profiles is estimated to be positive, in contrast with basic intuition about the source of such heterogeneity. An alternative test for the presence of profile heterogeneity is a positive auto-correlation of wage growth at long lags. There is no evidence of this in my sample. Note, however, that in this paper I model wages, whereas existing studies that have found evidence of profile heterogeneity have focussed on earnings. As a result, I focus on the persistent shocks explanation of an increasing variance of wages over the lifecycle from here on.

4.3 Structural Estimation

The remaining five parameters are the coefficient of relative risk aversion for consumption, γ, the inverse of the Frisch elasticity of substitution for leisure, σ, the mean and variance of the distribution of the relative taste for leisure, $E(\varphi)$ and $V(\log\varphi)$, and the variance of the initial wealth distribution, which I express in terms of the implied Gini coefficient, Gini(k_0). I estimate these parameters using a Laplace-type estimator with a GMM objective function and uniform priors.

In selecting the moments that are matched, I concentrate attention on the joint distribution of consumption, wages and hours. This generates 5 second moments (the variance of wages is already included in the system of moments from the first stage) which I express as deviations from their mean values over the lifecycle. Each of the 5 moments is targeted at 30 experience levels. In addition, I include the average fraction of time spent working in the economy. This last moment is included to help identify the mean of the distribution of the relative taste for leisure, which is not strongly identified from data on second moments alone. The value targeted is 0.45 and is matched exactly in the estimation. In total, the criterion function is comprised of 151 moments. In addition, β is calibrated within the estimation procedure so that the aggregate capital/income ratio in the model is the same as that in my PSID sample. This value is 3.115 and is in line with values targeted by Storesletten, Telmer and Yaron (2004) and Heathcote, Storesletten and Violante (2004).

Note that I exclude the lifecycle features of the wealth distribution from the system of targeted moments. This is for three reasons. First, the sample sizes are substantially smaller for the wealth sub-sample and the data more noisy than for the other moments. Second, wealth is not directly relevant for welfare - inequality in lifetime welfare in the model is generated through the joint distribution of consumption and hours. Hence these distributions are the focus of parameter estimation. Moreover, by excluding the wealth distribution as an explicitly targeted feature of inequality, it can then be used as an "out of sample" test of the model's

[25] In appendix E, I show that the parameters are still identified in the more flexible model that allows for profile heterogeneity

predictions. Third, it turns out that the salient features of the wealth distribution over the lifecycle are matched extremely well, even when it is not used in estimating parameters. It is the distribution of consumption, wages and hours that provides a challenge for this class of models.

In the benchmark estimation I choose an identity weighting matrix. In section 7, I explore the robustness of the estimates to a different choices of weighting matrix - one whose elements are the diagonal elements of the optimal weighting matrix. Because certain moments are in conflict with each other in the estimation (see section 5), the choice of weighting matrix is not negligible for the results. Given that the data come from two different data sets with very different sample sizes, an identity weighting matrix is preferred as the baseline.[26] The baseline parameter estimates are shown in Table 3. Standard errors are calculated as described in Appendix E.[27]

The estimated coefficient of relative risk aversion, 3.90, and the implied estimate of the intertemporal elasticity of substitution for consumption, 0.26, is fairly standard in the literature. The implied Frisch elasticity for leisure is 0.11, which is towards the low end of estimates for males in the PSID.[28] The estimate for the variance of preference heterogeneity in the taste for leisure, 0.03, implies that this type of heterogeneity does not play a large role in accounting for cross-sectional moments in the model.[29] This finding can be attributed to the fact that the estimation strategy targeted only the profile of hours inequality over the lifecycle, (and not the level of hours inequality), for which preference heterogeneity plays a minimal role.[30]

The estimated inequality in wealth, at the time of entry to the labor market, implies a Gini coefficient of 0.80. Recall that in the benchmark model, borrowing is not allowed and thus households have non-negative wealth. The corresponding figure for households in the PSID sample with less than 5 years experience is 0.88 and 0.73 when restricted to households with non-negative wealth. Given that these parameters are estimated entirely from the lifecycle patterns in the cross-sectional distribution of wages, hours and consumption (not using data on wealth) it is particularly encouraging that the model implies a realistic level of initial wealth inequality. The discount factor required to match the aggregate capital/income ratio in the data is 1.04.

[26] Sample sizes in the CEX are substantially smaller than in the PSID. This means that with the optimal weighting matrix, the moments that involve consumption are under-weighted relative to those that do not. An equally weighted objective function does not suffer from this bias.

[27] The standard errors that I report are estimates of the asymptotic standard errors, accounting for the choice of a non-optimal weighting matrix. The effect of the use of residuals from the first stage regression is accounted for by bootstrap. See Appendix E for details.

[28] See Blundell and MaCurdy (1999) for a survey of the literature on estimating labor supply elasticities.

[29] See figure 10 and the discussion in 6

[30] When the weighting matrix is chosen to overweight the PSID data compared to the CEX data, as is suggested by the optimal weighting matrix, a lower estimate is obtained for risk aversion (1.66) and a higher estimate for the Frisch elasticity (0.18). This is because a greater emphasis is placed on matching the moments that include hours, primarily the wage-hours correlation. These estimates are consistent with a flatter profile of this moment, but a larger increase in the variance of consumption.

5 Benchmark Model Fit

In this section I discuss the fit of the benchmark model and investigate the restrictions that the model places on the lifecycle properties of the distribution of consumption, wages and hours. The fit of the estimated moments in the benchmark model is shown in Figure 6. These are plotted as deviations from their mean values over the lifecycle.

5.1 Variance of Consumption

The increase in the variance of consumption in the model is slightly greater than the corresponding increase in the data (0.06 vs 0.03), although the increase is roughly within the confidence bounds. One interpretation of this finding is that even single worker male households in the US have access to more insurance possibilities than those afforded by just a risk-free bond. This finding is in accordance with recent studies[31] that suggest that economies with partial insurance may more closely resemble the risk-sharing possibilities individuals are faced with.

One reason why the model understates the lifecycle rise in consumption inequality, in contrast to the findings in Storesletten, Telmer and Yaron (2004), is the much flatter profile of consumption inequality that the model is being asked to generate. Another, more salient difference is that the results here are generated by attempting to match the profiles of all six moments in Figure 6 simultaneously. As is discussed in section 7, there are parameter values that can generate the same rise in consumption inequality as in the data, but these yield counterfactual predictions for other moments of interest.

5.2 Distribution of Earnings, Hours and Wages

The top right panel in Figure 6 shows the experience profile of the variance of labor earnings generated in the model and in the data, which is matched well. Earnings inequality can be separated into components due to wages and hours

$$V\left[\log y_{i,t}\right] = V\left[\log w_{i,t}\right] + V\left[\log h_{i,t}\right] + 2COV\left[\log w_{i,t}, \log h_{i,t}\right] \tag{1}$$

Since the distribution of wages is estimated independently in a first stage, the fit is perfect. Moreover, the middle right panel of Figure 6 shows that the model reproduces a declining profile in the covariance between wages and hours which is steeper than in the data. This implies that wage-hours covariance is overestimated at young ages and underestimated at older ages. Inspection of the variance of hours in the left middle panel then reveals that the reason why the models is nonetheless able to generate the correct profile of earnings inequality is that it fails to reproduce the downward sloping profile for the variance of hours.

Why does the covariance between wages and hours decline in the model? With the estimated value for risk aversion, $\gamma > 1$, the income effect of hours responses to wage changes is negative.

[31] Examples include Blundell, Pistaferri and Preston (2006), Heathcote, Storesletten and Violante (2006) and Attanasio and Pavoni (2006).

Agents face two types of wages shocks in the model. There is little or no income effect associated with transitory shocks. Hence hours should covary positively with transitory shocks due to the substitution effect. Persistent shocks, however, do generate an income effect. We saw from Figure 5, panel B, that the relative mix of persistent versus transitory shocks in wages increases substantially over the lifecycle. This generates a decreasing pattern in the covariance of wages and hours. A second effect operates through the pension system. Since the pension replacement rate is determined by the final level of the persistent component, as an agent nears retirement persistent shocks become more and more important for lifetime income, thus inducing larger (negative) hours responses. Lastly, I document below that the model matches the decreasing profile in inequality in wealth and the sharply increasing profile in mean wealth. All of these features serve to decrease the covariance of wages and hours at older ages.

Before exploring the model's counterfactual predictions for hours inequality early in the career, it will be helpful to look at the fit of the model in terms of the covariances involving consumption and the wealth distribution.

5.3 Covariance of Consumption with Wages and Hours

The bottom left panel of Figure 6 shows that the model generates an increasing covariance between wages and consumption. The magnitude of this increase is consistent with the data in the second half of the working life. However, the model significantly overstates the rise in this covariance at young experience levels. The reason for this difference is related to the pattern of hours inequality in the model and is discussed below. The covariance of consumption and hours decreases in the model, although the decrease is within the confidence bands for the profile in the data.

The explanation for these patterns is most easily seen by noting that consumption responds very little, or not at all, to transitory wage shocks, since these have only a small effect on total lifetime wealth. However, persistent wage shocks do induce a consumption response, which generates a positive correlation between consumption and wages. The increasing variance of wages and increasing fraction due to persistent shocks then generates an increasing cross-sectional covariance of consumption and wages. Moreover, since $\gamma > 1$, a positive persistent wage shock generates both a negative hours response and a positive consumption response. Hence the covariance of consumption with hours falls as the relative importance of persistent shocks rises over the life-cycle.

5.4 Additional Moments: Fit of Wealth Distribution

Recall that the model was estimated using the second moments of the distribution of consumption, hours and wages. In this section, I assess the fit of the model in terms of the wealth distribution. For the benchmark model, this is shown in Figure 8. Panel A shows that wealth inequality, as measured by the Gini coefficient, decreases at roughly the same rate in the data and in the model.[32] Panel B shows that the model generates a declining fraction of households with zero wealth. This proportion is lower than in the data because the benchmark model does not allow borrowing, so households with non-positive assets in the model are those for whom the no-borrowing constraint binds. In a robustness exercise in Section 7, I consider an alternative model in which households are allowed to borrow up to their natural borrowing limits. Finally, the bottom two panels show that the model is able to generate the same concave, increasing covariance between log wealth and log wages as in the data, as well as the slightly downward sloping profile of the covariance between log wealth and log hours.

With regards to the fit of the wealth distribution, it is important to bear in mind that none of these features are explicitly targeted in the structural estimation. That the model is able to replicate the broad patterns of wealth inequality over the lifecycle, with parameters that are identified from data on wages, hours and consumption, suggests that uninsured wage risk is an important factor in the determination of the distribution of wealth in the US economy, and should provide us with some confidence in the theory of precautionary savings as a motivation for wealth accumulation.

5.5 Understanding the Shortcomings of the Model

In this section I delve a little deeper into the mechanisms at work in the model. In order to understand where and how the model should be enriched, it is important to understand the sources of its failures, particularly in respect of the model's counterfactual predictions for the pattern of hours inequality over the lifecycle. Most of the important restrictions that the neo-classical labor supply model places on cross-sectional distributions can be seen clearly by manipulating the intratemporal first order condition for leisure. For a general utility function, $u(c_{i,t}, l_{i,t})$, this condition is

$$u_c(c_{i,t}, l_{i,t}, \varphi_i)\, w_{i,t} = u_l(c_{i,t}, l_{i,t}, \varphi_i) \tag{2}$$

For the special case of separable preferences of the form in the benchmark model, this condition can be re-arranged to give

$$\gamma \log c_{i,t} - \log w_{i,t} = \sigma \log l_{i,t} - \log \varphi_i \tag{3}$$

Equation (3) then implies the following restrictions on the evolution of the joint distribution of $(c_{i,t}, l_{i,t}, w_{i,t})$ over the lifecycle[33]:

$$\begin{pmatrix} \Delta V[\log l_{i,t}] \\ \Delta COV[\log l_{i,t}, \log w_{i,t}] \\ \Delta COV[\log l_{i,t}, \log c_{i,t}] \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2} & \frac{\gamma^2}{\sigma^2} & -\frac{2\gamma}{\sigma^2} & \frac{2\gamma}{\sigma^2} \\ -\frac{1}{\sigma} & 0 & \frac{\gamma}{\sigma} & 0 \\ 0 & \frac{\gamma}{\sigma} & \frac{-1}{\sigma} & \frac{1}{\sigma} \end{pmatrix} \begin{pmatrix} \Delta V[\log w_{i,t}] \\ \Delta V[\log c_{i,t}] \\ \Delta COV[\log w_{i,t}, \log c_{i,t}] \\ COV[\log \varphi_i, \Delta \log c_{i,t}] \end{pmatrix} \tag{4}$$

[32] The model slightly overestimates the reduction in wealth inequality during the first five years in the labor market. However the estimated Gini coefficient at the time of entry matches the value in the data. This implies that the overall level of wealth inequality predicted by the model is slightly lower than in the US.

[33] Similar sets of restrictions can be derived with more general preferences. See section 7.

The system of equations in (4) is a set of restrictions on changes over the lifecycle in the six moments of the distribution of consumption, wages and leisure, together with the cross-sectional covariance of heterogeneity in the relative taste for leisure with consumption changes. This system is arranged so that it shows the moments involving leisure as a function of the other moments in the system. These equations can be used to derive approximate equations for the moments involving hours. Starting with the covariance between wages and hours, we have:

$$\Delta COV\left[\log h_{i,t}, \log w_{i,t}\right] \approx \frac{1}{\sigma}\Delta V\left[\log w_{i,t}\right] - \frac{\gamma}{\sigma}\Delta COV\left[\log w_{i,t}, \log c_{i,t}\right] \tag{5}$$

Equation (5) shows that in order for the model to reconcile a downward sloping covariance between hours and wages with an increasing variance of wages, it is necessary that the covariance between wages and consumption must be increasing. However since the size of this increase is restricted by the data, the slope of $COV\left[\log h_{i,t}, \log w_{i,t}\right]$ identifies the relative size of γ and σ.

To see why this implies a difficulty in matching the decreasing profile of hours inequality, consider an approximation of the first equation in (4):

$$\begin{aligned}\sigma^2 \Delta V\left[\log h_{i,t}\right] \quad &\approx \quad \Delta V\left[\log w_{i,t}\right] + \gamma^2 \Delta V\left[\log c_{i,t}\right] \\ &\quad -2\gamma \Delta COV\left[\log w_{i,t}, \log c_{i,t}\right] + 2\gamma COV\left[\log \varphi_i, \Delta \log c_{i,t}\right]\end{aligned} \tag{6}$$

The first three moments on the right hand side of (6) are all positive in the data. The fourth moment, which describes the relationship between the relative taste for leisure and consumption changes, is negative, but very close to zero and can be ignored to a first order approximation.[34] Two things follow from equation (6). First, the only way for the model to generate a decline in the variance of hours at young ages is if the covariance of wages and consumption is increasing at young ages. Unfortunately, the data reveals that this moment is essentially flat over the first half of the working life. Second, increasing the value of γ, which was argued above to help flatten the profile of consumption inequality, requires an even bigger increase in $COV\left[\log w_{i,t}, \log c_{i,t}\right]$ to generate the same slope for $V\left[\log h_{i,t}\right]$. The result of the estimation is to balance these two outcomes - a very mild incline as opposed to a decline in hours inequality, but an overstated rise in $COV\left[\log w_{i,t}, \log c_{i,t}\right]$ in the first half of the career.

6 Decomposing Inequality

In this section I decompose cross-sectional variation in consumption, hours, wealth and welfare in the benchmark model into its various components. I start by splitting total variation in $X \in \{\log c, \log h, k\}$, denoted by $V(X)$ into a between-age component, $V\left[E\left(X|t\right)\right]$, and a within-age component, $E\left[V\left(X|t\right)\right]$. Table 4 shows that around 57% of the observed cross-sectional variation in consumption is among agents with the same experience levels. The corresponding numbers for the cross-sectional variance of hours and wealth, are 54% and 67% respectively.

[34] Away from the borrowing constraints, agents in the model act like permanent-income consumers. This implies that consumption is close to a random walk and consumption changes are not correlated with fixed individual characteristics such as preference heterogeneity.

Next I decompose the within-age component into the fractions that are due to differences in preferences, initial wealth endowments, fixed skills and the cumulative effect of wage shocks (persistent and transitory). Denote one of these four factors by q. The within-age component of cross-sectional inequality can then be decomposed as:

$$E\left[V\left(X|t\right)\right] = E\left[V\left(X|t,q\right)\right] + E\left\{V\left[E\left(X|t,q\right)|t\right]\right\} \tag{7}$$

The second term on the right hand side of (7) can be interpreted as the between-q component - the component of total within-age inequality in X that is due to cross-sectional differences in q. Table 5 reports $\frac{E\{V[E(X|t,q)|t]\}}{E[V(X|t)]}$ for each of the four shocks.[35] Just over half of the within-age variation in consumption, hours and wealth can be attributed to wage shocks (56%, 51% and 53% respectively). The split of the remaining fraction differs among the three variables. For consumption, fixed skills account for almost all the remaining variation, with financial wealth endowments accounting for only 5%. For labor supply inequality, 2% is due to differences in preferences, 25% is due to fixed skills and 20% can be explained by differences in initial wealth positions.[36] Wealth differences, on the other hand are explained primarily by wage shocks (53%) and initial wealth (39%), with differences in skills accounting for only 7%.

The bottom row in Table 5 shows a decomposition of the variance of discounted realized utility for agents with the mean relative taste for leisure in the economy. Slightly more of the variance is accounted for by wage shocks realized after entry to the labor force (59%) than by fixed skills (39%). Interestingly, initial financial wealth accounts for only 1% of inequality in lifetime welfare. Storesletten,Telmer and Yaron (2004) perform a similar exercise and also conclude that life-cycle shocks account for slightly more of the variation in lifetime utility than does fixed endowments (53%). Huggett, Ventura and Yaron (2006) reach a different conclusion, that more of the variation in lifetime utility is explained by initial conditions than by lifecycle shocks. However, unlike the model here, their model allows for endogenous accumulation of human capital in the determination of wages. Hence some of the variation attributed to life-cycle wage shocks in my model is traced back to its roots in differences in initial human capital endowments and learning abilities in their model.

The discussion until now has ignored the fact that there may be significant differences in the relative contribution of each of the four factors across ages. In Figure 10, I examine these differences by plotting the fraction of variation in each of consumption, hours and wealth that

[35] This method does not deliver an orthogonal decomposition of the four components. However, it turns out that in all cases, the decomposition is approximately orthogonal - the explained fractions sum to between 98% and 100%.

[36] Note that the reason why preference heterogeneity accounts for such a small proportion of overall labor supply inequality is due to the fact that only the lifecycle profiles, and not the levels, of inequality were targeted in estimation. Accordingly, the estimated magnitude of preference heterogeneity is small. The focus in this paper is on the relative contributions of different factors across ages. Targeting levels as well as profiles of inequality, as in Heathcote, Storesletten and Violante (2006), will lead to a different conclusion about the importance of preference heterogeneity in accounting for overall cross-sectional dispersion in labor supply. However it will not change the conclusion that the fraction of hours inequality accounted for by preference heterogeneity does not vary with age.

is explained by each of the four factors, at each point in the lifecycle. The top left panel shows the decomposition of the variance in consumption over the lifecycle. When agents start working, 60% of consumption inequality can be attributed to fixed individual skills. However, the effect of skills slowly wears off and the effect of wage shocks takes over - by the time agents approach retirement, the cumulative effect of wage shocks accounts for over 70% of cross-sectional variation in consumption. The small initial contribution of financial wealth also wears off quickly.

The decomposition of variation in hours tells a different story. Again, the effect of wage shocks increases with age: by retirement it accounts for around 70% of variation of hours in the model. However for hours, initial wealth and fixed effects are also important, particularly at young ages, accounting for around 40% and 30% of the cross-sectional variation, respectively.

The decomposition of wealth inequality is shown in the bottom left panel of Figure 10. From this figure we can learn about the rate at which the effect of initial wealth inequality on overall wealth inequality wears off as a cohort of households age. After 15 years in the labor market, differences in initial wealth endowments still account for half of the total cross-sectional variance in wealth. However by the time that agents reach retirement, 74% of wealth differences can be attributed to wage shocks realized during the working years, and 13% to each of initial endowments of financial wealth and fixed skills.

Finally we can also measure cross-sectional dispersion in the present value of future realized utility for households of different ages in the model. As reported above, at the time of entry to the labor market fixed skills and lifecycle shocks account for around 40% and 60% of the cross-sectional differences in welfare respectively. However, measured as agents enter retirement, almost 80% of these differences can be attributed to lifecycle shocks and just over 20% to fixed skills.

7 Robustness Exercises

In this section I explore the robustness of the estimated parameter values and the conclusions regarding the fit of the model to a number of alternative assumptions. The general result is that the choice of weighting matrix is the most important assumption that is driving the parameter estimates. This is consistent with the finding that no single set of parameter values for this model can lead to results that are consistent with all the dimensions of the data simultaneously. The conclusion that there is an inherent conflict in the model between matching the various moments is robust to the choice of borrowing limits, the specification for wage shocks and preferences and the definition of consumption.

7.1 Weighting Matrix

In section 4.3 it was argued that the identity weighting matrix is appropriate for the estimation strategy adopted. Here I consider an alternative, which is also a diagonal matrix, but whose elements correspond to those on the diagonal of the fully optimal weighting matrix.[37] This matrix has the advantage of placing more weight on moments that are more reliably estimated in the data. One implication is that more weight is placed on the moments that are constructed using the PSID than those estimated from the CEX, since the PSID contains bigger sample sizes. The estimated parameter values are shown in column 2 of Table 3 and the fit of the model in Figure 9.

With less weight placed on the moments involving consumption, the estimation produces a better fit for the covariance between wages and hours and the variance of hours. The downward slope in the wage-hours covariance now matches the data exactly. However, this is achieved through a lower value for risk aversion which dampens the negative income effect of persistent wage shocks and a higher labor supply elasticity which increases the relative size of hours responses to transitory wage shocks. These parameter values result in a significantly worse fit for the moments that involve consumption. In particular, the low risk aversion (high intertemporal elasticity of substitution for consumption) generates an increase in the variance of consumption over the lifecycle of 0.10 compared with 0.06 in the benchmark and 0.03 in the data.

7.2 Borrowing Limits

The benchmark model generated a rise in consumption inequality over the lifecycle that is about one and a half times as big as that in the CEX data. One possible interpretation of this result is that the benchmark model does not afford agents with sufficient mechanisms to smooth the effects of wage fluctuations. This suggests that the extreme assumption of no borrowing (or borrowing up to the value of total assets) may be too restrictive. Here I consider the opposite extreme - unlimited borrowing and lending, subject only to the constraint that with probability 1, agents do not die in debt. This is implemented by allowing agents to borrow up to their natural borrowing limits, which vary by age and state. For the sake of comparability with the benchmark results, and to avoid introducing further complications, I continue to assume that the wealth distribution of the young generation is lognormally distributed with the same average wealth/earnings ratio.[38]

The fit of the model with natural borrowing limits and the estimated parameter values are shown in Figure 10 and Table 3, respectively. The elasticity estimates are both higher than in the benchmark and there is slightly less estimated heterogeneity in both preferences and the variance of the initial wealth distribution. However comparing Figures 6 and 10, it is evident that the overall fit of the model in terms of the targeted moments is almost identical to the fit in the benchmark case. The same is true of the additional moments that involve the distribution

[37] The main reason for restricting attention to a diagonal matrix is that almost all of the off-diagonal elements are either equal to zero or not realiably estimated from the data available. For example the elements that refer to moments that involve consumption at two different experience levels are zero since the CEX is a cross-sectional survey. For those moments from the PSID, the sample sizes become very small, particularly when comparing experience levels that are far apart.

[38] In other words, even though borrowing is allowed agents are restricted to not start their working lives in debt.

of wealth. The results in this section extend the conclusion drawn by Storesletten, Telmer and Yaron (2004), that the choice of borrowing limit is not important for the lifecycle profile of consumption inequality, to a setting with endogenous labor supply. Moreover, the same is true for other dimensions of the distribution of consumption, hours and wages.

7.3 High Risk Aversion

To further stress the point that the failure of the model is not in reproducing consumption inequality *per se*, but rather its inability to produce parameter values that can match all moments simultaneously, Figure 11 shows the fit of the model when the level of risk aversion in consumption is increased (or alternatively, the intertemporal elasticity of substitution is decreased) from 3.90 to 9.0.[39]. Storesletten, Telmer and Yaron (2004) argue that when risk aversion is increased by the same amount (from 2.0 to 7.0), holding aggregate wealth constant, there is no effect on the increase in consumption inequality. However, Figure 11 shows that in the model with endogenous labor supply, the opposite is true. In fact, when γ is increased by this amount, the model generates exactly the same increase in consumption inequality as in the data. Part of the reason for the difference in results is that shocks in Storesletten, Telmer and Yaron (2004) are permanent, so that it is extremely difficult to use savings to smooth earnings shocks, no matter how much agents dislike intertemporal fluctuations in consumption. However, part of the difference is also due to the effects of endogenous labor supply. When hours are flexible, agents are able to respond to an increased desire to smooth consumption by working more when productivity is low and working less when productivity is higher. This is another way of saying that the negative income effect of wage shocks gets stronger as γ is pushed further above 1. Figure 11 also shows that with high risk aversion, although the rise in consumption inequality is as small as in the data, the decrease in the correlation between wages and hours is far stronger than in the data. This is because an increase in γ leads to a stronger negative wealth effect of wage shocks. As the cohort ages, persistent shocks account for an increasing fraction of total cross-sectional wage variation, which generates an increasingly negative correlation between wages and hours.

7.4 Initial Wealth Distribution

In this subsection I explore the role of the initial wealth distribution in accounting for hours inequality. Recall that in the benchmark estimation, the average wealth-to-earnings ratio of the young generation was set at its value in the PSID, 0.89. Figure 12 shows the fit of the benchmark model when the wealth endowment of the initial generation is increased to 20 times average earnings. A high level of initial wealth acts to break the link between consumption and wages at young ages and in doing so generates a lower correlation between wages and consumption and larger cross-sectional inequality in hours early in the career. The left middle

[39] The remaining parameters are left at their benchmark estimates, with the exception of β, which is recalibrated to ensure that aggregate wealth in the economy is the same as in the benchmark.

panel of Figure 12 shows that at the expense of the fit of the other moments, particularly the covariance between consumption and wages at young ages, the model is able to generate decreasing hours inequality over the lifecycle. However, the biggest problem with generating variation in hours using financial wealth is that it simultaneously generates the counterfactual prediction of decreasing average wealth with age. This is because the total amount of wealth in the economy is fixed, but a large amount is held by young agents. This suggests that the introduction of a second form of wealth, such as human capital, may provide an explanation for the observed pattern of hours variation. A model which features learning-by-doing or on-the-job training along the lines of Imai and Keane (2004) may be able to generate decreasing inequality in hours over the course of career. This is something that I leave for future research to investigate.

7.5 Specification for Wage Shocks

One feature of the assumed stochastic process for wages which is non-standard is the inclusion of experience effects in the variance of the transitory shocks. Column 3 of Table 2 shows the parameter estimates for the wage process when the transitory variance is restricted to be constant across experience levels. The estimates for the other wage parameters are essentially the same as in the benchmark case. The resulting structural parameter estimates are shown in column 4 of Table 3. With the exception of a slightly higher labor supply elasticity (driven by the smaller transitory shocks at young ages), the parameter estimates are very close to those from the benchmark estimation. The fit of the model without experience effects in the transitory shocks is displayed in Figure 13. With the exception of the variance of earnings at young ages, the fit is almost identical to the fit of the benchmark.

Another feature of the wage process that warrants discussion is the choice not to include heterogeneity in deterministic experience-wage profiles. However, in section 4.2 it was argued that there is very little evidence for the presence of such effects in this data. As such, estimation of the full model with profile heterogeneity (which would necessitate an extra exogenous state variable) could not be expected to generate very different results to the benchmark specification. Nonetheless, it is important to bear in mind that a specific information structure has been assumed throughout this paper, whereby agents are endowed with no more information about future wage shocks than is evident to the econometrician. Endowing agents with more information than the econometrician in the spirit of Cunha, Heckman and Navarro (2006), or with less information such as in Guvenen (2006), is an important avenue for future work on lifecycle inequality.

7.6 Non-separable Preferences

One may justifiably ask whether the particular specification that I have adopted for preferences is responsible for the restrictions that are inconsistent with patterns of inequality in the data. In particular, could complementarities between consumption and leisure help to generate the downward sloping profile of labor supply inequality at young ages? To address this question, I briefly explore the implications of two alternative specifications for preferences that allow for non-separabilities between consumption and leisure. The conclusion is that neither Cobb-Douglas, nor Constant Elasticity of Substitution (CES) preferences are able to improve on the results with separable preferences.

7.6.1 Cobb-Douglas Preferences

As noted in section 3, the benchmark preference specification is not consistent with balanced growth. It is well-known that Cobb-Douglas preferences of the form

$$u(c,l) = \frac{\left(c^{\eta_i} l^{1-\eta_i}\right)^{1-\gamma}}{1-\gamma}$$

are consistent with balanced growth. To allow for preference heterogeneity in time devoted to work at constant consumption and wage levels, in a manner consistent with that assumed in the separable case, I assume that $\log \varphi_i \sim N\left(\mu_\varphi, \sigma_\varphi^2\right)$ where $\eta_i = \frac{1}{1+\varphi_i}$. The parameter estimates are shown in column 5 of Table 3. For an individual with $\varphi_i = E[\varphi]$, the implied coefficient of relative risk aversion is 18.45 and the implied Frisch elasticity for leisure is 0.46.[40].

The fit of the model is shown in Figure 14. It is immediately evident that the fit is significantly worse than in the benchmark specification - consumption inequality rises by too much over the life-cycle, as does both earnings inequality and the covariance between consumption and hours. Moreover, the covariance between wages and hours and the variance of hours both increase in the model, whereas they decrease in the data. An analogous expression to equation (6) can be derived for the case of Cobb-Douglas preferences:

$$\begin{aligned} \Delta V[\log h_{i,t}] \approx\ & \Delta V[\log w_{i,t}] + \Delta V[\log c_{i,t}] \\ & + 2\Delta COV[\log w_{i,t}, \log c_{i,t}] + 2COV[\log \varphi_i, \Delta \log c_{i,t}] \end{aligned} \tag{8}$$

Examination of (8) reveals that rather than allowing for more flexibility in the joint distribution of consumption, hours and wages through non-separability, the strong structure imposed by the balanced-growth nature of Cobb-Douglas preferences places even more constraints on this distribution. Noting again that the final term in this expression is small relative to the other terms, it is apparent that in order to achieve increasing consumption inequality, and an increasing profile of $COV[\log w_{i,t}, \log c_{i,t}]$, hours inequality must increase by *at least as much* as wage inequality, which is strongly counter-factual.

7.6.2 CES Preferences

One drawback of Cobb-Douglas preferences is that they rely on a single parameter (η_i) to determine both the overall relative taste for leisure and the degree of complementarity between

[40] With Cobb-Douglas preferences, the coefficient of relative risk aversion for an individual i is given by $1-\eta_i+\eta_i\gamma$ and the Frisch elasticity for leisure is given by $\frac{1-\eta_i+\eta_i\gamma}{\gamma}$.

consumption and leisure. A class of preferences which allows more flexibility in this respect are CES preferences[41] of the form

$$u\left(c,l\right)=\frac{1}{1-\gamma}\left[\frac{c^{\nu}}{1+\varphi_i}+\frac{\varphi_i l^{\nu}}{1+\varphi_i}\right]^{\frac{1-\gamma}{\nu}}$$

with $\nu\in(-\infty,1)$. When $\nu<(>)\,0$, consumption and leisure are complements (substitutes). $\varphi_i>0$ determines the relative taste for leisure. These preferences generate an analogous expression to (6) and (8) which is given by

$$\begin{aligned}(1-\nu)^2\,\Delta V\left[\log h_{i,t}\right] \quad\approx\quad & \Delta V\left[\log w_{i,t}\right]+(1-\nu)^2\,\Delta V\left[\log c_{i,t}\right] \\ & +2\left(1-\nu\right)\Delta COV\left[\log w_{i,t},\log c_{i,t}\right]-2\left(1-\nu\right)COV\left[\log\varphi_i,\Delta\log c_{i,t}\right]\end{aligned} \tag{9}$$

Inspection of (9), noting that $1-\nu>0$, reveals that neither complementarity or substitutability is able to reconcile decreasing hours inequality with increasing consumption inequality and an increasing profile of $COV\left[\log w_{i,t},\log c_{i,t}\right]$.[42]

8 Conclusions

This paper started by documenting some facts about the evolution of inequality over the lifecycle for single-worker households with male heads in the USA. These facts reveal a number of interesting and important patterns in the cross-sectional distribution of consumption, hours, wages and wealth. I have argued that attempting to gain an understanding of the sources of these patterns within the context of a model of economic decision making is an important and challenging goal for economic research. The challenge is not in providing a coherent explanation for any one of these facts, but in accounting for all the salient features of inequality simultaneously - consumption choices, labor supply behavior and wealth accumulation.

The model under consideration is the Bewley-Hugget-Aiyagari incomplete markets model with endogenous labor supply. The main mechanisms at work are the precautionary savings motive and labor supply flexibility in the face of persistent uninsurable wage risk. The findings suggest that these mechanisms are consistent with a number of features of the data - a smaller rise in consumption inequality than wage inequality over the lifecycle; a decreasing covariance between wages and hours; and an increasing covariance between consumption and wages. The model also generates predictions for the evolution of the wealth distribution over the lifecycle that are in line with the data. However the data also reveal a number of puzzles which this class of models does not adequately address - the most stark of which is its inability to reconcile the sharp decrease in the cross-sectional variance of hours worked during the first 20 years in the labor market with a large increase in the variance of wages and earnings over the same period.

[41] Although CES preferences are more flexible than Cobb-Douglas in that they allow for any degree of complementarity between consumption and leisure, they do not nest the separable preferences assumed in the benchmark model.

[42] Simulation results for CES preferences are available from the author on request.

The paper also makes a methodological contribution in illustrating a feasible approach for moving beyond calibration to structural estimation for computationally demanding models. This is achieved by applying the pseudo-Bayesian methods in Chernozhukov and Hong (2003) that allow Monte-Carlo Markov Chain techniques to be used in a setting where simulated moment conditions, rather than a likelihood function, form the basis of estimation.

The findings in this paper that are most directly relevant for policy concern the underlying sources of inequality. The quantitative findings indicate that at least as much of the cross-sectional variance in total lifetime inequality is due to shocks realized after entry to the labor market, as is due to factors fixed at the beginning of the career. This suggests that policy interventions targeted at working-age individuals may potentially impact the distribution of welfare in the economy. Of the fixed factors, it was found that initial financial wealth is less important for total lifetime welfare than human wealth in the form of fixed individual skills.

Finally, the issues raised in this paper suggest a number of potentially fruitful areas in which the research agenda for understanding lifecycle inequality should head. First, one could allow for a more realistic model of the labor market, allowing for search frictions and participation decisions along the lines of Low, Meghir and Pistaferri (2006). Second, the scope of the investigation could be expanded to include households with access to female labor supply as a source of additional insurance as in Attanasio, Low and Sanchez-Marcos (2004). Third, allowing for investments in human capital either on- or off-the-job as in Imai and Keane (2004) or Hugget, Ventura and Yaron (2006) could help match the observed patterns in the variance of hours. Fourth, recent work by Cunha, Heckman and Navarro (2005) suggests that individuals may have more information about future wages than is implied by the stochastic process for wages in this paper. Allowing for these sorts of phenomena will impact the pattern of inequality over the lifecycle. Lastly, a particular asset market structure that includes only a risk-free bond has been assumed in this paper. Allowing for alternate trading environments that include risky assets and durable consumption goods may also be important for understanding lifecycle inequality.

A Data

A.1 PSID Data

The PSID sample is drawn from the 1968 to 1997 waves of the Panel Study of Income Dynamics. Table A.1 shows the selection criteria applied, along with the number of individual-year observations lost at each stage. I include an observation in the sample in each year that it satisfies all of the selection criteria. Hence households may drop in and out the panel. The following selection criteria warrant further comment:

- The SEO sample is the Survey of Economic Opportunities, a sub-sample that oversamples poor households. The selection is implemented by dropping observations whose 1968 household identifier is less than 3000.
- I define potential labor market experience as age minus years of education minus six and keep households whose head has between 3 and 38 years of experience and is aged between 20 and 60.
- The single worker criterion is implemented by by choosing households where the male head is either the only worker or the female spouse works but earns less than half the annual earnings of the head.
- I drop observations where the head worked less than 520 or more than 5096 hours in the sample year. The hours criteria correspond to a minimum of 10 hours per week on average over the year and a maximum of of 14 hours per day, 7 days a week in every week of the year.
- The minimum wage criterion is implemented by dropping observations where the average hourly wage was less than half the corresponding legal minimum wage for that year.
- From 1968-1979, annual earnings were top-coded in the PSID at $\$99,999$. This was increased to $\$999,999$ in 1980 and to $\$9,999,999$ in 1981.

	PSID, 1968-1997	**CEX, 1980-2003**
Raw sample	**290,375**	**45,925**
Drop SEO sample	(136,078)	N/A
Heads of households	(64,897)	N/A
Male	(16,939)	N/A
Drop full-time students	(14)	(0)
Age $\in [20, 60]$	(4,980)	(10,290)
Potential experience $\in [3, 38]$	(5,681)	(2,136)
Single-worker household	(19,259)	(18,030)
Annual hours $\in [520, 5096]$	(611)	(255)
1997 or earlier	N/A	(3,913)
Minimum wage criterion	(342)	(60)
Drop top-coded earnings	(60)	N/A
Final sample	**41,514**	**11,241**

In total, the sample comprises 4,629 individuals who are present in between 1 and 30 waves. The distribution of number of observations per individual is shown in Table A.2

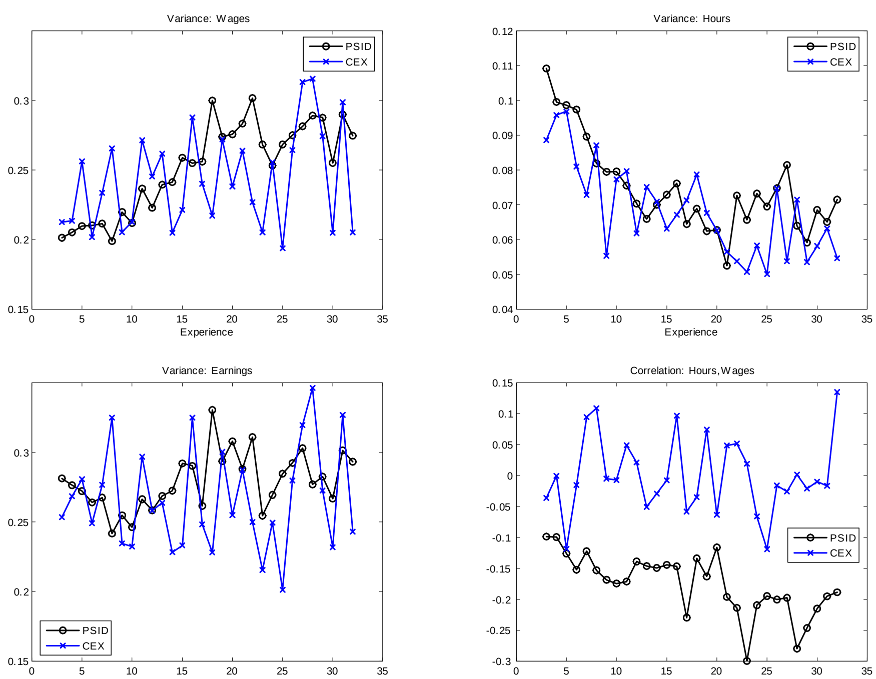
Number of years present	Number of Individuals
1-5	1974
6-10	996
11-15	718
16-20	524
21-25	289
26-30	128
Total	4629

A.2 CEX Data

The CEX sample is drawn from the 1980 to 1997 waves of the Consumer Expenditure Survey. Table A.1 shows the selection criteria applied, along with the number of observations lost at each stage. The raw sample is defined as households which were interviewed for four consecutive quarters. Following Kruger and Perri (2006) I assign a household to a survey year if the fourth interview took place before April of the following year. I define the head as the CEX reference person if he is a male and as the spouse if the the CEX reference person is female and the head is a spouse. This is for consistency with the PSID in which the male is always referred to as the head in male/female couples.

A.3 Comparison of CEX and PSID Samples

Figure A1 shows a comparison of the lifecycle properties of the joint distribution of wages and hours in the two data sets. Both the level and patten in the variance of wages, earnings and hours are remarkably similar in the two samples. The wage-hours correlation is negative in the PSID with a notable decrease, while it is positive in the CEX and relatively flat.



B Construction of Moments

To fix ideas I first define the population moment that we which to estimate. Let $z_{i,t}^1$ and $z_{i,t}^2$ be observations on two variables for indiviudal i with potential labor market experience t. They may be wages, hours, consumption, wealth or dummy variables for positive wealth. We can split the covariance betweeen them at each point in the lifcycle into a component due to observable characteristics, X, and a residual component of unexplained variation.

$$COV\left[z^1, z^2|t\right] = COV_X\left[E\left[z^1|X,t\right], E\left[z^2|X,t\right]|t\right] + E_X\left[COV\left[z^1,z^2|X,t\right]\right] \tag{10}$$

In this paper I am concerned with the properties of the second, residual, component. This can be simplified further as

$$\begin{aligned} E_X\left[COV\left[z^1,z^2|X,t\right]\right] &= E\left[\left(z^1 - E\left[z^1|X,t\right]\right)\left(z^2 - E\left[z^2|X,t\right]\right)|t\right] \\ &= E\left[\varepsilon^{z_1}\varepsilon^{z_2}|t\right] \end{aligned} \tag{11}$$

where $\varepsilon_{i,t}^z = z_{i,t} - E\left[z|X_i,t\right]$. Estimation of this moment procedes in two steps

1. For each z estimate $E\left[z|X,t\right]$ by regressing z on a set of time dummies, time-education interactions (4 education categories), 3 race dummies, a set of experience dummies and experience/year slope interactions. Note that the choice to control for time effects, rather than cohort effects in the first stage is arbitrary. Because I use dummie variables and that are interacted with experience, the partitions created by experience/time controls is exactly the same as those created by experience/cohort controls. I denote the residual from this first stage regression by $\hat{\varepsilon}_{i,t}^z$. For each relevant pair of variables, z_1,z_2, I denote the product $\varepsilon_{i,t}^{z_1}\varepsilon_{i,t}^{z_2}$ by $x_{i,t}$ and its estimate by $\hat{x}_{i,t} = \hat{\varepsilon}_{i,t}^{z_1}\hat{\varepsilon}_{i,t}^{z_2}$.

2. The outer expectation in (11) can then be written as $E\left[x_{i,t}|t\right]$, which is consistently estimated by $\frac{1}{n^{x_t}}\sum_{i=1}^{n^{x_t}}\hat{\varepsilon}_{i,t}^{z_1}\hat{\varepsilon}_{i,t}^{z_2}$. These are the estimates plotted in the figures in the main text.

Confidence intervals are calculated by bootstrap, stratified by experience levels, using 250 replications. These account for any additional estimation error induced by the use of residuals from the first-stage regression.

C Solution of Model

The decision problem of an agent can be written recursively as:

$$V_t\left(\alpha, z_t, \varepsilon_t, k_h\right) = \max_{c_t, h_t, k_{t+1}}\left\{u_t\left(c_t, l_t\right) + \beta s_t E\left[V_{t+1}\left(\alpha, z_{t+1}, \varepsilon_{t+1}, k_{t+1}\right)\right]\right\}$$

subject to the following constraints:

$$c_t + s_t k_{t+1} \leq I_t$$

$$I_t = \begin{cases} We_t h_t + Rk_t & \text{if } t < T^{ret} \\ Wp & \text{if } t \geq T^{ret} \end{cases}$$

$$k_{t+1} \geq 0$$

$$0 \leq h_t \leq 1 \quad \text{if } t < T^{ret}$$

$$\log e_t = \alpha + z_t + \varepsilon_t$$

I compute the optimal decision rules by backward induction using the method of endogenous grid points. I approximate policy functions in capital using piece-wise linear interpolants, with an exponentially spaced grid. All other state variables are assumed discrete. The three components of the wage process are approximated using a discrete-state markov chain, with an age-varying state. Values and transition probabilities are chosen to match the age-varying unconditional variance and dependence structure of each component to that implied by the continuous process. I assume $n_\alpha = n_\varepsilon = n_\varphi = 3$ grid points for the fixed effect, transitory shocks and preference heterogeneity, $n_Z = 7$ grid points for the persistant component of wages and $n_k = 35$ grid points for capital holdings.

Equilibrium distributions are then simulated using the calculated policy rules and exogenous death probabilties. I use $10,000$ simulations. This procedure (calculation of policy rules and simulation) is then repeated with different values of β until the aggregate capital/income ratio is equal to 3.115. Further details and code are available from the author on request.

D Identification of Wage Process

Benchmark Model Let $\omega_{i,t}$ be the estimated log wage residual from the first-stage regression. The benchmark specification is given by

$$\begin{aligned} \omega_{i,t} &= \phi\alpha_i + z_{i,t} + \tau_t\eta_{i,t} \\ z_{i,t} &= \rho z_{i,t-1} + \pi\varepsilon_{i,t} \\ z_{i,0} &= 0 \end{aligned}$$

where α_i, $\eta_{i,t}$ and $\varepsilon_{i,t}$ are all standard normal random variables. Assume that $\omega_{i,t}$ is observed for $t = 1...T$, $T \geq 4$. A sub-set of the elements of the autocovariance matrix are given by

$$\begin{aligned} \sigma_{jj} &= \phi^2 + \tau_j^2 + \frac{1-\rho^{2j}}{1-\rho^2}\pi^2 \\ \sigma_{1j} &= \phi^2 + \rho^{j-1}\pi^2; \qquad j \geq 2 \end{aligned}$$

The following three moment conditions then uniquely identify ρ, π^2 and ϕ^2:

$$\begin{aligned} \frac{\sigma_{14}-\sigma_{13}}{\sigma_{13}-\sigma_{12}} &= \rho \\ \sigma_{13}-\sigma_{12} &= \rho\left(\rho-1\right)\pi^2 \\ \sigma_{12} &= \phi^2 + \rho\pi^2 \end{aligned}$$

Finally, τ_t is identified from σ_{tt}.

Profile Heterogeneity Model The profile heterogeneity model specification is given by

$$\begin{aligned} \omega_{i,t} &= \alpha_i + \beta_i t + z_{i,t} + \tau_t\eta_{i,t} \\ z_{i,t} &= \rho z_{i,t-1} + \pi\varepsilon_{i,t} \\ z_{i,0} &= 0 \end{aligned}$$

where $\eta_{i,t}$ and $\varepsilon_{i,t}$ are standard normal random variables and the distribution of $(\alpha_i, \beta_i)'$ is given by

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} \sim N\left(0, \begin{matrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ & \sigma_\beta^2 \end{matrix}\right)$$

Assume that $\omega_{i,t}$ is observed for $t = 1...T$, $T \geq 4$. A sub-set of the elements of the autocovariance matrix are given by

$$\sigma_{jj} = \sigma_\alpha^2 + 2j\sigma_{\alpha\beta} + j^2\sigma_\beta^2 + \tau_j^2 + \frac{1-\rho^{2j}}{1-\rho^2}\pi^2$$

$$
\begin{aligned}
\sigma_{1j} &= \sigma_\alpha^2 + (1+j)\,\sigma_{\alpha\beta} + j\sigma_\beta^2 + \rho^{j-1}\pi^2 \\
\sigma_{2j} &= \sigma_\alpha^2 + (2+j)\,\sigma_{\alpha\beta} + 2j\sigma_\beta^2 + \rho^{j-2}\left(1+\rho^2\right)\pi^2 \\
\sigma_{kj} &= \sigma_\alpha^2 + (k+j)\,\sigma_{\alpha\beta} + kj\sigma_\beta^2 + \rho^{j-k}\frac{1-\rho^{2k}}{1-\rho^2}\pi^2; \qquad j>k
\end{aligned}
$$

The following two moment conditions identify ρ and π^2

$$
\begin{aligned}
\frac{(\sigma_{15}-\sigma_{14})-(\sigma_{14}-\sigma_{13})}{(\sigma_{14}-\sigma_{13})-(\sigma_{13}-\sigma_{12})} &= \rho \\
(\sigma_{14}-\sigma_{13})-(\sigma_{13}-\sigma_{12}) &= \rho\,(\rho-1)^2\,\pi^2
\end{aligned}
$$

σ_α^2, $\sigma_{\alpha\beta}$ and σ_β^2 are then identified from

$$
\begin{aligned}
(\sigma_{34}-\sigma_{24})-(\sigma_{23}-\sigma_{13}) &= \sigma_\beta^2 + \text{ terms in }\left(\rho,\pi^2\right) \\
\sigma_{23}-\sigma_{13} &= \sigma_{\alpha\beta} + 3\sigma_\beta^2 + \text{ terms in }\left(\rho,\pi^2\right) \\
\sigma_{12} &= \sigma_\alpha^2 + 3\sigma_{\alpha\beta} + 2\sigma_\beta^2 + \rho\pi^2
\end{aligned}
$$

Finally, τ_t is identified from σ_{tt}.

E Details of Structural Estimation Procedure

Define $x_{i,t} = \left(z_{i,t}^1 - E\left[z^1|X_{i,t},t\right]\right)\left(z_{i,t}^2 - E\left[z^2|X_{i,t},t\right]\right)$ for two variable $z^1, z^2 \in \{\log c, \log h, \log w\}$ as in Appendix B. Let $\theta \in \Theta$ denote the $K\times 1$ vector of parameters to be estimated in the structural estimation phase. I assume that Θ is convex and compact. For each pair of variables represented by x, and each experience level, t, define the corresponding cross-sectional second moment predicted by the model at parameter vector θ by $\alpha^{x,t}(\theta)$.

Using the notation and setup of Chernozhukov and Hong (2003) (hereon CH) define the criterion function for the estimation as

$$
\begin{aligned}
L_n(\theta) &= -\frac{1}{2}\sum_{x,t} w^{x,t} g_n^{x,t}(\theta)^2 \\
g_n^{x,t}(\theta) &= \left(\frac{1}{n^{x,t}}\sum_{i=1}^{n^{x,t}} x_{i,t} - \frac{1}{T}\sum_{s=1}^{T}\frac{1}{n^{x,s}}\sum_{i=1}^{n^{x,s}} x_{i,s}\right) - \left(\alpha^{x,t}(\theta) - \frac{1}{T}\sum_{s=1}^{T}\alpha^{x,s}(\theta)\right) \\
&= \frac{1}{n^{x,t}}\sum_{i=1}^{n^{x,t}}\left(\tilde{x}_{i,t} - \tilde{\alpha}^{x,t}(\theta)\right)
\end{aligned}
$$

where $\bar{x}_t = \frac{1}{n^{x,t}}\sum_{i=1}^{n^{x,t}} x_{i,t}$, $\tilde{x}_{i,t} = x_{i,t} - \frac{1}{T}\sum_{s=1}^{T}\bar{x}_s$ and $\tilde{\alpha}^{x,t}(\theta) = \alpha^{x,t}(\theta) - \frac{1}{T}\sum_{s=1}^{T}\alpha^{x,s}(\theta)$ are the deviations from the average moments over the lifecycle in the data and the model repsectively. Note that for simplicity of notation I have restricted the setup to include only diagonal weighting matrices. Note also that two factors complicate the notation: (1) the moments conditions are expressed as deviations from their mean values over the lifecycle; (2) each moment (x,t) is calculated in the data using a different number of observations, $n^{x,t}$.

The GMM Laplace Type Estimator for θ is defined as

$$
\begin{aligned}
\hat{\theta} &= \int_\Theta \theta p_n(\theta)\,d\theta \\
p_n(\theta) &= \frac{e^{L_n(\theta)}}{\int_\Theta e^{L_n(\theta)}d\theta}
\end{aligned}
$$

This decision rule corresponds to a square loss function with respect to the the quasi-posterior distribution from a uniform prior over θ.

E.1 MCMC Simulation for Quasi-Posterior Distribution

CH show that the estimator defined above is asymptotically equivalent to the usual non-linear GMM estimator. Parameter estimates are calculated through Monte-Carlo simulation of the quasi-posterior distribution. I use a Gibbs-Metropolis-Hastings algorithm to generate a sequence of parameter draws whose limiting distribution is $p_n(\theta)$. The parameter estimate is then taken as the sample mean of this simulated distribution. I use a Metropolis-Hastings algorithm with a Gibbs procedure for updating each element of the parameter vector draws sequentially. The jumping distribution for each parameter is a univariate normal distribution, centered on the current state of the markov chain, with a variance that is adjusted to maintain a rejection probability between 50% and 60%. I draw $3000 \times K$ rounds of parameters, where K is the dimension of the parameter vector. The first 2000 are discarded as a "burn-in" phase to ensure that the limiting distribution is reached.

E.2 Confidence Intervals for $\hat{\theta}$

Assumption 2 of CH is satisfied through the choice of penalty function above. I assume that Assumptions 1 of CH (that the true parameter lies in the interior of the parameter space) are satisfied at $\theta = \theta_0 \in \Theta$. I assume that the true parameters uniquely satisfy the moment conditions from the model:

$$E[x_{i,t}] - \alpha^{x,t}(\theta) = 0 \text{ iff } \theta = \theta_0$$

for all x, t. This implies that the identifiability condition, Assumption 3 in CH, is satisfied. I use $\alpha(\theta)$ to denote the $1 \times M$ vector of the individual population moments, $\alpha^{x,t}(\theta)$. Below, wherever I make reference to aymptotic behaviour of any quantities as $n \to \infty$, this (slightly abused) notation refers to $n^{x.t} \to \infty$ for all x, t.

Expanding $L_n(\theta)$ around θ_0 gives

$$\begin{aligned}
L_n(\theta) - L_n(\theta_0) &= (\theta - \theta_0)' \Delta_n(\theta_0) - \frac{1}{2}(\theta - \theta_0)' \left[J_n^*\left(\bar{\theta}\right)\right](\theta - \theta_0) \\
\underset{K\times 1}{\Delta_n(\theta_0)} &= \sum_{x,t} w^{x,t} g_n^{x,t}(\theta_0) \nabla_\theta \tilde{\alpha}^{x,t}(\theta_0) \\
\underset{K\times K}{J_n^*(\theta_0)} &= \nabla_{\theta\theta} L_n(\theta_0)
\end{aligned}$$

where $\bar{\theta} = \tau\theta + (1-\tau)\theta_0$ for some $\tau \in [0,1]$. We now examine the limiting behavior of $\Delta_n(\theta_0)$. By Slustsky's theorem and the central limit theorem for $\bar{x}_{\cdot,t}$ we have for each (x,t)

$$\frac{\sqrt{n^{x,t}} g_n^{x,t}(\theta)}{\sqrt{V[\tilde{x}_{i,t}]}} = \frac{g_n^{x,t}(\theta)}{\sqrt{V\left[\bar{\tilde{x}}_t\right]}} \to N(0,1)$$

where $\bar{\tilde{x}}_t = \frac{1}{n^{x,t}} \sum_{i=1}^{n^{x,t}} \tilde{x}_{i,t} = \frac{1}{n^{x,t}} \sum_{i=1}^{n^{x,t}} \left[x_{i,t} - \frac{1}{T} \sum_{s=1}^{T} \bar{x}_{\cdot,s}\right]$.

To write this in the notation of CH, let $\nabla_\theta \tilde{\alpha}(\theta_0)$ be the $K \times M$ matrix of derivatives of each of the M moment deviations with respect to the K parameters. Let $V\left[\bar{\tilde{x}}\right]$ be the $M \times M$ covariance matrix of the means of the deviations of the data from their average moments over the lifecycle. Let w be the $M \times 1$ vector of the weights. It then follows that

$$\Delta_n(\theta_0) \to N(0, \Omega_n^*(\theta_0))$$

with

$$\underset{K\times K}{\Omega_n^*}(\theta_0) = \nabla_\theta \tilde{\alpha}(\theta_0)\, V\left[w\cdot\bar{\tilde{x}}\right]\nabla_\theta\tilde{\alpha}(\theta_0)'$$

Note that in the notation of CH, we have $\Omega_n^*(\theta) = n\Omega_n(\theta)$ and $J_n^*(\theta) = nJ_n(\theta)$[43].

To construct confidence intervals we apply Theorem 4 in CH and use the variance of the quasi-posterior distribution as an estimate of the inverse of the population hessian, $J_n^*(\theta_0)^{-1}$. Let

$$\underset{K\times K}{\hat{J}_n^*\left(\hat{\theta}\right)^{-1}} = \int_\Theta \left(\theta-\hat{\theta}\right)\left(\theta-\hat{\theta}\right)' p_n(\theta)\,d\theta$$

then by CH, Theorem 4, $\hat{J}_n^*\left(\hat{\theta}\right) J_n^*(\theta_0)^{-1} \xrightarrow{p} I$ and the following are valid confidence intervals for θ:

$$c_n(\alpha) \equiv \hat{\theta} + q_{\frac{\alpha}{2}}\operatorname{diag}\left(\sqrt{\hat{J}_n^*\left(\hat{\theta}\right)^{-1}\Omega_n^*\left(\hat{\theta}\right)\hat{J}_n^*\left(\hat{\theta}\right)^{-1}}\right)$$

where $\hat{\Omega}_n\left(\hat{\theta}\right)$ is such that $\hat{\Omega}_n\left(\hat{\theta}\right)\Omega_n(\theta_0)^{-1} \xrightarrow{p} I$ and q_α is the α-quantile of the standard normal distribution. To calculate these intervals we use the variance-covariance matrix of the limiting distribution of the MCMC sequence together with the following estimate of $\Omega_n^*\left(\hat{\theta}\right)$:

$$\hat{\Omega}_n^*\left(\hat{\theta}\right) = \nabla_\theta\tilde{\alpha}(\theta_0)\,\hat{V}\left[w\cdot\bar{\tilde{x}}\right]\nabla_\theta\tilde{\alpha}(\theta_0)'$$

where $\hat{V}\left[w\cdot\bar{\tilde{x}}\right]$ is an estimate of the $M\times M$ weighted variance-covariance matrix of the moments of the deviations of the data from their life-cycle averages. This matrix contains variances and covariances of means of sample second moments, constructed using residuals from first stage regressions. Analytic expressions for such a matrix are difficult or impossible to derive. To overcome this problem, I estimate this covariance matrix using a stratified bootstrap estimator that accounts for both the additional estimation error induced by the use of residuals from the first stage and the fact that sample second moments refer to deviations rather than raw residuals. 250 bootstrap repetitions are used.

E.3 Simulation-Based Estimation

The above discussion has proceeded under the assumption that $\alpha^{x.t}(\theta)$ is known for all x,t and $\theta\in\Theta$. In reality, the true moments from the model are not known and are instead approximated through simulation as $\hat{\alpha}^{x,t}(\theta) = \frac{1}{R}\sum_{r=1}^{R}\alpha_r^{x,t}(\theta)$, where R is the number of simulations. I assume that $R\to\infty$ at rate that is faster than $\left(n^{x,t}\right)^{\frac{1}{4}}$. This is sufficient to ensure that the simulation-induced error in the moments from the model does not impact the confidence intervals for the parameter estimates.

[43] This is an abuse of notation, but the meaning of this should be obvious to the reader.

References

[1] Abowd, J.M. and D.Card (1989), "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, 57, 411-445.

[2] Altonji, J.G. and J.M. Segal (1996), "Small-sample Bias in GMM Estimation of Covariance Structures", *Journal of Business and Economic Statistics*, 14(3), 353-366.

[3] Aiyagari (1994), "Uninsured Idiosyncratic Risk and Aggregate Savings", *Quarterly Journal of Economics*, 109(3), 659-684.

[4] Attanasio, O., H. Low and V. Sanchez-Marcos (2004), "Explaining Changes in Female Labor Supply in a Lifecycle Model", *Mimeo*, University College London.

[5] Attanasio, O. and N. Pavoni (2006), "Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption", *Mimeo*, University College London.

[6] Baker (1997), "Growth-Rate Heterogeneity and the Covariance Structure of Lifecycle Earnings", *Journal of Labor Economics*, 15(2), 338-375.

[7] Bewley (1986), "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers", in *Contributions to Mathematical Economics in Honor of Gerard Debreu,* ed by W.Hildenbrand, and A. Mas-Colell. North Holland, Amsterdam.

[8] Blundell, R., L. Pistaferri and I. Preston (2006), "Consumption Inequality and Partial Insurance", *Mimeo*, Institute for Fiscal Studies.

[9] Blundell, R. and T. MaCurdy (1999), "Labor Supply: A Review of the Alternative Approaches", in *Handbook of Labor Economics,* ed by O. Ashenfelter, and D. Card, Elsevier Science.

[10] Carroll (2005), "The Method Of Endogenous Grid Points for Solving Dynamic Stochastic Optimization Problems", *Mimeo*, John Hopkins University.

[11] Chernozhukov and Hong (2003) "An MCMC Approach to Classical Estimation", *Journal of Econometrics*, 115, 293-346.

[12] Cunha, Heckman and Navarro (2005), "Separating Uncertainty from Heterogeneity in Lifecycle Earnings", *Oxford Economic Papers*, 57(2), 191-261.

[13] Deaton, A. and C. Paxson (1994), "Intertemporal Choice and Inequality", *Journal of Political Economy*, 102, 437-467.

[14] Diaz-Gimenez, J., V. Quadrini and J.V. Rios-Rull (1997), "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income and Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2), 3-21.

[15] Domeij, D. and J. Heathcote (2004), "On the Distributional Effects of Reducing Capital Taxes", *International Economic Review*, 45(2), 523.

[16] Fernandez-Villaverde, J. and D. Kruger (2005) , "Consumption and Saving over the Lifecycle: How Important are Consumer Durables?", *Mimeo*, University of Pennsylvania.

[17] Floden, M. (2006), "Labor Supply and Saving Under Uncertainty", *Economic Journal*, 116, 721-737.

[18] Gourinchas, P. and J.A. Parker (2000), "Consumption Over the Lifecycle", *Econometrica*, 70(1), 47.

[19] Guvenen (2006), "Learnings Your Earning: Are Labor Income Shocks Really Very Persistent?", *American Economic Review*, Forthcoming.

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

[21] Heathcote, J., K. Storesletten and G.L.Violante (2004), "The Macroeconomic Implications of Rising Wage Inequality in the U.S.", *Mimeo*, New York University.

[22] Heathcote, J., K. Storesletten and G.L.Violante (2005), "Two Views of Inequality Over the Life Cycle", *Journal of the European Eonomic Association*, 3(2-3), 765-775.

[23] Heathcote, J., K. Storesletten and G.L.Violante (2006), "Consumption and Labor Supply with Partial Insurance: An Analytical Framework", *Mimeo*, New York University.

[24] Huggett (1993), "The Risk Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies", *Journal of Economic Dynamics and Control*, 17, 953-969.

[25] Huggett (1996), "Wealth Distribution in Life-cycle Economies", *Journal of Monetary Economics*, 38(3), 469-494.

[26] Huggett, M., G. Ventura and A. Yaron (2006), "Sources of Lifetime Inequality", *Mimeo*, Georgetown University.

[27] Imai and Keane (2004), "Intertemporal Labor Supply and Human Capital Accumulation", *International Economic Review*, 45(2), 601.

[28] Kochanek, K., S. Murphy, R. Anderson and C. Scott (2004), "Deaths: Final Data for 2002", *Nattional Vital Statistics Reports*, 53(5).

[29] Kruger and Fernandez-Villaverde (2005), "Consumption and Savings over the Life-Cycle: How Important are Consumer Durables?", *Mimeo*, University of Pennsylvania.

[30] Kruger, D. and F. Perri (2006), "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory", *Review of Economic Studies*, 73(1), 191-261.

[31] Low, H (2005), "Self-Insurance in a Life-Cycle Model of Labor Supply and Savings", *Review of Economic Dynamics*, 8(4), 945-975.

[32] Low, M., C. Meghir and L. Pistaferri (2006), "Wage Risk and Employment Risk Over the Lifecycle", *IFS Working Paper*, WP06/27.

[33] Marcet, A., F.Obiols-Homs and P.Weil, "Incomplete Markets (2002), Labor Supply and Capital Accumulation", *UPF Economics and Business Working Paper*, No. 659.

[34] Pakes, A. and D. Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators", *Econometrica*, 57(5), 1027-1057.

[35] Pijoan-Mas, J (2006), "Precautionary Savings or Working Longer Hours", *Review of Economic Dynamics*, 9(2), 326-352.

[36] Quadrini, V. and J.V. Rios-Rull (1997), "Understanding the U.S. Distribution of Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2), 22-36.

[37] Rodriguez, S. J. Diaz-Gimenez, V. Quadrini and J.V. Rios-Rull (2002), "Dimensions of Inequality: Facts on the U.S. Distributions of Earnings, Income and Wealth", *Federal Reserve Bank of Minneapolis Quarterly Review*, 26(3), 2-35.

[38] Slesnick, D. and Ulker (2004), "Inequality and the Life-Cycle: Age, Cohort Effects and Consumption", *Mimeo*, University of Texas - Austin.

[39] Storesletten, Telmer and Yaron (2001), "How Important Are Idiosyncratic Shocks? Evidence from Labor Supply", *American Economic Review Papers and Proceedings*, 91(2), 413-417.

[40] Storesletten, Telmer and Yaron (2004), "Consumption and Risk Sharing Over the Lifecycle", *Journal of Monetary Economics*, 51, 609-633.

Table 1: Exogenously Fixed Parameters

Parameter	Description	Value
θ	capital share in production	0.33
δ	depreciation rate	0.06
R	after rax risk-free rate	4.5%
τ_L	labor tax rate	40%
τ_K	capital rax rate	27%
$\frac{K}{Y}$	aggregate capital/income ratio	3.115
$\frac{K^{young}}{E^{young}}$	wealth/earnings ratio of young generation	0.888

Table 2: Parameter Estimates From Wage Model

Parameter	Benchmark	Profile Heterogeneity	No Age Effects in Trans Variance
ρ	0.946	0.937	0.948
	(0.010)	(0.024)	(0.011)
σ^2_η	0.019	0.020	0.018
	(0.002)	(0.004)	(0.003)
σ^2_α	0.056	0.057	0.059
	(0.007)	(0.014)	(0.006)
σ^2_ε	0.072^+	0.069^+	0.072
			(0.014)
σ^2_β	n/a	$1.35E-5$	n/a
		$(8.77E-5)$	
$\sigma_{\alpha\beta}$	n/a	$6.44E-5$	n/a
		$(7.29E-4)$	

Notes: Standard errors are computed using a block bootstrap with 250 replications.

$^+$: For models with age effects in the transitory variance, σ^2_ε is the average transitory variance over the lifecycle.

Table 3: Structural Parameter Estimates

Parameter	Benchmark	Alt. Weight Matrix	With Borrowing	No Age Effects Trans. Shocks	Cobb-Douglas Preferences
	(1)	(2)	(3)	(4)	(5)
γ	3.90	1.66	2.92	4.04	39.49
	(0.17)	(0.02)	(0.05)	(0.17)	(8.55)
σ	9.20	5.55	7.43	8.87	n/a
	(0.10)	(0.01)	(0.03)	(0.20)	
$E(\varphi)$	0.14	0.13	0.12	0.20	1.21
	(0.02)	(0.00)	(0.00)	(0.01)	(3.13)
$V(\log\varphi)$	0.036	0.012	0.020	0.036	0.282
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Gini(k_0)	0.80	0.80	0.78	0.80	0.76
β	1.04	1.00	1.06	1.05	0.82

Notes: Standard errors in parenthesis, calculated as described in Appendix E. Standard errors are not reported for wealth gini since the estimated parameter is the the variance of log wealth Estimated values for log wealth, along with standard errors are available from the author on request. β is calibrated so that the ratio of aggregate domesitc capital to aggregate income is 3.115.

Table 4: Age Component of Cross-Sectional Variance

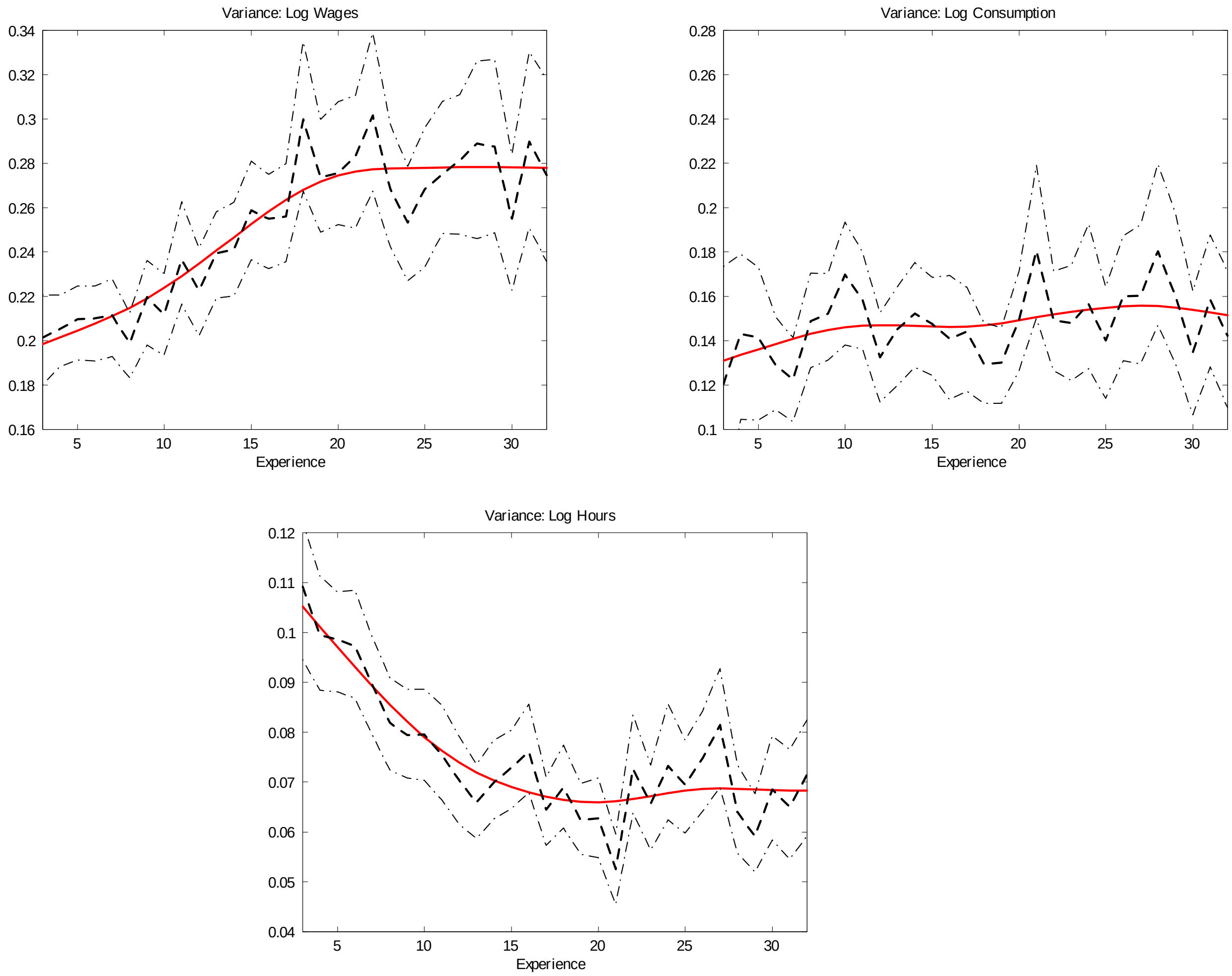
	Between Age $V[E(X \mid t)]$	Within Age $E[V(X \mid t)]$
Log Consumption	43%	57%
Log Hours	46%	54%
Wealth	33%	67%

Table 5: Overall Variance Decomposition

	Preference Heterogeneity	Fixed Effect	Initial Wealth	Wage Shocks
Log Consumption	0%	37%	5%	56%
Log Hours	2%	25%	20%	51%
Wealth	0%	7%	39%	53%
PV Utility		39%	1%	59%

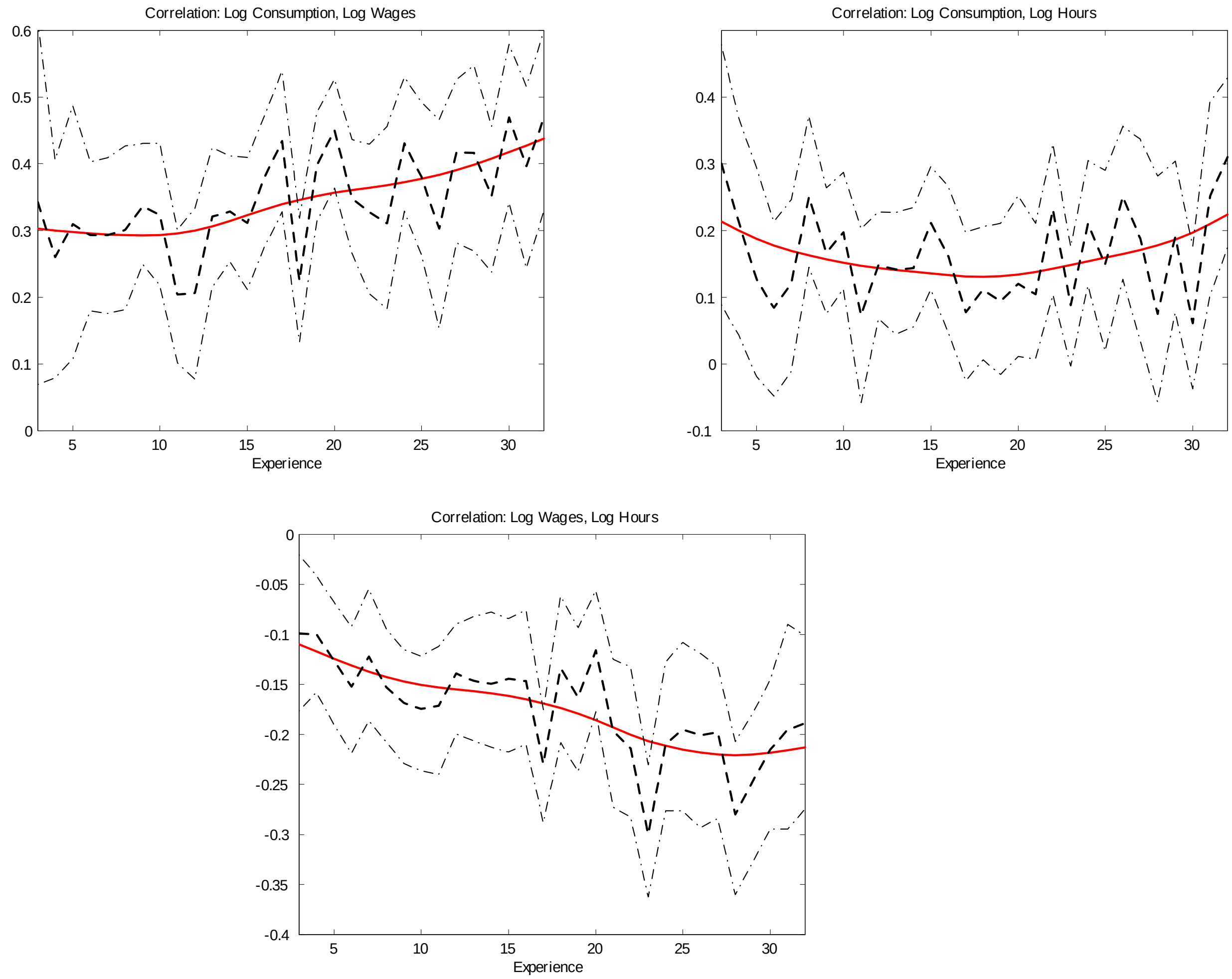
Notes: PV Utility refers to the cross-sectional variance of discounted realized utility for an individual with the average relative taste for leisure in the economy.

Figure 1: Residual Inequality in Wages, Consumption and Hours Over the Lifecycle



Figures show the cross-sectional variance of log wages, log consumption and log hours as a function of potential labor market experience. Thick black dashed lines are point estimates, calculated as described in Appendix B. Thin black dashed lines are 95% bootstrap confidence intervals. Solid red line is a smoothed version of the point estimates, calculated using an HP filter with smoothing parameter equal to 100. Figures for consumption and wages are shown on same scale. Inequality in hours is graphed using a different scale.

Figure 2: Joint Moments of Wages, Consumption and Hours Over the Lifecycle



Figures show the pairwise correlations between log wages, log consumption and log hours as a function of potential labor market experience. Thick black dashed lines are point estimates, calculated as described in Appendix B. Thin black dashed lines are 95% bootstrap confidence intervals. Solid red line is a smoothed version of the point estimates, calculated using an HP filter with smoothing parameter equal to 100. Figures for correlation of consumption with hours and wages are shown on same scale. The correlation between wages and hours is graphed using a different scale.

Figure 3: Alternate Definitions of Consumption

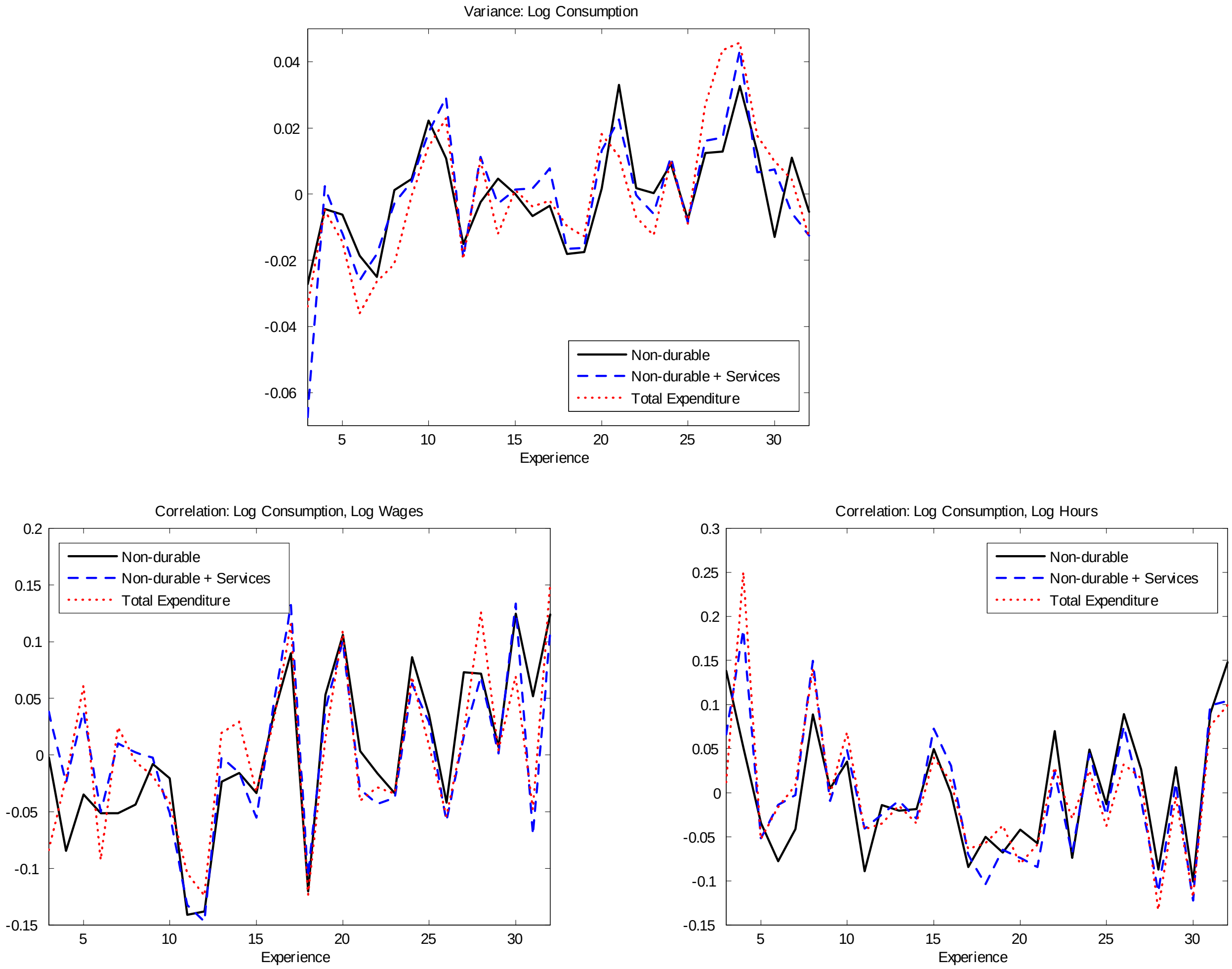


Figure 4: Distribution of Wealth Over the Lifecycle

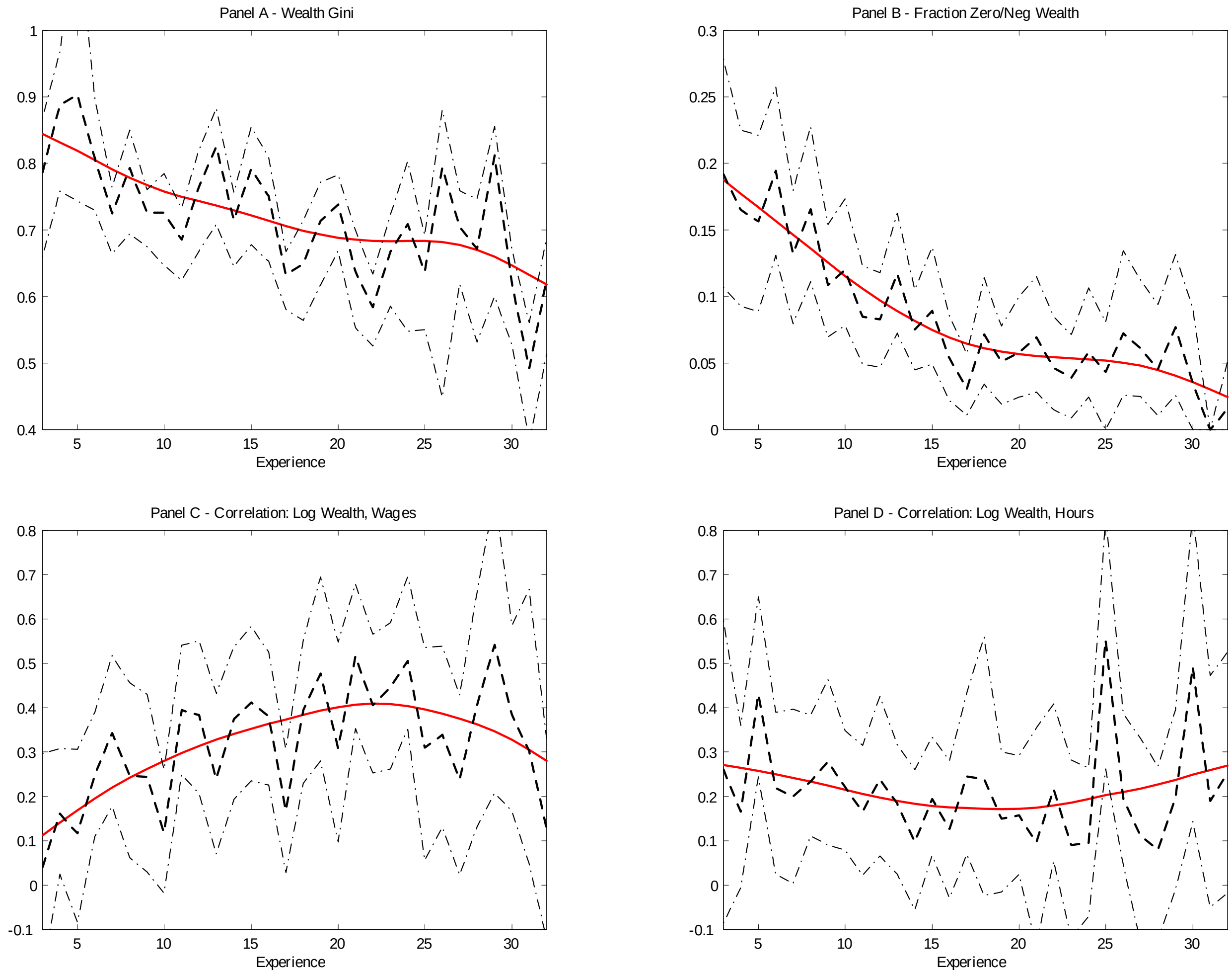


Figure 5: Results from Estimation of Wage Process

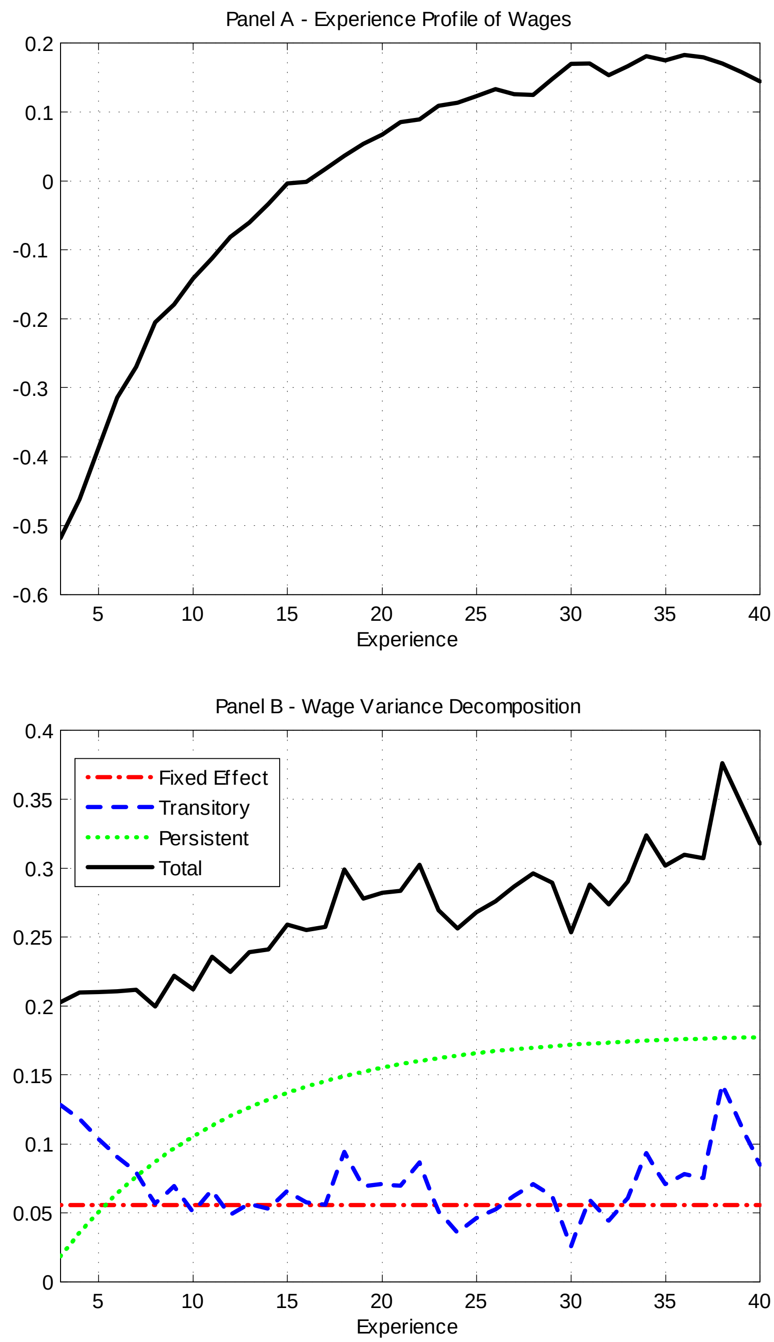


Figure 6: Benchmark Model Fit

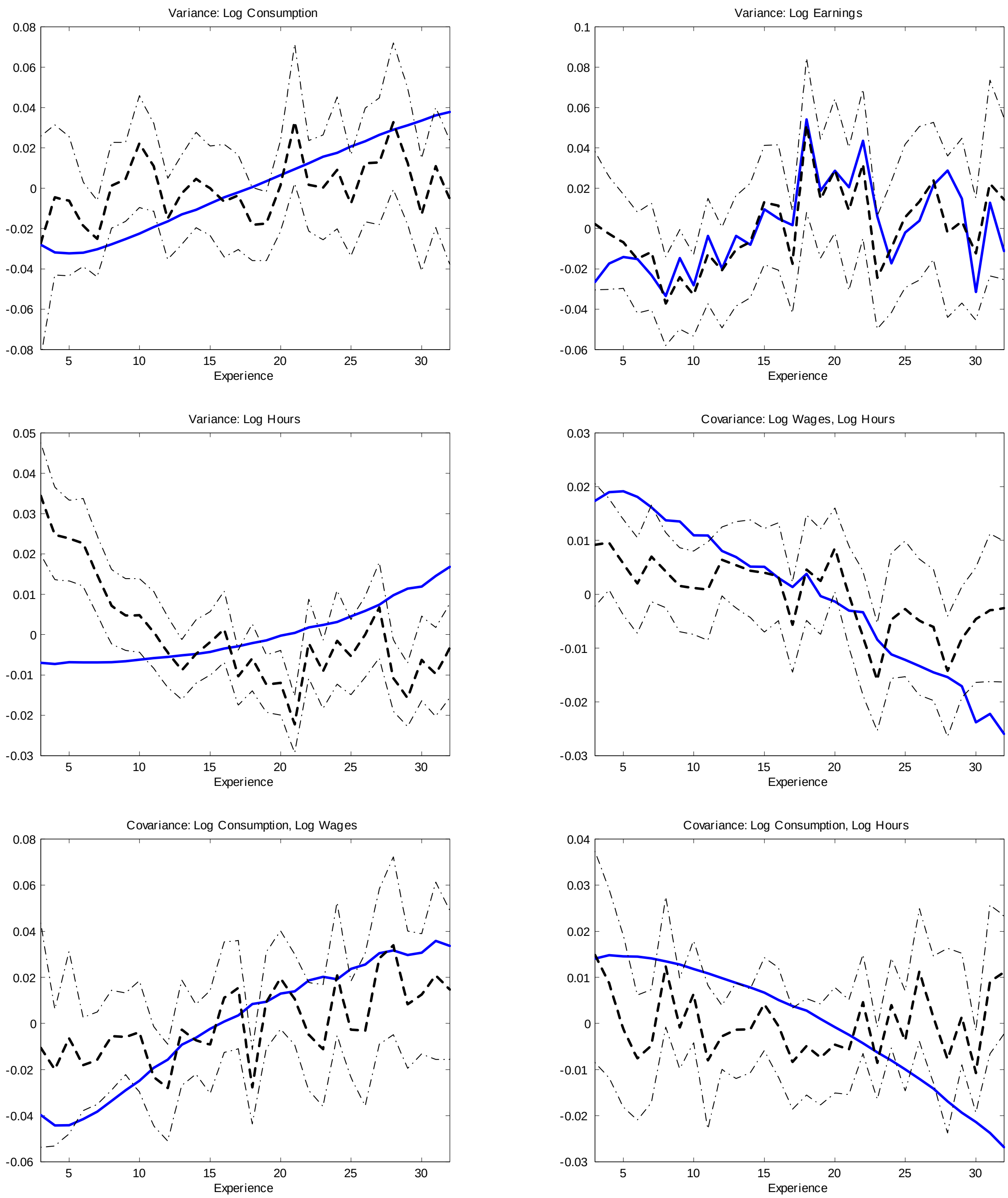


Figure 7: Benchmark Model Fit - Wealth Distribution

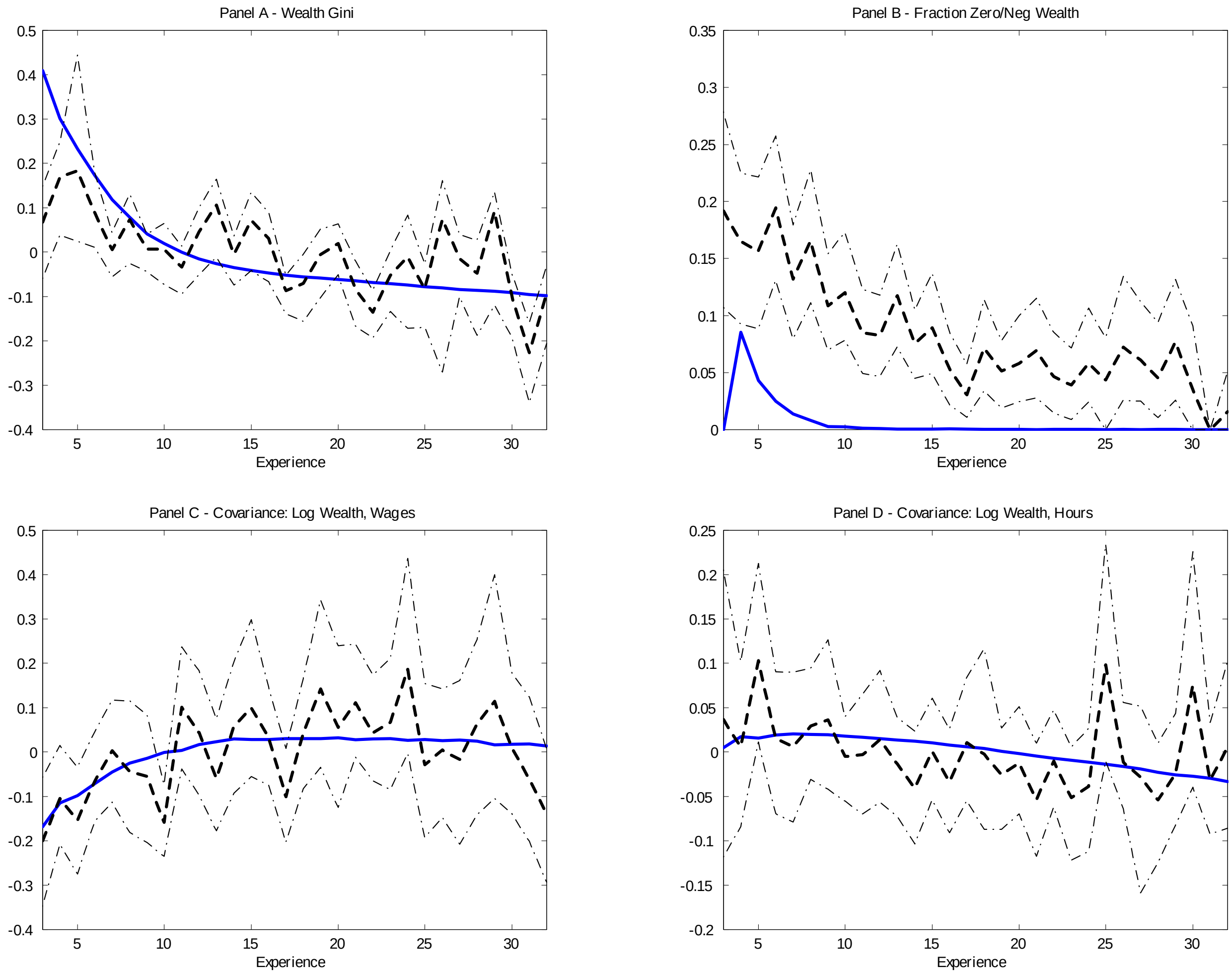


Figure 8: Variance Decomposition - Benchmark Model

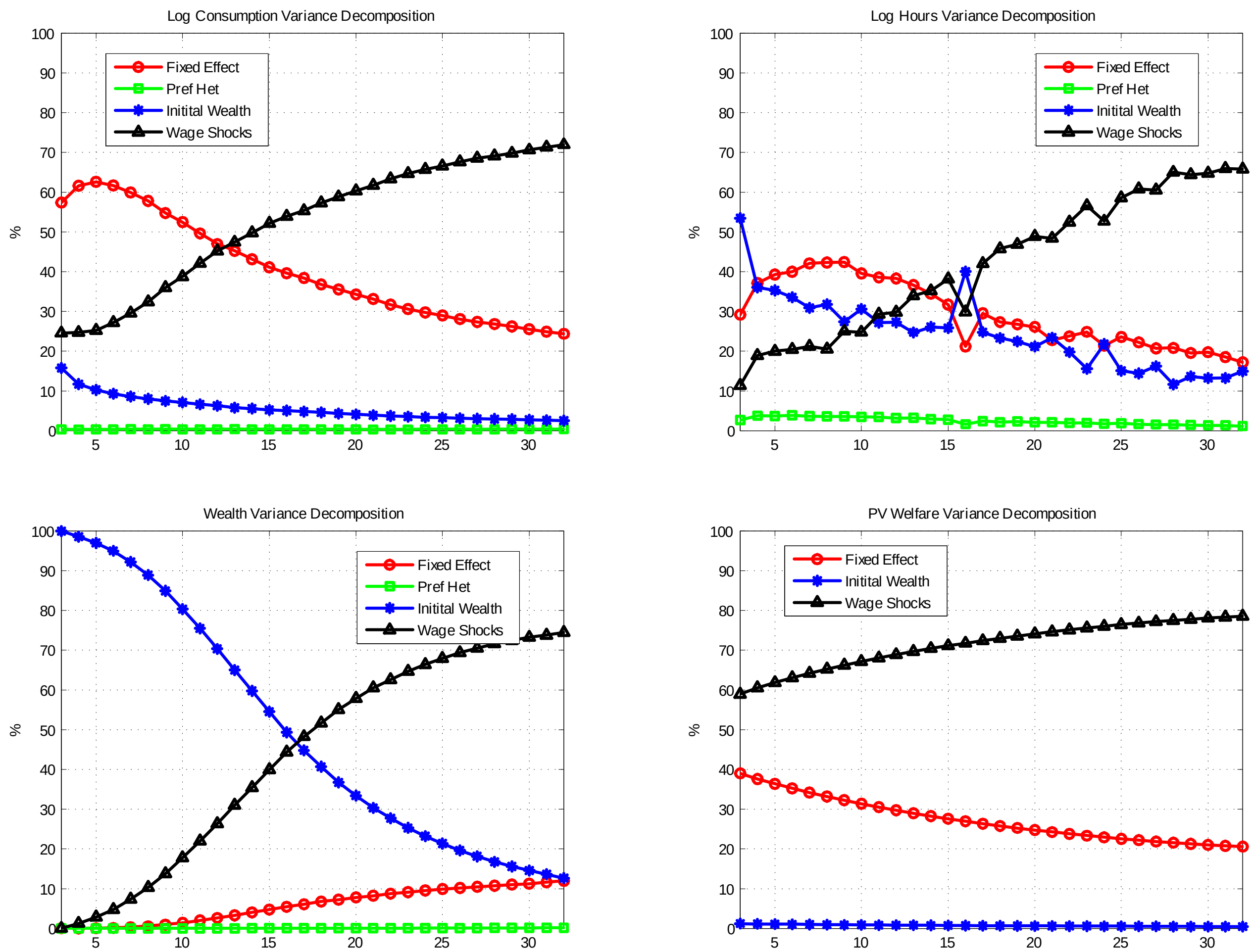


Figure 9: Robustness Exercise - Optimal Weighting Matrix

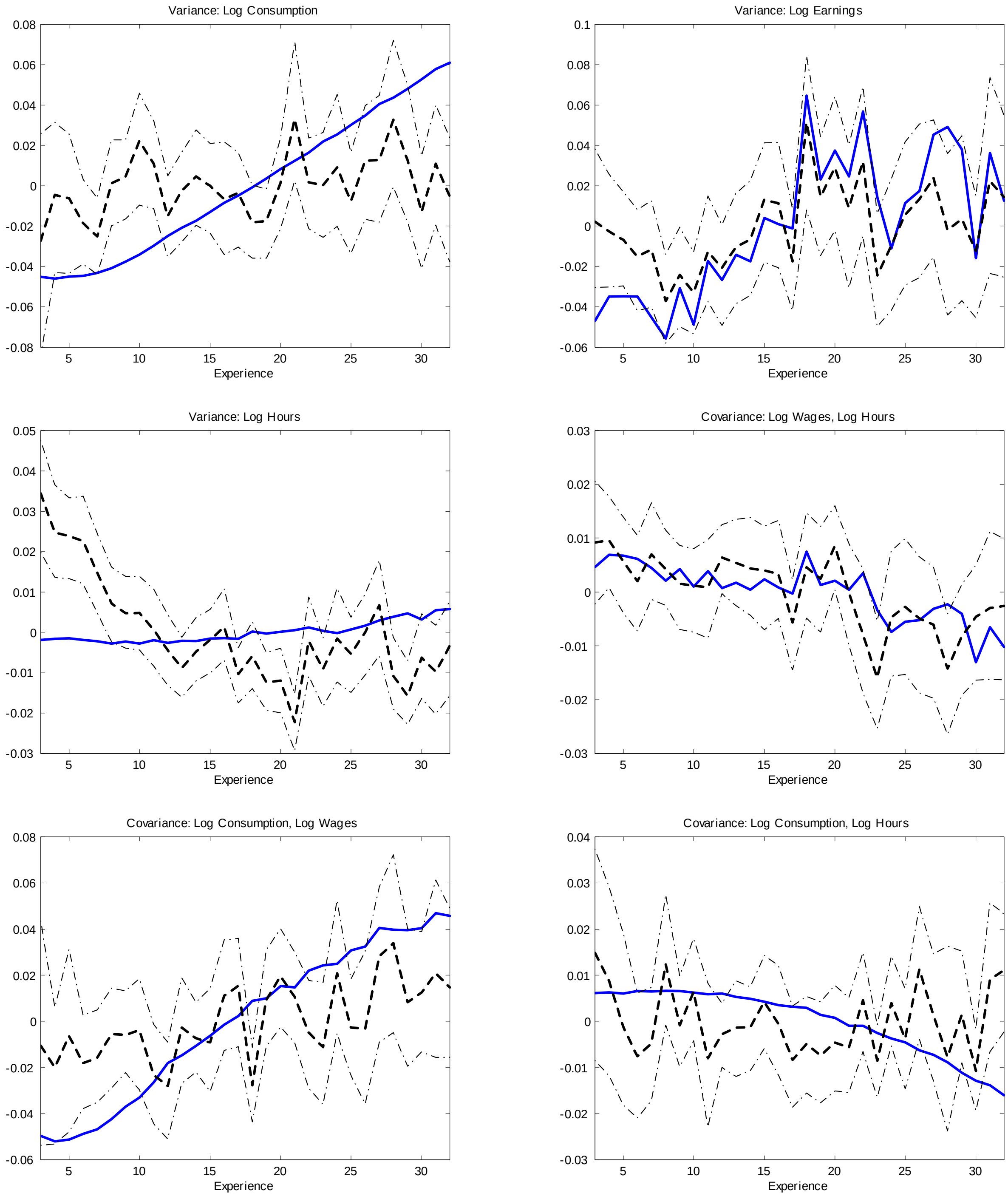


Figure 10: Robustness Exercise - Natural Borrowing Limits

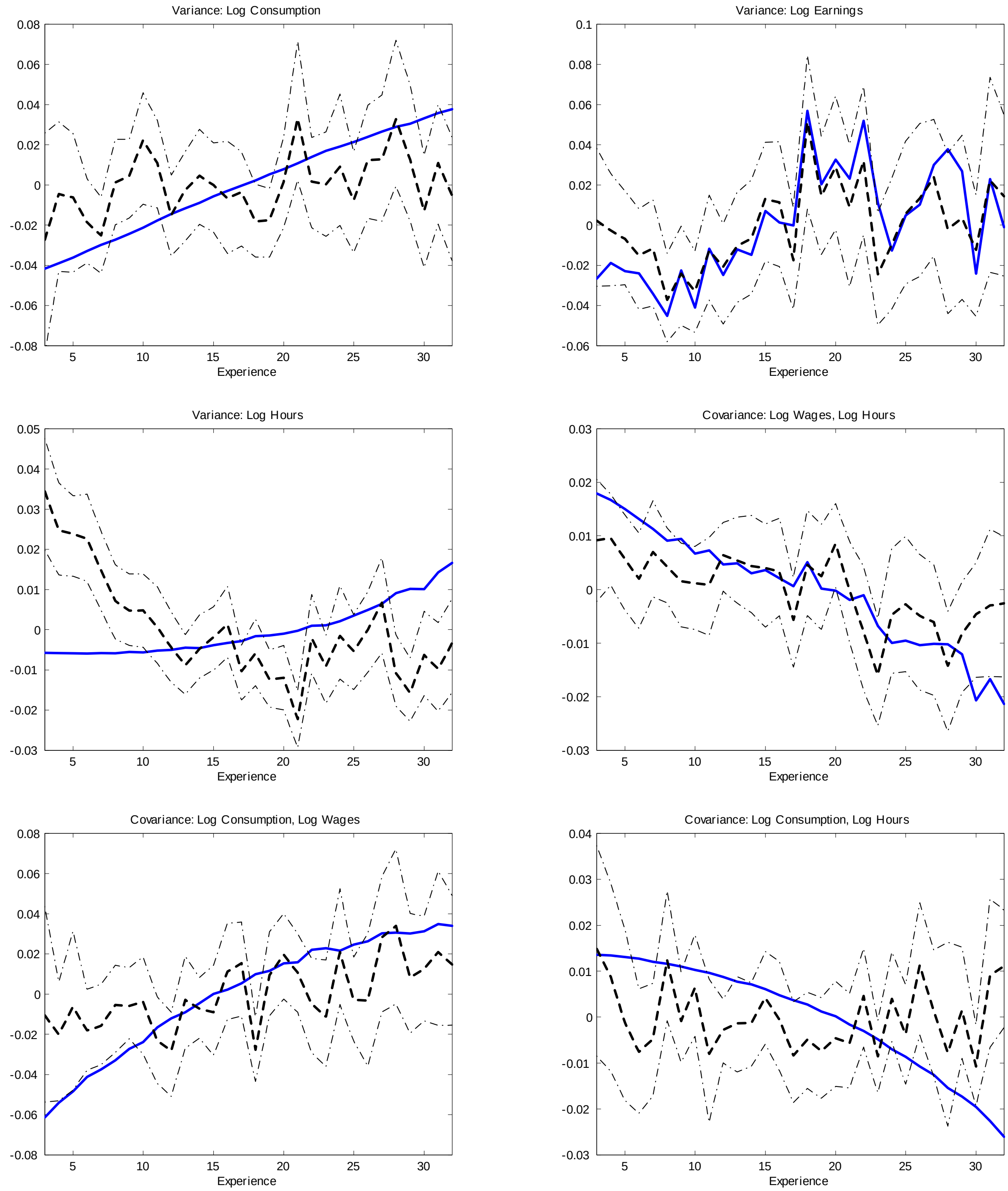


Figure 11: Robustness Exercise - High Risk Aversion

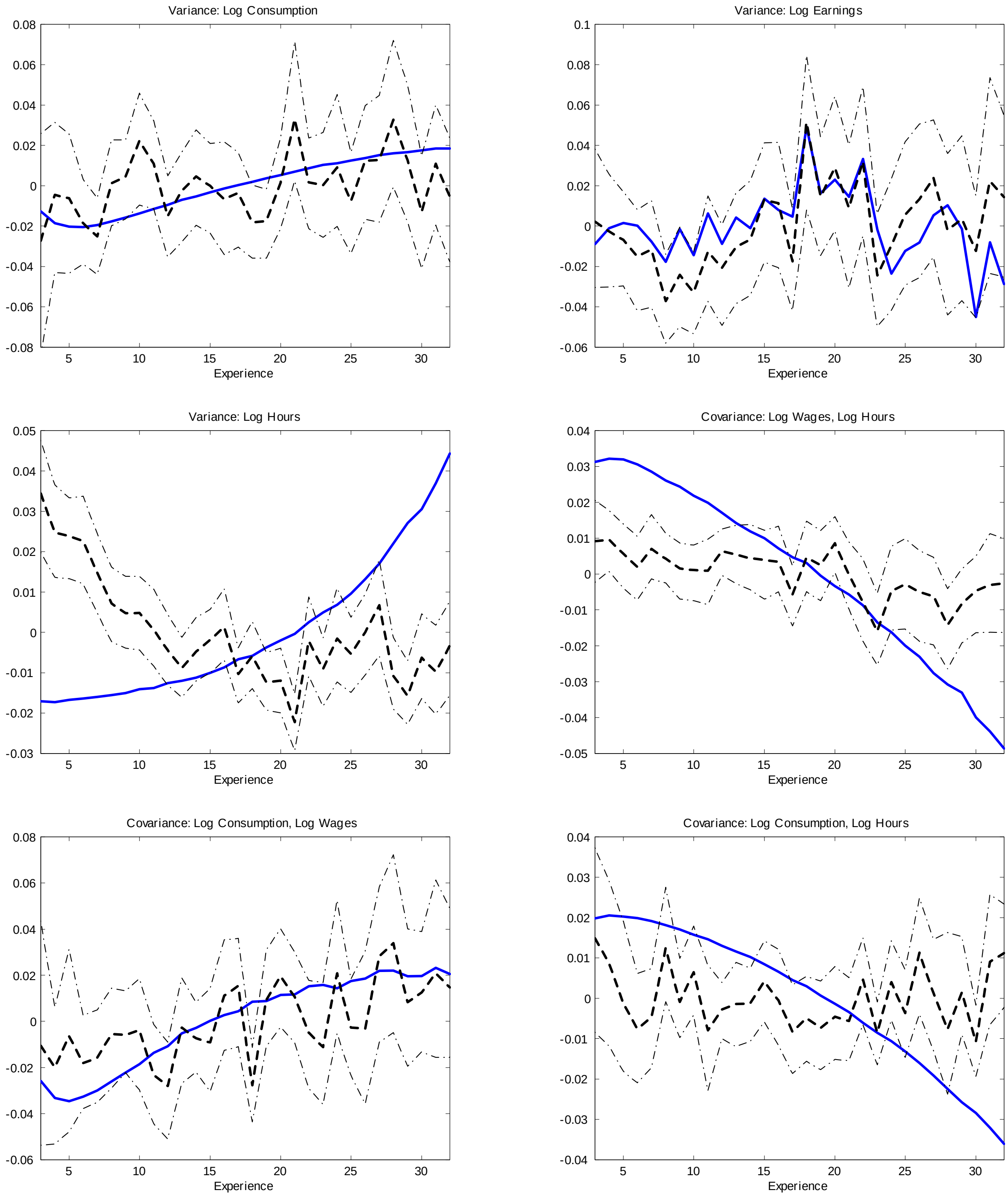


Figure 12: Robustness Exercise - High Initial Wealth

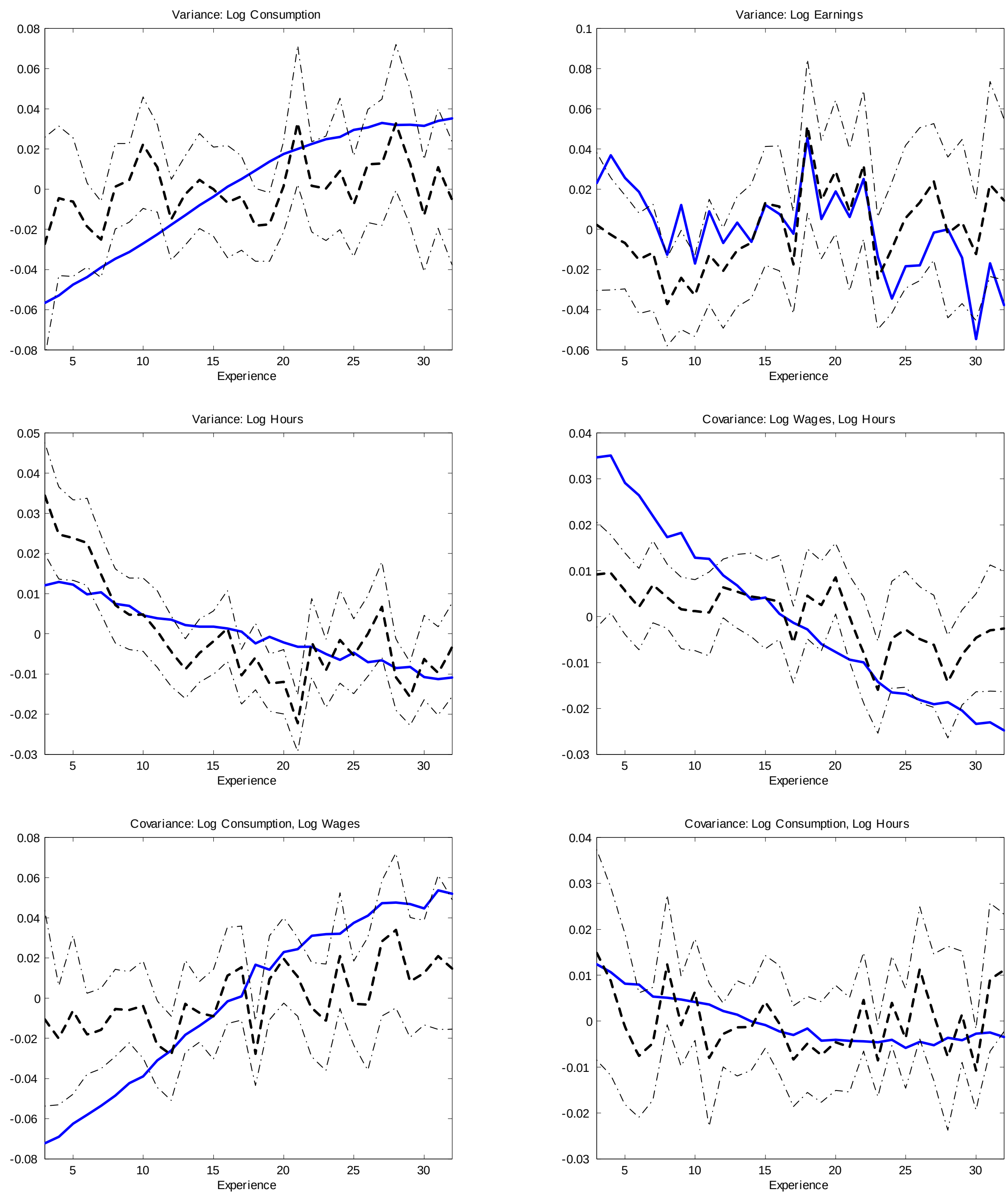


Figure 13: Robustness Exercise - No Age Effects in Transitory Variance

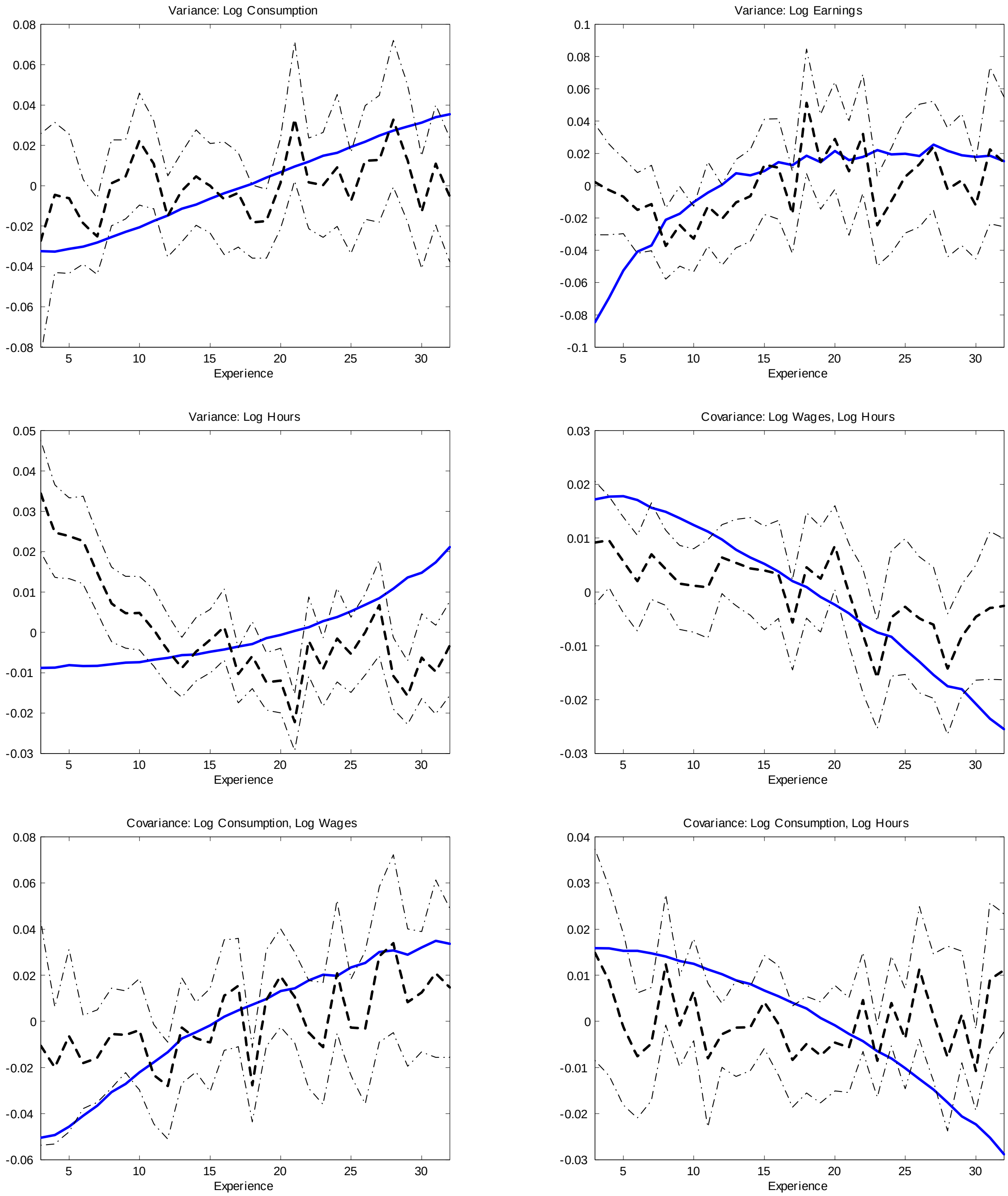


Figure 14: Robustness Exercise - Cobb-Douglas Preferences