Interpreting Life-Cycle Inequality Patterns as an Efficient Allocation: Mission Impossible?

Alejandro Badel[*] Mark Huggett[†]

March 28, 2012

Abstract

The life-cycle patterns of consumption, wage and hours inequality observed in U.S. data are commonly viewed as incompatible with a Pareto efficient allocation. We ask whether these patterns are consistent with Pareto efficiency within a model with preference shocks, wage shocks and full information. Our answer is yes, under two conditions. First, the Frisch elasticity of labor supply has to be below 1 to account for the relatively flat variance profile of log hours and the relatively flat covariance profile of log wages and log hours. Second, preference shocks that impact the marginal utility of consumption must have an increasing variance with age and an increasing covariance with wages with age.

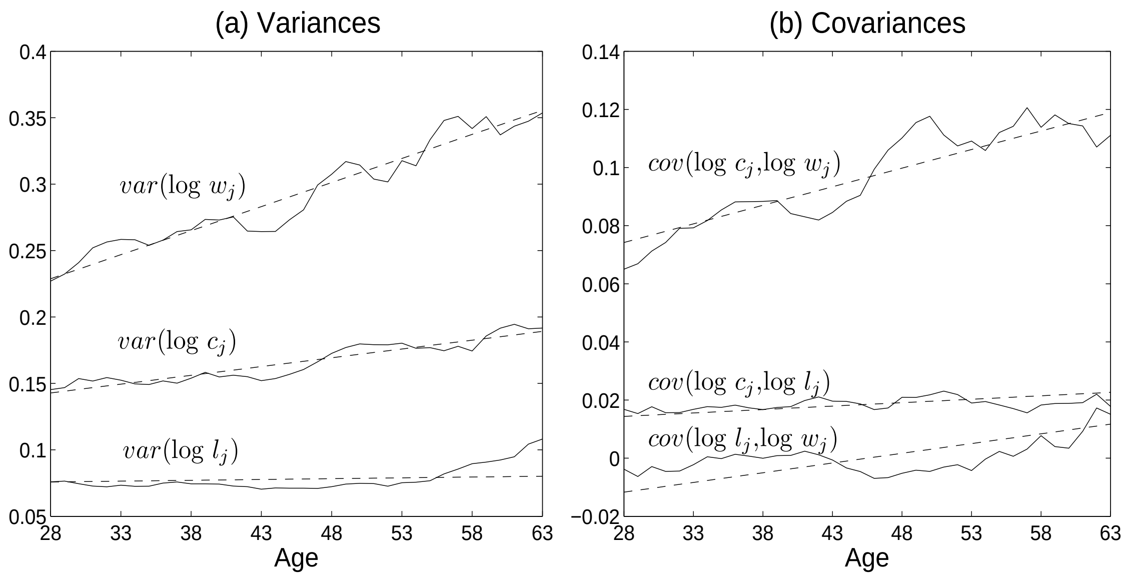
JEL Classification: E21, D91, D52
Keywords: Life Cycle, Inequality, Efficient Allocation, Preference Shocks

[*]Federal Reserve Bank of St. Louis. email: badel@stls.frb.org.
[†]Georgetown University. email: mh5@georgetown.edu.

1 Introduction

The solid lines in Figure 1 display the life-cycle inequality patterns found in micro data from the U.S. Consumer Expenditure Survey (CEX).[1] Figure 1 is consistent with the view that the variance of log consumption and the variance of log wages increase with age while the variance of log hours is relatively flat over much of the life cycle.[2]

Figure 1: Life-Cycle Inequality Patterns



Source: Author's calculations based on CEX data 1981-2003.

A commonly-held view is that a plausible explanation for the cross-sectional variance profiles in Figure 1(a) will involve the introduction of some friction so that some portion of the idiosyncratic shocks that impact wages is transmitted to consumption. The literature has examined the following frictions: incomplete markets, limited commitment and private information.[3]

[1]All empirical profiles in Figure 1 are estimated controlling for time effects. Data, sample selection criteria, and details on statistical methodology are described in the Data Appendix.

[2]We focus on the slope of covariances by age as opposed to the slope of correlations in order to avoid attenuation bias.

[3]See Huggett (1996), Storesletten, Telmer and Yaron (2004), Guvenen (2007), Heathcote, Storesletten and Violante (2009), Huggett, Ventura and Yaron (2011) and Kaplan (2011) for work

The literature contains almost no attempt to challenge or refine this commonly-held view. Storesletten, Telmer and Yaron (2001) is, to our knowledge, the only exception. They ask if an efficient allocation in a model where all agents have the same non-separable preferences between consumption and leisure can generate profiles like those in Figure 1(a), given the increasing pattern of wage dispersion. Such models are sometimes called full-insurance or full-information models as no friction beyond resource feasibility inhibits risk sharing. Nonseparability might help to produce the variance patterns in Figure 1(a) as higher hours of work may increase the marginal utility of consumption leading those with high wages to have high hours and high consumption. However, they find that there are no utility function parameters for which their model produces the observed rise in consumption dispersion with only a "small" rise in hours dispersion.[4]

The work of Heathcote et al. (2009) and Kaplan (2011), based on the incomplete-markets friction, has gone the furthest in offering a quantitative account of the cross-sectional facts in Figure 1. Their models feature exogenous random elements: agents are hit by idiosyncratic shocks to wages (i.e productivity) and to preferences. The preference shocks in Heathcote et. al. (2009) follow a random walk and, thus, display increasing dispersion as a cohort ages. The preference shocks in Kaplan (2011) are age invariant and, thus, do not display increasing dispersion with age.

The fact that two leading explanations of the cross-sectional facts in Figure 1 feature wage and preference shocks leads us to ask the following question: are the patterns in Figure 1 inconsistent with a model of Pareto efficient allocations under full information when idiosyncratic wage (i.e productivity) and preference shocks are allowed? Our answer is displayed in Figure 1, where the dotted lines in Figure 1 are produced by our model. Visually, the dotted lines capture much of the age patterns in the data. Thus, we conclude that the basic patterns in U.S. cross-sectional inequality data displayed in Figure 1 do not rule out models of Pareto efficient allocations with full information when such models allow both productivity and preference shocks.

How does our model, with additively separable preferences between consumption and labor, account for the variance patterns in Figure 1? First, an agent works more the larger the agent's idiosyncratic wage shock. Thus, the variance of log hours and log wages both rise with age. A low Frisch elasticity of labor accounts for the relatively flat hours dispersion profile. Second, an agent consumes more the larger the agent's idiosyncratic preference shifter (i.e. the preference shock). The rise in the variance of the preference shifter by age accounts for the rise in consumption

based on incomplete markets. Krueger and Perri (2006) analyze limited commitment - a situation where an agent can walk away from the terms of a contract. Ales and Maziero (2009) analyze efficient allocations when idiosyncratic shocks are privately observed.

[4]They define a "small" rise in the variance of log hours over the life cycle to be .08 or less.

dispersion. These features account for the patterns in Figure 1(a).

How does our model account for the covariance patterns in Figure 1(b)? First, the age trend in the covariance between log wages and log hours is positive in the data. Due to additive separability, the model implies that this covariance must rise with age, given that the wage variance rises with age. The rise in the model covariance is determined by the product of the Frisch elasticity and the rise in the variance of wages. Second, in the data the covariance between log consumption and log wages rises with age, whereas the age trend for the covariance between log consumption and log hours is fairly flat. To account for these facts, the model requires that the covariance between preference shocks and wages rises with age. In addition, the model also requires that the parameter governing the Frisch elasticity must be well below 1 so that the rise over the life cycle in the covariance between consumption and labor hours is fairly flat as it is in U.S. data.

The paper is organized as follows. Section 2 contains the model and a theorem which describes the nature of the inequality patterns in the model. Section 3 describes the quantitative implications of the model for the data patterns in Figure 1. Section 4 concludes.

2 Framework

We analyze an overlapping generations economy. A continuum of agents is born at each time t. The size of each birth cohort is denoted by N_t. Agents are characterized by their age j, their year of birth b, and their own shock history $s^j = (s_0, s_1, ..., s_j)$. At any age $j = 0, 1, ..., J$ there are a finite number of possible shock histories s^j for the agent that occur with probability $P(s^j)$. An agent's productivity $w(s^j) > 0$ and preference shifter $z(s^j)$ at age j are determined by their shock history.

Agents care about expected utility derived from consumption and labor. The functions $c_b(s^j)$ and $l_b(s^j)$ denote age j consumption and labor in history s^j for an agent born in year b. Expected utility is additively separable, where u is a period utility function, β is a discount factor and φ_j is the probability of surviving up to age j:

$$U(c_b, l_b) = E\left[\sum_{j=0}^{J} \varphi_j \beta^j u(c_b(s^j), l_b(s^j), z(s^j))\right]$$

2.1 Planning Problem

At time $t = 1$ the planning objective is to maximize the weighted sum of individual expected utilities. The objective includes all cohorts born in $t = 1, 2, 3...$ and the cohorts born before $t = 1$ which have members alive at $t = 1$. The objective is as follows, where $\gamma_b > 0$ is a planning weight assigned to agents from the cohort born at time b:

$$\sum_{b=-[J-1]}^{\infty} \gamma_b N_b U(c_b, l_b)$$

The Planning Problem is to maximize this objective subject to a resource constraint. This problem is stated below, where we restate the objective by using the fact that time t, birth year b and the age j satisfy $b = t - j$. We also drop terms related to consumption and labor choices occurring before $t = 1$, as they are assumed out of the planner's control at time $t = 1$. The resource constraint says that total consumption equals total output at each time period.

$$\textit{Problem P1} : \max \sum_{t=1}^{\infty} E\left[\sum_{j=0}^{J} \gamma_{t-j} N_{t-j} \varphi_j \beta^j u(c_{t-j}(s^j), l_{t-j}(s^j), z(s^j))\right]$$

subject to

$$\sum_{j=0}^{J} E\left[c_{t-j}(s^j) - l_{t-j}(s^j) w(s^j)\right] N_{t-j} \varphi_j = 0, \forall t \geq 1$$

We make the following assumptions:

A1: The period utility function u is additively separable, strictly concave, continuously differentiable. Furthermore, u is strictly increasing in consumption and decreasing in labor and satisfies the Inada conditions so that the range of the marginal utilities of consumption and labor are $(0, \infty]$ and $[0, -\infty)$ respectively.

A2: $N_t = (1 + n)^t$ and $\beta(1 + n) < 1$.

A3: The Planning weights are set to $\gamma_b = \beta^b, \forall b$.

We rewrite the Planner's objective below, making use of assumptions A2 and A3. This highlights the fact that the Planner faces effectively a sequence of static maximization problems with the same period objective function and the same resource constraint.

$$\sum_{t=1}^{\infty} [\beta(1 + n)]^t E\left[\sum_{j=0}^{J} \frac{\varphi_j}{(1 + n)^j} u(c_{t-j}(s^j), l_{t-j}(s^j), z(s^j))\right]$$

Theorem 1: Assume A1 - A3. Then (i) and (ii) below hold.

(i) *There exists a unique allocation* $(c^*(s^j), l^*(s^j))$ *that solves the problem of maximizing* $E\left[\sum_{j=0}^{J} \frac{\varphi_j}{(1+n)^j} u(c(s^j), l(s^j), z(s^j))\right]$ *subject to the period resource constraint* $\sum_{j=0}^{J} E\left[c(s^j) - l(s^j)w(s^j)\right] \frac{\varphi_j}{(1+n)^j} = 0$.

(ii) *The time-invariant allocation* $(c^*(s^j), l^*(s^j))$ *is the unique solution to Problem P1.*

Proof: See the Appendix.

2.2 Inequality Implications of the Model

Theorem 2 presents the inequality implications of the model. Theorem 2 says that the variances and covariances that were analyzed in Figure 1 rise linearly with age in the model economies. These implications are based on the additively separable, isoelastic utility function with multiplicative preference shifters stated in Assumption $A1'$. Similar specifications are widely used in the empirical literature.

A1′: The period utility function is $u(c, l, z) = \exp(z)\frac{c^{1-\rho}}{1-\rho} - \frac{l^{1+\phi}}{1+\phi}$.

Theorem 2: Assume A1′, A2 and A3. Furthermore, assume that the variances and covariance of log wages and the preference shifter rise linearly with age according to the following variance-covariance equations

$$\begin{aligned} var(\log w(s^j)) &= g_w + v_w j \\ var(z(s^j)) &= g_z + v_z j \\ cov(\log w(s^j), z(s^j)) &= g_{z,w} + v_{z,w} j \end{aligned}$$

where constants (g_w, v_w, g_z, v_z) *are positive and* $(g_{z,w}, v_{z,w})$ *can be either positive or negative. Then the unique solution* $(c^*(s^j), l^*(s^j))$ *to Problem P1 has the property that, within a birth cohort or in cross section, variances and covariances of the logs of consumption, hours and wages evolve linearly with age according to*

$$\Delta var(\log c^*(s^j)) = \frac{1}{\rho^2}\Delta var(z(s^j)) = \frac{1}{\rho^2}v_z \quad (1)$$

$$\Delta var(\log l^*(s^j)) = \frac{1}{\phi^2}\Delta var(\log w(s^j)) = \frac{1}{\phi^2}v_w \quad (2)$$

$$\Delta cov(\log c^*(s^j), \log w(s^j)) = \frac{1}{\rho}\Delta cov(z(s^j), \log w(s^j)) = \frac{1}{\rho}v_{z,w} \quad (3)$$

$$\Delta cov(\log c^*(s^j), \log l^*(s^j)) = \frac{1}{\phi\rho}\Delta cov(z(s^j), \log w(s^j)) = \frac{1}{\phi\rho}v_{z,w} \quad (4)$$

$$\Delta cov(\log l^*(s^j), \log w(s^j)) = \frac{1}{\phi}\Delta var(\log w(s^j)) = \frac{1}{\phi}v_w. \quad (5)$$

Proof: The uniqueness and time invariance of the solution is established in Theorem 1. The linear evolution of the variances and covariances follows in five steps. First, plug assumption A1′ into the necessary conditions for an interior solution to the problem stated in Theorem 1(i). This is done below, where Λ is the Lagrange multiplier on the resource constraint.

$$u_c(c^*(s^j), l^*(s^j), z(s^j)) = \Lambda \quad \Rightarrow \quad \exp(z(s^j))c^*(s^j)^{-\rho} = \Lambda$$
$$-u_l(c^*(s^j), l^*(s^j), z(s^j)) = \Lambda w(s^j) \quad \Rightarrow \quad l^*(s^j)^{\phi} = \Lambda w(s^j),$$

Second, take the log of each condition.

$$\log c^*(s^j) = \frac{1}{\rho} z(s^j) - \frac{1}{\rho}\log\Lambda \tag{6}$$
$$\log l^*(s^j) = \frac{1}{\phi}\log w(s^j) + \frac{1}{\phi}\log\Lambda. \tag{7}$$

Third, obtain all variances and covariances within age groups listed on the left hand sides of conditions (1)-(5) by applying conditions (6) and (7) and basic properties of variances and covariances.[5] Fourth, apply the linear assumptions on variances and covariances of exogenous variables stated in the theorem. Fifth, take first difference across ages to obtain conditions (1)-(5).⋄

3 Quantitative Implications of the Model

We now relate the model implications in Theorem 2 to the data patterns in Figure 1. Specifically, we want to see if there are model parameters that produce the magnitudes of the linear age trends in the data. We do so in two steps. First, we discuss the variances in Figure 1(a). The total increase in the variance of log consumption and log hours over ages 28-63 in Figure 1(a) is about .05 and .025, respectively. Theorem 2 says that this rise is related to the preference parameters (ρ, ϕ) and the variance parameters (v_z, v_w) as follows:

$$0.05 = \Delta var(\log c_j) \times J = \left(\frac{1}{\rho^2}\right) \times (v_z \times J)$$

$$0.025 = \Delta var(\log l_j) \times J = \left(\frac{1}{\phi^2}\right) \times (v_w \times J) = 0.2 \times 0.125$$

[5]The properties are $var(a+bx) = b^2 var(x)$, $cov(y, a+bx) = bcov(y,x)$ and $cov(x,x) = var(x)$.

The first condition, which governs consumption, is satisfied for any positive value of ρ (i.e. the coefficient of relative risk aversion) as the increase in the variance of the preference shifters v_z is unobserved. The second condition, which governs labor hours, places restrictions on preference parameters. Given that the rise in the variance of wages is about 0.125 over the life cycle in Figure 1 and the rise in the variance in hours is about 0.025, the square of the Frisch elasticity needs to be 0.2 for this restriction to hold. Thus, the Frisch elasticity $1/\phi$ needs to be about 0.45. Values of this magnitude or less are commonly found in the labor literature that focuses on male labor supply. Smaller Frisch elasticities will also produce patterns in Figure 1 which are visually very similar.[6]

We now see if there are model parameters that produce all of the magnitudes of the linear age trends in variances and covariances in Figure 1. Using model restrictions (1)-(5) from Theorem 2, we now proceed more formally. We pose a set of empirical moment conditions of the form $m_j = \vec{0}$ that would be satisfied by our model. These conditions hold regardless of the presence of classical measurement error as the variances of the measurement error terms are differenced away because we employ differences across age groups, as opposed to their levels.[7]

$$m_j \equiv \begin{pmatrix} \Delta var(\log w_j) - v_w \\ \Delta var(\log c_j) - \frac{1}{\rho^2} v_z \\ \Delta var(\log l_j) - \frac{1}{\phi^2} v_w \\ \Delta cov(\log c_j, \log w_j) - \frac{1}{\rho} v_{zw} \\ \Delta cov(\log c_j, \log l_j) - \frac{1}{\phi\rho} v_{zw} \\ \Delta cov(\log l_j, \log w_j) - \frac{1}{\phi} v_w \end{pmatrix}$$

The system does not separately identify ρ, v_z and v_{zw}. These three variables appear exclusively in three of these moment conditions. The second condition in this system pins down the product $\frac{1}{\rho^2} v_z$, while the fourth and fifth condition both determine the product $\frac{1}{\rho} v_{zw}$. Therefore, varying the value of ρ would imply a renormalization of the parameters (v_z, v_{zw}). For this reason, we fix the value of the coefficient of relative risk aversion ρ at a standard value of 2.5 and choose the remaining parameter values with these moment conditions in mind.

[6] The age profile of the variance of log labor hours has been characterized as being increasing, U-shaped or decreasing with age. The Appendix discusses some of the literature on this profile and highlights the point that in CEX data this estimated profile can take on any of these properties depending on sample selection criteria. Given this fact, we simply stress that our model can produce a relatively flat profile over the life cycle.

[7] Let a measured (log) variable, say $\log \hat{x}$, be equal to the true value plus a measurement error: $\log \hat{x} = \log x + \epsilon^x$, for $x = c, l, w$. Then the measurement error variance is differenced away when measurement errors ϵ^x_j are independent over time, across households and across x, and identically distributed.

Table 1: Parameter Values

ρ	$\frac{1}{\phi}$	v_z	v_w	v_{zw}
2.5	0.1847	0.0083	0.0036	0.0032

In order to examine the ability of the model to produce the patterns in Figure 1, we choose model parameters $(\frac{1}{\phi}, v_z, v_w, v_{zw})$ to minimize the following objective function mIm', where $m = (m_1', m_2'..., m_J')$ and I is the identity matrix. The parameter values which solve this problem are listed in Table 1 and the data moments implied by these model parameters have already been graphed in Figure 1. Thus, the dotted lines in Figure 1 are those produced by the model using the parameter values in Table 1.

4 Conclusion

Regarding the distribution of consumption, hours and wages across households, we close the paper with the following points:

1. Accounting for the variance profiles in Figure 1 within full-insurance models is not a challenge once one allows for preference shocks. This can be done using standard utility functions with standard values for the relevant elasticities. However, our model requires that the variance of the preference shifters increase with age.

2. Accounting for the variance and covariance profiles in Figure 1 is a greater challenge. Our model accounts for some of the main features of these profiles. However, it requires that preference shocks and wage shocks have an increasing covariance with age. The model uses this to produce an increase in the covariance between consumption and wages with age.

3. Cochrane (1991) rejects the full-insurance model using panel data on individual household consumption growth. The rejection in his work relies on the auxiliary assumption that some specific idiosyncratic variable is cross-sectionally independent of preference shifters. Specifically, his data contains information on involuntary job loss and he assumes that involuntary job loss and preference shifters are cross-sectionally independent. As consumption growth is (negatively) correlated with involuntary job loss, he rejects the full-insurance model. Our work considers cross-sectional data. Our full-insurance model

cannot produce an increase in the covariance between consumption and wages across ages, as observed in Figure 1, under the auxiliary assumption that there is no increase in the covariance between preference shocks and wage shocks with age.

References

Ales, L. and P. Maziero (2009), Accounting for Private Information, manuscript.

Cochrane, J. (1991), A Simple Test of Consumption Insurance, *Journal of Political Economy,* 99, 957- 76.

Guvenen, F. (2007), Learning your Earning: Are Labor Income Shocks Really Very Persistent?, *American Economic Review,* 97, 687- 712.

Heathcote, J., Storesletten, K. and G. Violante (2009), Consumption and Labor Supply with Partial Insurance: An Analytical Framework, NBER Working Paper 15257.

Huggett, M. (1996), Wealth Distribution in Life-Cycle Economies, *Journal of Monetary Economics,* 38, 469 - 94.

Huggett, M., Ventura, G. and A. Yaron (2011), Sources of Lifetime Inequality, forthcoming *American Economic Review.*

Kaplan, G. (2011), Inequality and the Lifecycle, manuscript.

Krueger, D. and F. Perri (2006), Does Income Inequality Lead to Consumption Inequality? Evidence and Theory, *Review of Economic Studies,* 73, 163- 93.

Storesletten, K., Telmer, C. and A. Yaron (2001), How Important Are Idiosyncratic Shocks? Evidence from Labor Supply, *American Economic Review Papers and Proceedings,* 91, 413- 17.

Storesletten, K., Telmer, C. and A. Yaron (2004), Consumption and Risk Sharing Over the Life Cycle, *Journal of Monetary Economics,* 51, 609- 33.

A-1 Appendix

A-1.1 Data

We employ the publicly available micro data of Krueger and Perri (2006). Their dataset comes from the CEX and contains household-level measures of consumption, male hours and male wages for the 1981-2003 period. We use their "ND" consumption measure, which includes nondurable components of household expenditures. To control for family size and composition, we normalize household consumption using a standard adult equivalent scale.[8] Annual consumption is computed as the sum of quarterly flows, restricting our sample to households completing all four interviews of the CEX. Annual hours of work of the head of household ("reference person" in the CEX) are computed as the product of reported "hours usually worked per week" and reported "number of weeks worked by reference person full or part time in last 12 months, including paid vacation and paid sick leave". We measure hourly wages as total annual labor income divided by annual hours of work.

Our basic sample selection criteria are the same as in the "benchmark sample" of Krueger and Perri (2006).[9] We additionally restrict the sample to households whose reference person (which is the household head in the CEX) is a white male between 25 and 65 years of age who is not retired, is not self employed, working for the army, or in the forestry/fishing sector. We also restrict the household head's annual hours of work to be between 520 and 5096, reflecting high attachment to the labor force. In order to reduce measurement error in hours, we further restrict the sample to reference persons who report working 24 hours or more per week and 16 weeks or more per year.

A-1.2 Estimating Life-Cycle Profiles

The life-cycle dispersion profiles displayed in Figure 1 is constructed in two steps. Step 1 is to calculate variances or covariances of log consumption, log hours and log wages for households grouped into age-year cells. To obtain reasonable sample sizes, we define a household to belong to cell (a, y) if the interview year is in $[y-1, y+1]$ and the age of the head of household is in $[a-2, a+2]$. This means, for example, that the relevant variance

[8]The number of adult equivalents is given by $\sqrt{a+0.5c}$, where a is the number of people over 15 years of age and c is the number of people 15 years of age or younger living in the household.

[9]We include all reference persons that responded all four interviews of the CEX, reported to live in an urban area, reported a positive dollar ammount for at least one major income category, reported an hourly wage greater than one half of the interview year's minimum wage and reported after-tax labor earnings plus transfers ("LEA+" in Krueger and Perri (2006)) equal or greater than zero.

for the age=28 and year=1985 cell was calculated using data on all households of ages 26 to 30 interviewed in years 1984 to 1986.

The Krueger-Perri data set contains interviews from the first quarter of 1980 to the first quarter of 2004. Since we consider only households with four complete quarterly interviews, our data set starts in the last quarter of 1980. Also, as in Krueger and Perri (2006), we consider households whose fourth interview was conducted in the first quarter of any year t as part of year $t-1$'s data. We drop the 1980 observations because of cell size considerations. Our sample therefore contains data for ages $a = 25, 26, ..., 65$ and years $y = 1981, 1982, ..., 2003$. Consistent with our definition of age-year groups, the resulting cells are available for $a = 27, 28, ..., 63$ and $y = 1982, 1983, ..., 2002$. The resulting cell sizes range from 66 to 545 observations, with an average of 335 observations.

Step 2 in the construction of dispersion profiles is to run an ordinary least squares regression of the relevant variance in cell (y, a) on age dummy variables and year dummy variables. The regression equation is additive in age and year and there are no cross age-year terms. The age coefficients in Figure 1 are normalized so that the variance at age 40 equals the unconditional sample variance of each variable at age 40.

A-1.3 Hours Dispersion Profiles

Our model implies a positive increase in log hours dispersion over the life cycle given an increase in wage dispersion over the life cycle - see Theorem 2. The empirical literature contains at least four different versions of the graph relating age to log hours dispersion, in addition to ours: Storeseletten et. al. (2001) and Heathcote et. al. (2009) use PSID data, whereas Kaplan (2011) uses both CEX and PSID data. Examination of these profiles and comparison to ours, leads us to conclude that, varying sample selection criteria and methodology, one can obtain increasing, decreasing, and U-shaped profiles from the same basic data set.

A-1.3.1 Profiles Based on PSID Data

Storesletten et. al. (2001) present a profile similar to ours in shape but with a larger rise in dispersion. They focus on ages 24 to 60 and find that dispersion increases by about .08 across these ages compared to .03 in our work. Most of the rise in dispersion is concentrated after age 50. Heathcote et. al. (2004) focus on ages 25 to 60 and find a U-shaped profile, where the variance of log hours decreases by around .03 between age 25 and age 40, and then rises by a similar amount until age 60. Kaplan (2011) analyzes hours dispersion with respect to potential work experience rather than age. In PSID data, he finds that dispersion decreases with potential experience by around .05 over the experience cycle.[10] His profile includes males ages 20 to 60 but with potential experience

[10]Kaplan defines *potential experience = age-years of education-6.*

levels between 3 and 32 years.

A-1.3.2 Profiles Based on CEX Data

The CEX based profile in Kaplan (2011) directly contrasts with ours. Hours dispersion in his work decreases with experience by around .05 over the experience cycle. Households age 20 to 60 are included provided that potential experience is between 3 to 32 years. Most of the decline occurs in the first 10 years of experience (around .03) with the rest of the experience cycle characterized by ups and downs of around .02 in magnitude.

To trace the sources of differences, we recalculate age profiles under three variations of the sample selection criteria. The main results are shown in Figure A1 and our conclusions are listed below:

(i) Since we do not restrict the sample based on experience, our sample includes a larger collection of older workers compared to Kaplan's. This explains why our profile *rises* at the end of the life cycle while Kaplan's does not.[11] If we take Kaplan's sample selection criteria and apply our methodology, we obtain Profile 1 (see Figure A1). The variance profile is decreasing, but ends early in the life cycle.

(ii) A criterion in our sample selection procedure, not present in Kaplan's, excludes households with less than 24 hours worked per week or less than 16 weeks worked per year. Profile 2 is produced by applying our sample selection criteria but with two changes: (1) we eliminate the hours per week and weeks per year restriction and (2) we extend the sample to include households age 20-24 when calculating age cells.

(iii) Profile 3 is produced by applying our sample selection criterion but extending the sample to include households age 20-24 when calculating age cells.

A-1.4 Proof of Theorem 1

Theorem 1, Proof:

(i) By A1, the following two functions are well defined for $\Lambda > 0$:

$$c(s^j;\Lambda) \equiv u_c(\cdot, z(s^j))^{-1}(\Lambda)$$
$$l_j(s^j;\Lambda) \equiv u_l(\cdot, z(s^j))^{-1}(-w(s^j)\Lambda)$$

These functions are continuous in Λ by the continuous differentiability of the period utility function and are strictly decreasing and increasing, respectively, by concavity. The resource constraint is strictly decreasing in Λ. Assumption A1 implies that there are values Λ for which the constraint is strictly positive and different values Λ for which the

[11]The profile in Kaplan (2011) includes workers with 3 to 32 years of potential and 20 to 60 years of age. Following his sample selection criteria as closely as possible, we find that only 53% of workers older than 45, only 30% of workers older than 49, and only 6% of workers older than 54, have less than 33 years of potential experience.

constraint is strictly negative. The Intermediate Value Theorem then implies that there is a positive value Λ^* at which the resource constraint holds with equality.

The candidate allocation is $(c^*(s^j), l^*(s^j)) = (c(s^j;\Lambda^*), l(s^j;\Lambda^*))$. This allocation satisfies the Kuhn-Tucker conditions for a solution to this problem. As these conditions are sufficient conditions for finite-dimensional, concave maximization problems, the candidate allocation solves the problem. To establish uniqueness, note that if there were a different feasible allocation solving this problem, then a convex combination of the two solutions would be feasible, by the convexity of the resource constraint, and would increase the objective since the objective is strictly concave. Contradiction.

(ii) The time-invariant allocation $(c^*(s^j), l^*(s^j))$ from Theorem 1(i) satisfies the resource constraint to Problem P1 each period and delivers a finite value for the objective function in the Planning Problem by assumptions A2 and A3. We now argue that there do not exist solutions which deliver an infinite value for the objective function. Suppose by way of contradiction that there is such a solution, then in some time period the value of the period objective must exceed the value implied by the solution constructed in Theorem 1(i). Contradiction.

We now argue that $(c^*(s^j), l^*(s^j))$ is the unique solution to Problem P1. First, any feasible allocation leading to a greater value of the objective must produce a greater value in some period. By Theorem 1(i) this can not hold. Thus, $(c^*(s^j), l^*(s^j))$ solves Problem P1. Second, it is unique as any alternative feasible allocation must by Theorem 1(i) deliver strictly less utility at some time period. $\diamond$