Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

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

Alejandro Badel
and
Mark Huggett

Working Paper 2010-046C
http://research.stlouisfed.org/wp/2010/2010-046.pdf

October 2010
Revised February 2014

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

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

Alejandro Badel and Mark Huggett*

February 19, 2014

Abstract

The life-cycle patterns of consumption, wage and hours inequality observed in U.S. cross-section data are commonly viewed as incompatible with a Pareto efficient allocation. We determine the extent to which these qualitative and quantitative patterns can or cannot be produced by Pareto efficient allocations in models with preference shocks, wage shocks and full information.

JEL Classification: E21, D91, D52

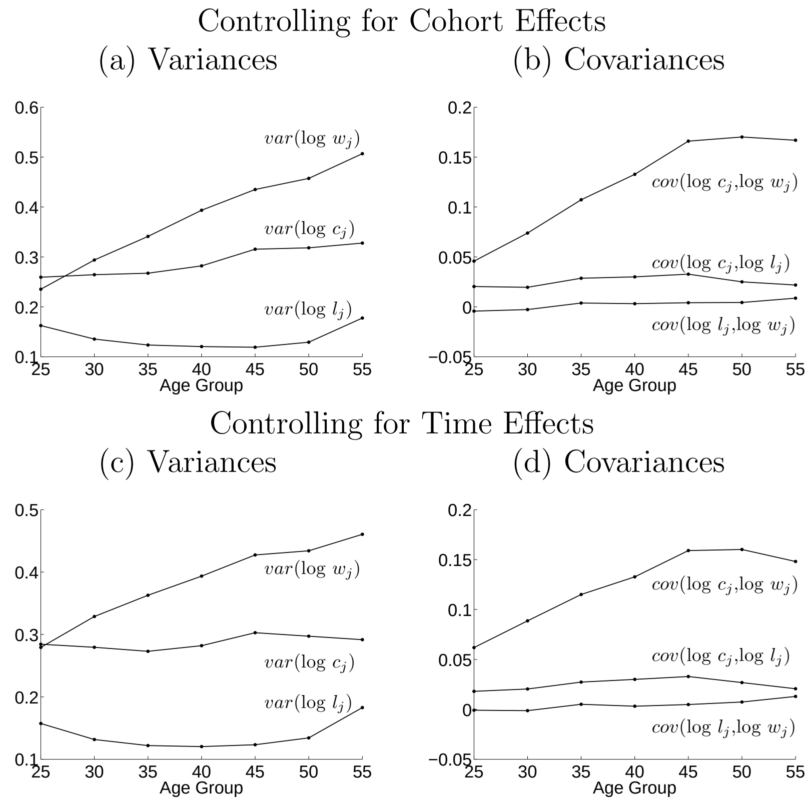
Keywords: Life Cycle, Inequality, Efficient Allocation, Preference Shocks

*Affiliation: Federal Reserve Bank of St. Louis and Georgetown University, respectively. We thank two referees, Roozbeh Hosseini, Gustavo Ventura, Jonathan Heathcote and especially the editor, Gianluca Violante, for valuable comments. Corresponding author: Alejandro Badel. E-mail address: badel@stls.frb.org. Telephone: 1(314)413-9995. Fax: 1(314)444-8731. Mailing Address: Research Division, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166.

1 Introduction

Figure 1 shows the life-cycle inequality patterns 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 in cross-section data while the variance of log hours is relatively flat over much of the life cycle.

Figure 1: Life-Cycle Inequality Patterns



Source: Authors' calculations based on Heathcote, Perri and Violante (2010) and their publicly available CEX data covering years 1980-2006. Each age group includes the five consecutive years of age starting at the group's label. All profiles are displaced vertically to match the corresponding sample variance or covariance for the 40-44 age group.

[1]Data, sample selection criteria, and details of the statistical methodology for controlling for time or cohort effects are described in the Appendix.

A common view is that a plausible explanation for the slopes of the cross-sectional variance profiles in Figure 1 involves the introduction of some type of friction so that a portion of the idiosyncratic shocks that affect wages and lead to increasing wage dispersion is transmitted to consumption. The literature has examined the following frictions: incomplete markets, limited commitment and private information.[2]

In this paper we examine the extent to which Pareto efficient allocations in models without frictions are able to produce the qualitative and quantitative patterns shown in Figure 1. More specifically, we consider models with preference and wage (i.e. productivity) shocks and determine the extent to which they can match features of how the variances and covariances in Figure 1 change with age.

One reason why such an exercise is interesting is that the interpretation of Figure 1 has profound normative implications. In an incomplete-markets interpretation of Figure 1, substantial welfare gains are possible. For example, Huggett and Parra (2010) analyze an incomplete-markets model and find social insurance reforms that deliver welfare gains that are equivalent to increasing consumption by 4 percent. Heathcote, Storesletten and Violante (2008) analyze an incomplete-markets model and find that welfare gains from hypothetically completing all markets are equivalent to increasing consumption by 40 percent. These implications motivate further inquiry into the key features of cross-sectional or panel data that indicate that models without frictions are untenable.[3]

The main findings of this paper are in two parts. First, we show that a simple model without frictions can produce the rough magnitudes of the average slopes of all the profiles displayed in Figure 1. This holds for a standard parametric class of preferences when the key preference parameters are within the range estimated in a large applied literature. Thus, contrary to some conventional wisdom, the simplest models without frictions are not so

[2]See Huggett (1996), Storesletten, Telmer and Yaron (2004), Guvenen (2007), Huggett, Ventura and Yaron (2011), Kaplan (2012) and Heathcote, Storesletten and Violante (2013) for work based on incomplete markets. Krueger and Perri (2006) analyze limited commitment - a situation in which an agent can walk away from the terms of a contract. Ales and Maziero (2009) analyze efficient allocations when idiosyncratic shocks are privately observed.

[3]We acknowledge that frictionless models and models featuring constrained-optimal allocations, such as private-information models, have similar welfare implications when compared to an incomplete markets interpretation of Figure 1.

easily eliminated as positive explanations for the slopes of the cross-sectional profiles in Figure 1. Second, we show that the same class of models is able to produce both the levels and the slopes of the profiles in Figure 1 when measurement errors, with a plausible nature and magnitude, are taken into consideration

The paper is organized as follows. Section 2 presents the model and provides a theorem that describes the nature of the inequality patterns implied by solutions to a planning problem. Section 3 describes the extent to which the model can produce the data patterns in Figure 1. Section 4 concludes.

1.1 Related Literature

Our work is most closely linked to three related literatures. The first is the literature on formal tests of full consumption insurance. Cochrane (1991) and Attanasio and Davis (1996) provide evidence that models of Pareto efficient allocations without frictions cannot produce some aspects of U.S. data. For example, Cochrane (1991) tests whether household consumption growth is cross-sectionally independent of some variables that one might conjecture are independent of household preference shocks. He finds that consumption growth is not independent of some variables (e.g. a long illness or involuntary job loss) and, thus, rejects full consumption insurance based on this evidence and the assumption that preference shocks are independent of a long illness or an involuntary job loss. However, he does not reject independence for other variables such as a spell of unemployment, loss of work due to strike and an involuntary move. Our work differs because our main interest is on cross-sectional rather than panel data. In addition, our question is quite different: we ask whether several main features of cross-sectional, inequality data can or cannot be produced by a class of full-insurance models.

Second, the work of Storesletten, Telmer and Yaron (2001) is close in spirit to our exercise. They ask whether variance profiles like those in Figure 1 can be produced by a Pareto efficient allocation in a model in which all agents have the same non-separable preferences between consumption and leisure, given the increasing pattern of wage dispersion. Nonseparability might help to produce the variance patterns as more hours of work may increase the marginal utility of consumption, leading those with high wages to have high hours and high consumption. They find that there are no utility function parameters for which their model produces the observed rise in consumption dispersion with a rise in hours dispersion that is consistent with data. They conclude that the cross-sectional rise in variance profiles alone is sufficient to imply that a class of models without frictions cannot produce some of the main inequality patterns found in U.S. data. The Appendix revisits this issue based on the updated facts in Figure 1. While we find that their model can pass their test based on the updated facts in Figure 1, we find that their model cannot produce both the slopes of the variance and covariance profiles in Figure 1.

Third, the work of Kaplan (2012) and Heathcote et. al. (2013), 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.[4] 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 without frictions when idiosyncratic wage and preference shocks are allowed?

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 $P(s^j)$. We assume that $P(s^j)$ is also the fraction of agents in a birth cohort receiving shock history 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 determined by the period utility function u, a discount factor β and the probability φ_j of surviving up to age j:

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

[4]We acknowledge that preference shocks may be said to play a more important role in our work for producing the cross-section facts than in either of these papers.

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 $t = 1$ whose members are alive at $t = 1$. The weights associated with agents born in year b and individual history s^0 are determined by the product of cohort size N_b and Pareto weight $\hat{\gamma}_b(s^0)P(s^0)$.

This problem is stated below, where we reorganize the planning objective so that all the utility flows of the different cohorts that are alive in period t enter into the square-bracketed term. Our notation uses the fact that time t, birth year b and the age j satisfy $b = t - j$. The resource constraint states that total consumption equals total output at each time period.

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

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 and continuously differentiable. Furthermore, u is strictly increasing in consumption and decreasing in labor and satisfies the Inada conditions so that the ranges of the marginal utilities of consumption and labor are $(0, \infty]$ and $(-\infty, 0]$ respectively.

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

A3: Planning weights satisfy $\hat{\gamma}_b(s^0) = \beta^b \exp(\gamma(s^0)), \forall b$, where $\gamma(s^0)$ is a random variable that determines the shock-specific component of the Pareto weight.

We rewrite the planner's objective below, using 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 period resource constraint.

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

Theorem 1: Assume A1 - A3.

(i) *There exists a unique allocation* $(c^*(s^j),l^*(s^j))$ *that solves the problem of maximizing* $E\left[\sum_{j=0}^{J}\frac{\varphi_j \exp(\gamma(s^0))}{(1+n)^j}u(c(s^j),l(s^j),z(s^j))\right]$ *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. These implications can be compared to the age patterns in Figure 1. They are based on the additively separable, iso-elastic utility function with multiplicative preference shifters stated in assumption $A1'$. This functional form is widely used in the empirical literature. It implies that in a solution to problem P1 (log) consumption and labor depend linearly on the exogenous variables (i.e. log wages, Pareto weights and the preference shock). This in turn implies that the variance-covariance properties of the observable variables (i.e. log consumption, log hours and log wage) are then linear combinations of the elements of the variance-covariance matrix of the exogenous variables. This simple fact leads us to directly posit in assumption $A4$ that the variances and covariances of the exogenous variables move linearly with age. The consequence is then that variances and covariances of observable variables in a solution to P1 will rise or fall linearly with age.[5]

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

[5] Note that the linear evolution assumed in A4 implicitly restricts both the Pareto weight variance and the preference shifter-Pareto weight covariance to be constant with age. This implies that the slopes of the linear profiles of endogenous variables will be controlled by the four slopes of exogenous linear profiles $(v_w, v_z, v_{wz}, v_{w\gamma})$ and the parameters of the utility function.

A4: The variance-covariance matrix $\Omega_j = var(z_j, \log w_j, \gamma)$ is positive semi-definite and evolves linearly with age.

$$\Omega_j = \begin{pmatrix} g_z + v_z j & g_{wz} + v_{wz} j & g_{z\gamma} \\ g_{wz} + v_{wz} j & g_w + v_w j & g_{w\gamma} + v_{w\gamma} j \\ g_{z\gamma} & g_{w\gamma} + v_{w\gamma} j & g_\gamma \end{pmatrix}$$

Theorem 2 shows how the variances and covariances analyzed in Figure 1 rise or fall linearly with age in the model economies under two alternative, additional restrictions on exogenous variables. The first restriction says that Pareto weights do not display a changing covariance with productivity across age. The second restriction says that preference shocks do not display a changing covariance with productivity across age. Both restrictions allow for positive or negative covariation among the exogenous variables but restrict how this covariation moves with age. These restrictions lead to different possible ways to account for age variation in the empirical variances and covariances examined in Figure 1.

Theorem 2: Assume A1′ and A2 - A4.

(i) If there is no change across age in the covariance between Pareto weights and productivity (i.e. $v_{w\gamma} = 0$), then the variance-covariance matrix $\Sigma_j \equiv var(\log c_j^*, \log l_j^*, \log w_j)$ produced by a solution to P1 evolves linearly with age according to the increment

$$\Sigma_{j+1} - \Sigma_j = \begin{pmatrix} (\frac{1}{\rho})^2 v_z & (\frac{1}{\rho\phi}) v_{wz} & (\frac{1}{\rho}) v_{wz} \\ (\frac{1}{\rho\phi}) v_{wz} & (\frac{1}{\phi})^2 v_w & (\frac{1}{\phi}) v_w \\ (\frac{1}{\rho}) v_{wz} & (\frac{1}{\phi}) v_w & v_w \end{pmatrix}$$

(ii) If there is no change across age in the covariance between preference shocks and productivity (i..e. $v_{wz} = 0$), then the variance-covariance matrix $\Sigma_j \equiv var(\log c_j^*, \log l_j^*, \log w_j)$ produced by a solution to P1 evolves linearly with age according to the increment

$$\Sigma_{j+1} - \Sigma_j = \begin{pmatrix} (\frac{1}{\rho})^2 v_z & (\frac{1}{\rho\phi}) v_{w\gamma} & (\frac{1}{\rho}) v_{w\gamma} \\ (\frac{1}{\rho\phi}) v_{w\gamma} & (\frac{1}{\phi})^2 [v_w - 2v_{w\gamma}] & (\frac{1}{\phi}) [v_w - v_{w\gamma}] \\ (\frac{1}{\rho}) v_{w\gamma} & (\frac{1}{\phi}) [v_w - v_{w\gamma}] & v_w \end{pmatrix}$$

Proof: Theorem 1 established that the allocation $(\log c_j^*, \log l_j^*)$ solves P1. The linear evolution of the variances and covariances follows in three steps.

First, substitute assumption A1$'$ into the necessary conditions for an interior solution to the problem stated in Theorem 1. This is done below, where λ is the Lagrange multiplier on the resource constraint.

$$\begin{aligned}\exp(\gamma(s^0))u_c(c^*(s^j), l^*(s^j), z(s^j)) = \lambda &\Rightarrow \exp(\gamma(s^0))\exp(z(s^j))c^*(s^j)^{-\rho} = \lambda \\ -\exp(\gamma(s^0))u_l(c^*(s^j), l^*(s^j), z(s^j)) = \lambda w(s^j) &\Rightarrow \exp(\gamma(s^0))l^*(s^j)^{\phi} = \lambda w(s^j)\end{aligned}$$

Second, take logs.

$$\log c^*(s^j) = \frac{1}{\rho}[z(s^j) + \gamma(s^0) - \log \lambda] \quad (1)$$

$$\log l^*(s^j) = \frac{1}{\phi}[\log w(s^j) - \gamma(s^0) + \log \lambda]. \quad (2)$$

Third, obtain each element of the variance-covariance matrix Σ_j by applying the basic properties of variances and covariances to the result of the previous step and by applying the assumptions on variances and covariances of exogenous variables stated in assumption A4. Clearly, only the parameters governing how the covariances of exogenous variables change with age enter into the difference $\Sigma_{j+1} - \Sigma_j$. $\diamond$

3 Quantitative Implications of the Model

We present the quantitative implications of the model in four steps. First, we outline a method for relating model slopes to data slopes. Second, we present the findings based on this method. Third, we present findings based on using both the slopes and levels of the profiles in Figure 1. Fourth, we present findings after imposing a particular mechanism governing how individual productivity and Pareto weights covary with age.

3.1 Method

We acknowledge that (log) consumption, hours and wages in CEX data are measured with error. This motivates us to focus on moments of the data that are insensitive to a plausible error structure. We posit that measurement errors for consumption, hours and wages are additive in logs, uncorrelated with true values and that the distribution of these errors is age and time invariant. If so, then the change in a specific variance or covariance across ages or across time periods should be due only to the change in the statistic of the true data in large samples. This leads us to focus on the change across age groups in variances and covariances.

We now provide a specific method for relating model parameters to data. Define the vector of moment conditions m_j as indicated below, using the first set of restrictions from Theorem 2. For a specific age j this condition involves the difference (which equals the slope) of the empirical variance and covariance profiles across adjacent ages from Figure 1 (e.g. $\Delta var(\log c_j) \equiv var(\log c_j) - var(\log c_{j-1})$) and the slope implied by the model economy.[6]

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

We choose model parameters $(\frac{1}{\phi}, v_z, v_w, v_{wz})$ to minimize the objective function mIm', where $m = (m_1', m_2' ..., m_J')$ and I is the identity matrix. In choosing model parameters, we fix the preference parameter ρ at $\rho = 2.5$, which is a plausible value from the literature. We fix it because the system of moment conditions does not separately identify parameters (ρ, v_z, v_{wz}). To see this note that these three parameters appear exclusively in three of the moment conditions: the first condition pins down the product $\frac{1}{\rho^2} v_z$, while the fourth and fifth conditions both determine the product $\frac{1}{\rho} v_{wz}$. Varying the value of ρ would therefore just imply a renormalization of parameters (v_z, v_{wz}).

[6] The moment conditions involve the change in the variance or covariance of measured variables and the change in the model-produced (i.e. true) variance or covariance. The assumptions on measurement errors imply that the change in the variance or covariance of measured variables equal the change in the true variance or covariance. More specifically, all variances of measured variables equal true variances plus measurement error variances. To make statements about the levels of covariances of measured variables compared to true variables, we need stronger assumptions. If the measurement errors for consumption, labor and earnings are also uncorrelated, then the level of the covariance of measured and true variables are equal, except for $cov(\log l_j, \log w_j)$ as measurement error in labor hours impacts both measured labor and measured wage.

The method we use for relating model parameters to data is similar for the second set of restrictions indicated in Theorem 2. The only differences are that the model-implied slope conditions differ and that the vector of model parameters is $(\frac{1}{\phi}, v_z, v_w, v_{w\gamma})$ instead of $(\frac{1}{\phi}, v_z, v_w, v_{wz})$.

Table 1: Parameter Values

Case 1: $v_{w\gamma} = 0$

Parameter	Cohort Effects	Time Effects
ρ	2.5 (−)	2.5 (−)
$\frac{1}{\phi}$	0.04621 (0.00793)	0.08674 (0.00962)
v_z	0.0142 (0.00072)	0.0015 (0.00037)
v_{wz}	0.0101 (0.00025)	0.0072 (0.00013)
v_w	0.0091 (0.00013)	0.006 (0.00008)

Case 2: $v_{wz} = 0$

Parameter	Cohort Effects	Time Effects
ρ	2.5 (−)	2.5 (−)
$\frac{1}{\phi}$	0.001 (0.00105)	0.001 (0.00013)
v_z	0.0142 (0.00072)	0.0015 (0.00037)
$v_{w\gamma}$	0.01009 (0.00025)	0.00718 (0.00013)
v_w	0.0091 (0.00013)	0.006 (0.00008)

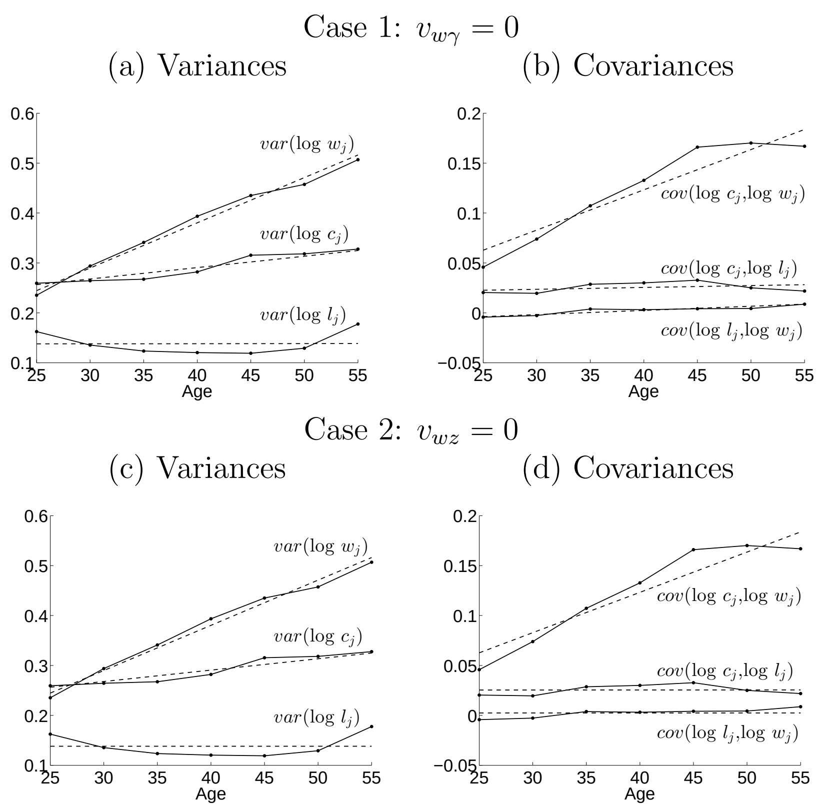
Note: Standard errors are reported in parentheses next to each parameter value. Standard errors are based on repeating the estimation procedure 200 times. Each estimation procedure resamples with replacement from each age-year cell in the CEX. The number of observations in each age-year cell is the same as in the CEX.

The results of applying this method to data are presented in Table 1. Parameter values are reported for the case of the time effects and the cohort effects specifications underlying Figure 1. We note that from the perspective of a large applied literature, the values of the model parameters $(\rho, 1/\phi)$ in Table 1 do not seem out of line. For example, Browning, Hansen and Heckman (1999) survey estimates of these parameters. They highlight that the Frisch elasticity $1/\phi$ varies from about 0 to typically below 0.5 for males.

3.2 Main Findings: Slopes

The main findings for our analysis of slopes are presented graphically in Figure 2 and Figure 3. To understand what is being plotted, focus for the moment on Figure 2. The solid lines plot the data properties previously highlighted in Figure 1 for the case of cohort effects. The dotted lines plot the model implications based on the parameter vector from Table 1. In plotting the model implications we employ a free vertical scaling of each variance and covariance as a visual aid to match the average level of the empirical profiles. The reader will recall that our method of relating model parameters to data is to treat the slopes in the graphs as the moments that are reliably estimated.

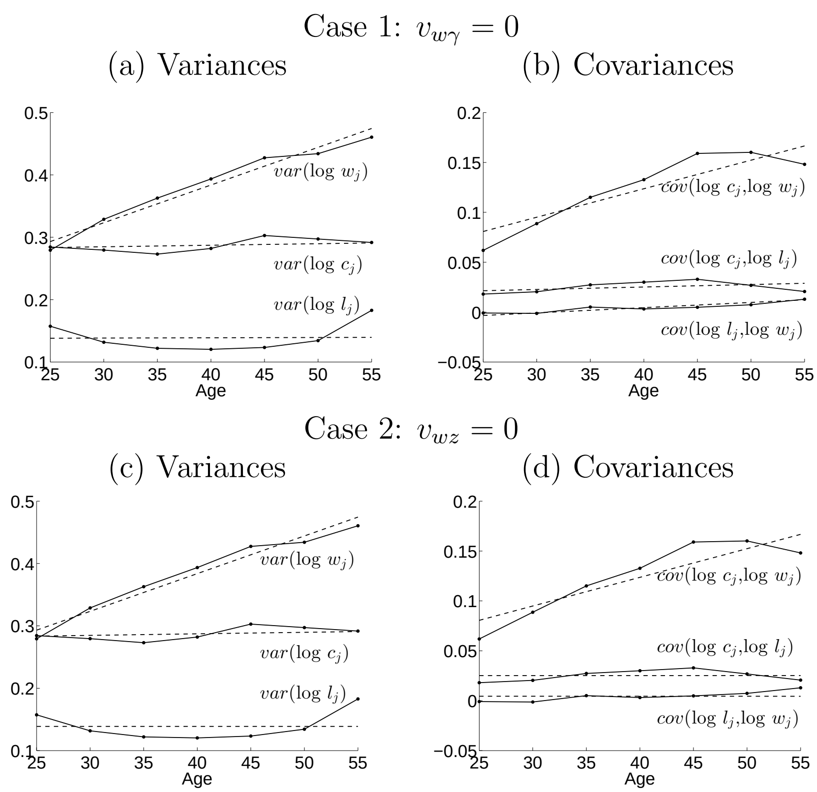
Figure 2: Rise in Life-Cycle Inequality: Data and Model (Cohort Effects)



NOTE: The solid lines show the data properties from Figure 1 based on cohort effects. The dashed lines show model implications based on Table 1. Each model-generated profile is vertically shifted, as a visual aid, to match the average level of the corresponding empirical profile.

Figure 2 is in two parts. The upper panel presents findings for the case when Pareto weights are restricted as described in the first part of Theorem 2, whereas the bottom panel presents findings for the case when preference shocks are restricted as described in the second part of Theorem 2. Figure 3 is analogous to Figure 2 but handles the case when the data moments are estimated using time effects rather than cohort effects.

Figure 3: Rise in Life-Cycle Inequality: Data and Model (Time Effects)



NOTE: The solid lines show the data properties from Figure 1 based on time effects. The dashed lines show model implications based on Table 1. Each model-generated profile is vertically shifted, as a visual aid, to match the average level of the corresponding empirical profile.

One reading of Figure 2 and 3 is that the model economies can roughly match the average slopes of the empirical profiles. This holds for either of the two restrictions on how covariances of exogenous model variables change with age. This also holds when the model is asked to match the empirical profiles for the case of either time or cohort effects. Thus, we do not find a spectacular failure of even the simplest models without frictions to match the rough patterns in how variances and covariances change with age in cross-section data.

We acknowledge that this reading of Figure 2 and 3 may not be universal. For example, some researchers may highlight the U-shaped pattern in the variance of log hours by age as a feature that the simple model cannot produce. While this observation is correct, we think that many researchers may be surprised that the simple model is not even more robustly incapable of producing the empirical slope properties in Figure 1. These researchers may want to probe more deeply to understand the mechanisms by which the model accounts for the data moments.

Rise in Variances

How does the model account for the average slopes in the variance of log wages, consumption and hours? Wages are exogenous in the model and the parameter v_w is selected partly on the basis of producing the empirical rise in the variance of log wages. One mechanism that produces this near linear rise assumes that agents are hit with highly persistent idiosyncratic productivity shocks. Thus, productivity dispersion spreads out for a cohort over time due to the accumulation of these period-by-period shocks. While this particular story for productivity dynamics is not essential for producing the model implications, we mention it as it is widely explored in the literature.

The rise in the variance of log consumption is entirely accounted for by the rise in the variance of preference shocks. To see this, note that from the proof of Theorem 2

$$\log c^*(s^j) = \frac{1}{\rho}[z(s^j) + \gamma(s^0) - \log \lambda].$$

This implies that, other things equal, consumption increases as the preference shifter $z(s^j)$ increases. Moreover, $\Delta var(\log c^*(s^j)) = (\frac{1}{\rho^2})v_z$ holds within the model. Thus, an increasing variance of preference shifters (i..e. $v_z > 0$) accounts for any rise in consumption dispersion.[7] The consumption variance

[7] Assumption A4 in Theorem 2 rules out a rise in the variance of log consumption

patterns restrict the values of the product $(\frac{1}{\rho^2})v_z$ but not either component as neither is observed.

The empirical variance profile of log hours is U-shaped. The model can produce either a linear increase or a decrease in dispersion. To see this, labor hours satisfies

$$\log l^*(s^j) = \frac{1}{\phi}[\log w(s^j) - \gamma(s^0) + \log \lambda]$$

in the proof of Theorem 2. Thus, $\Delta var(\log l^*(s^j)) = (\frac{1}{\phi^2})[v_w - 2v_{w\gamma}]$. Table 1 presents parameter estimates based on two cases that impose extra restrictions on model primitives. Case 1 imposes that any covariation between wages and Pareto weights does not change with age (i.e. $v_{w\gamma} = 0$). With this restriction and the magnitude of v_w pinned down by the empirical rise in the variance of log wages, the Frisch elasticity parameter $1/\phi$ needs to be well below 1 to match the empirical variance profile. Case 2 from Table 1 does not impose a restriction on the parameter $v_{w\gamma}$. Thus, Case 2 allows the model produced hours variance profile to increase or decrease linearly with age.

Rise in Covariances

We highlight two properties of covariances from Figure 1. First, the average slope of the consumption-productivity covariance is positive. Second, the average slope of the consumption-productivity covariance is substantially larger than the average slope of the consumption-hours covariance.

How does the model economy produce these two data properties? We start by recalling the necessary conditions from Theorem 2, which were displayed in the previous subsection. A quick look at these conditions reveals that they imply that $\Delta cov(\log c_j^*, \log w_j) = \frac{1}{\rho}[v_{wz} + v_{w\gamma}]$ and that $\Delta cov(\log c_j^*, \log l_j^*) = \frac{1}{\rho\phi}[v_{wz} + v_{w\gamma}]$. Thus, to get these two data properties we need two things: $1/\phi < 1$ and $v_{wz} + v_{w\gamma} > 0$. First, the Frisch elasticity $1/\phi$ needs to be well below 1 to match the substantial difference in the slopes. Earlier we argued that the Frisch elasticity needed to be below 1 to account for the fact that the slope of the hours variance profile is flatter than that for wages. Second, either the preference shock-productivity covariance or the Pareto weight-productivity covariance needs to increase with age

arising from an increasing covariance between preference shocks $z(s^j)$ and Pareto weights $\gamma(s^0)$.

so that the sum is positive. This fact led us to usefully consider two types of restrictions $v_{wz} = 0$ or $v_{w\gamma} = 0$ in Theorem 2.

This discussion also highlights the fact that the model is not empirically vacuous. The model economies covered in Theorem 2 do not allow for arbitrary linear patterns in variances and covariances. For example, $cov(\log c_j^*, \log w_j)$ and $cov(\log c_j^*, \log l_j^*)$ must either both increase or both decrease with age. They cannot go in opposite directions. Put somewhat differently, if they went in opposite directions in the data, then they could not be produced by a Pareto efficient allocation within this model.

3.3 Main Findings: Levels and Slopes

Given that the model can produce the rough magnitudes of the slope patterns in Figure 1, we ask whether additional moment conditions from cross-sectional data highlight the empirical implausibility of the model. Thus, we ask if the model can produce both the level and slope patterns in Figure 1. To carry out this exercise, we make assumptions on the nature as well as the plausible magnitudes of measure error in the CEX data. The basic assumptions on measurement error in log earnings, hours and consumption are those listed in Section 3.1. Since wages are measured as total reported earnings per hour of reported work, measurement error in earnings and hours imply that measured log wage equals true log wage plus the error in earnings less the error in hours. This implies that measurement error acts to bias downwards the empirical wage-hours covariance. We choose model parameters so that model implied variances and covariances plus the effect of measurement error matches the levels of all the empirical variances and covariances in Figure 1.[8]

We use external magnitudes of the variance of measurement errors in earnings, hours and consumption. Several studies have attempted to determine the nature of measurement error in household surveys. One strategy compares survey responses to error-free measures of earnings, hours and wages. Bound and Krueger (1991) compare March CPS survey responses with the corresponding earnings reported to the Social Security Administration in 1978 for a sample containing 444 men. Bound, Brown, Duncan and Rodgers (1994) use the PSID validation study. In this study, a cohort of 418 workers

[8]A straightforward extension of Theorem 2 provides closed-form solutions for all model-implied variances and covariances as a function of the model primitives specified in the assumptions of Theorem 2.

Table 2: Estimates of Measurement Error Variances

Study	σ_e^2	σ_l^2	σ_c^2
Bound and Krueger (1991)	0.114		
Bound et al. (1994), data from 1982	0.01988	0.01232	
Bound et al. (1994), data from 1986	0.01464	0.02132	
Guvenen and Smith (2013)	0.022		0.126
Average	0.04263	0.01682	0.126

at a large firm were interviewed in 1983 and 341 of them, plus 151 new workers, were interviewed again in 1987. Workers were asked questions about annual earnings, hours, and tenure similar to those used by the PSID survey. Interview responses were compared to detailed company records containing accurate measures of earnings and hours worked.

A second strategy is via the structural estimation of life-cycle models. Guvenen and Smith (2013) estimate a structural consumption-savings model using data from the PSID and the CEX.[9] They separately identify measurement error in income and shocks to income by focusing on the annual changes in income and consumption of borrowing-constrained households. In their model consumption changes are equal to income shocks for these households. In order to capture initial differences in wealth and household size effects that are not present in their model, they allow measurement error in consumption to have a permanent component in addition to transitory measurement error. We focus exclusively on their estimates of the transitory component.

Table 2 lists the estimates of the variance of measurement error for earnings σ_e^2, hours σ_l^2 and consumption σ_c^2 from the work cited above. We use the averages in the last row of the table as external estimates in our exercise.

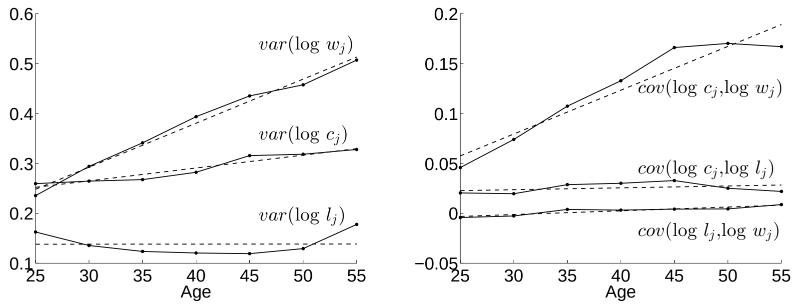
Table 3 contains the model parameter values for which the simple model together with measurement error best match the empirical profiles in Figure 1. The model parameters are those governing preferences $(\rho, 1/\phi)$ and those listed in Assumption A4 that govern the variances and covariances of preference shifters, wages and Pareto weights. The reader will recall that g-values

[9]Their study employs household consumption and after-tax non financial household income data for a sample of married couples in the PSID. The consumption measure is constructed using an imputation equation based on food consumption and other household characteristics. The parameters of the imputation equation are estimated using CEX data.

Figure 4: Life-Cycle Inequality: Data and Model (Cohort Effects)

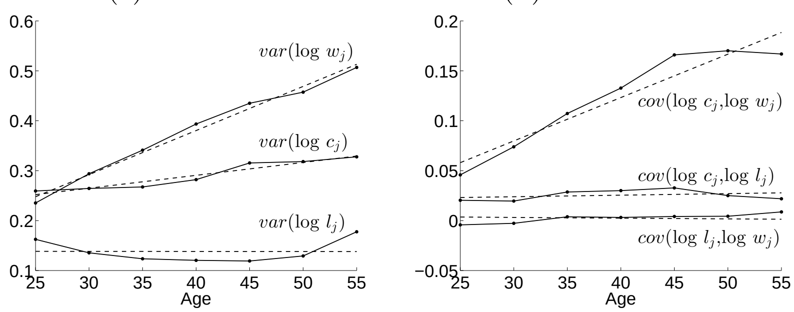
Case 1: $v_{w\gamma} = 0$

(a) Variances (b) Covariances



Case 2: $v_{wz} = 0$

(c) Variances (d) Covariances



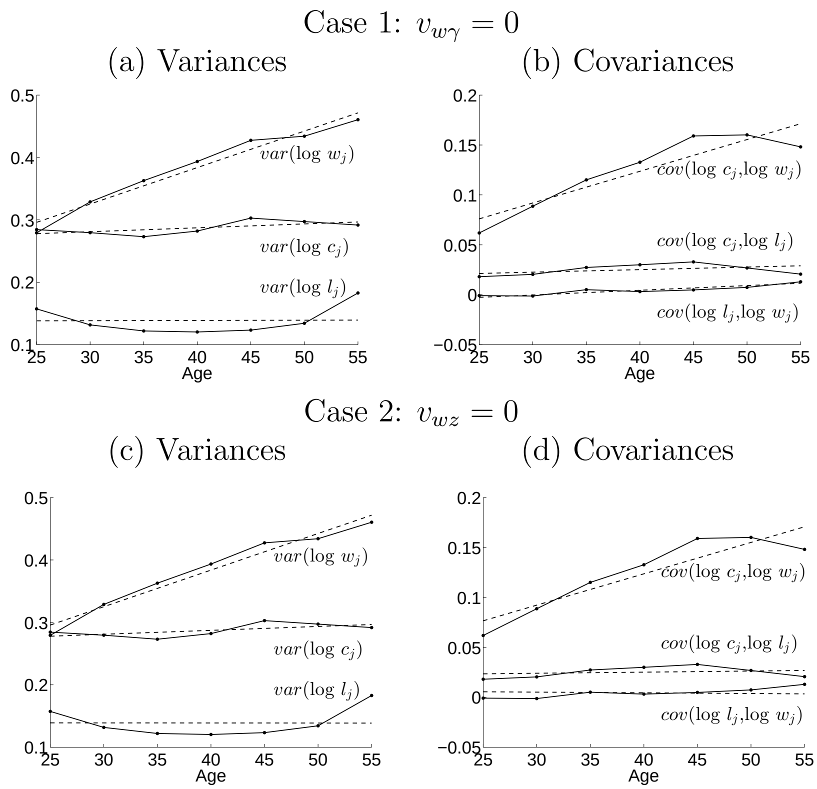
NOTE: The solid show the data properties from Figure 1. The dashed lines show model implications based on Table 2-3.

are intercepts and v-values determine slopes.[10]

Figure 4 and 5 show the model implications based on the results of Table 2 and 3. Unlike Figure 2-3, Figure 4-5 do not use a free vertical scaling. The model under Case 1 and Case 2 is able to produce the slopes and approximate levels of all the variance-covariance profiles from Figure 1. Thus, extending the list of moments to be explained to include the level and slopes of the variance-covariance profiles does not clearly eliminate the simple model as

[10]One notable implication of the parameter values in Table 3 is that the correlation between preference shifters z_j and Pareto weights γ is close to -1 over the life cycle. The log wage-preference shifter correlation is positive and roughly 0.1 at many ages. The log wage and Pareto weight γ correlation is close to zero at most ages.

Figure 5: Life-Cycle Inequality: Data and Model (Time Effects)



NOTE: The solid show the data properties from Figure 1. The dashed lines show model implications based on Table 2-3.

an explanation of the cross-section inequality facts in Figure 1.

3.4 Wage-Consumption Covariance: A Mechanism

We have two distinct stories for how the model produces the rise in the wage-consumption covariance with age. One relies on an increase in the wage-Pareto weight covariance with age (i.e. Case 2: $v_{w\gamma} > 0$), whereas the other relies on an increase in the covariance between preference shocks and wages with age (i.e Case 1: $v_{wz} > 0$). While both stories work in the limited sense demonstrated in Figures 2-5, we have not provided an economic mechanism that would produce either story nor have we provided direct measurement for either story.

This section does two things. First, it provides an economic mechanism that can produce an increase in the wage-consumption covariance with age via an increase in the wage-Pareto weight covariance. Second, it provides measurement consistent with the economic mechanism. This measurement provides empirical discipline that restricts how much the wage-Pareto weight covariance can increase with age.

Pareto weights attached to an agent's utility in Planning Problem P1 remain fixed upon the entry of an agent into the model. Thus, the only way for the wage-Pareto weight covariance to increase with age is if labor productivity moves more in sync with these fixed Pareto weights as a cohort ages. An economic mechanism behind this is that agents differ early in life in an attribute that governs the steepness of the mean earnings or the mean wage profile. Pareto weights would then need to be positively correlated with such an attribute for the wage-Pareto weight covariance to increase with age. One such mechanism, highlighted in the human capital literature, is that agents differ early in life in the ability to learn. Good learners have steeper sloped mean earnings and mean human capital profiles than those with lower learning ability as they invest heavily and produce new skills. Huggett, Ventura and Yaron (2011) analyze exactly this mechanism within a model with idiosyncratic risk to human capital. They argue that ability differences lead to differences in mean earnings profiles and help account for increases in earnings dispersion over the life cycle.[11]

This type of mechanism is related to two empirical literatures. One vast literature documents differences in the shape of age-mean earnings or age-mean wage rate profiles for groups that differ in years of formal education. The literature finds that groups with more years of formal education have steeper profiles. The other literature posits statistical models for earnings and determines whether there is support for individual-specific slopes (see Guvenen (2009) among others).

We now posit a statistical model governing wage rates that allows for individual-specific slope differences. The model is the standard statistical model used for log earnings by Guvenen (2009) and many others. We apply it to log wage rates. Log wage rates equal the sum of a deterministic component

[11] They build on the Ben-Porath (1967) model and the analysis of this model in Huggett, Ventura and Yaron (2006).

d_{jt}, an individual fixed effect α and individual slope parameter β, a persistent shock m_j and a transitory shock ϵ_j. Innovations (ϵ_j, η_j) are uncorrelated over time and with (α, β) and $(\alpha, \beta) \sim N(0, \Sigma), \epsilon_j \sim N(0, \sigma_\epsilon^2), \eta_j \sim N(0, \sigma_\eta^2)$.

$$\log w_{jt} = d_{jt} + \alpha + \beta j + m_j + \epsilon_j$$

$$m_j = \chi m_{j-1} + \eta_j \text{ and } m_{-1} = 0$$

We estimate this statistical model using PSID wage data from HPV (2010). The results are summarized in Table 4 and our methods are described in Appendix A-5. The size of the point estimates of the standard deviation of the individual slope term σ_β are smaller than results based on earnings data surveyed in Guvenen (2009).

Following the line of argument used in Theorem 2, it is clear that variance-covariance properties of the observable variables are determined from the variance-covariance matrix $\Omega_j = var(z_j, \log w_j, \gamma)$ of the exogenous variables and the two utility function parameters. We state matrix Ω_j below.[12]

$$\Omega_j = \begin{pmatrix} g_z + v_z j & g_{z\alpha} & g_{z\gamma} \\ g_{z\alpha} & \sigma_\alpha^2 + \sigma_\beta^2 j^2 + var(m_j) + \sigma_\epsilon^2 & g_{\gamma\alpha} + (r_{\gamma\beta}\sigma_\beta\sqrt{g_\gamma})j \\ g_{z\gamma} & g_{\gamma\alpha} + (r_{\gamma\beta}\sigma_\beta\sqrt{g_\gamma})j & g_\gamma \end{pmatrix}$$

It is assumed that the covariances between (i) Pareto weights and shocks are zero (i.e. $cov(\gamma, \eta) = cov(\gamma, \epsilon) = 0$) and between (ii) preference shifters and slopes and shocks are zero (i.e. $cov(z_j, \beta) = cov(z_j, \eta_j) = cov(z_j, \epsilon_j) = 0$).

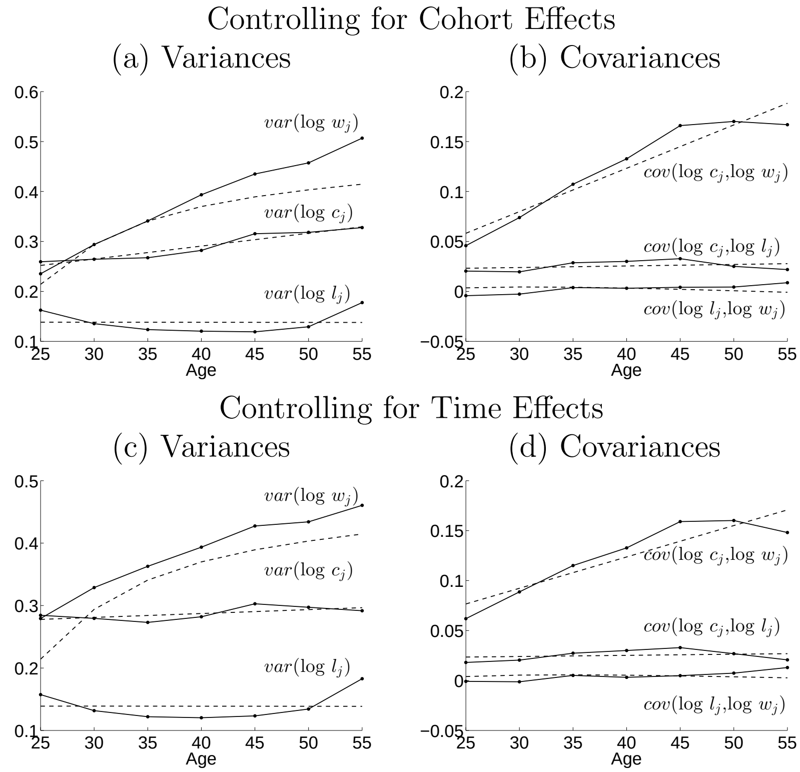
Thus, among the exogenous variables, the only covariance that is allowed to change with age is the wage-Pareto weight covariance. Furthermore, the degree to which this covariance is allowed to vary with age is restricted by the key standard deviation σ_β, which is estimated in Table 4. We let $r_{\gamma\beta}$ denote the correlation between the Pareto weight term γ and the individual slope term β.

Table 5 presents the values of the remaining model parameters that best fit the cross-sectional inequality profiles. The parameters of the wage process are preset to the values estimated in the first column of Table 4. The details for carrying out this exercise are in Appendix A-5.

Figure 6 shows the model implications for the observable variables. The model is still able to produce the qualitative properties of the data plots.

[12] The variance of the persistent shocks follows: $var(m_j) = \chi^2 var(m_{j-1}) + \sigma_\eta^2$.

Figure 6: Life-Cycle Inequality



NOTE: The solid lines show the data properties from Figure 1. The dashed lines show model implications based on Table 4-5.

The model only requires moderate positive correlation $r_{\gamma\beta}$ between Pareto weights and the individual slope terms in the wage process to match the linear rise in the consumption-wage covariance.[13]

[13]Perhaps it is surprising that the variance profile for wages is visually far from the data plots. However, model implications are entirely determined from the statistical model for wages estimated using PSID data, whereas the data plots are based on CEX data.

4 Conclusion

It may seem far-fetched that Pareto efficient allocations in a simple model without frictions might produce some of the main quantitative features of U.S. cross-sectional inequality data. If so, then it should be straightforward to highlight the data features that lead to this view. These data features could be added to the facts from panel data that help to reject models of Pareto efficient allocations under full information.

We find that first-best allocations in our model can produce the rough slope patterns for all the variance and covariance profiles in Figure 1. We also find that the model can produce the slope and level patterns for all the profiles in Figure 1 after accounting for a plausible structure and magnitude of measurement error. Thus, contrary to some conventional wisdom, we do not find it to be so straightforward to highlight some basic U.S. cross-sectional inequality facts that help to rule out simple, frictionless models.

References

[1] Ales, L. and P. Maziero (2009), Accounting for Private Information, manuscript, University of Pennsylvania.

[2] Attanasio, O. and S. Davis (1996), Relative Wage Movements and the Distribution of Consumption, Journal of Political Economy, 104, 1227-62.

[3] Ben-Porath, Y. (1967), The Production of Human Capital and the Life Cycle of Earnings, Journal of Political Economy, 75, 352-65.

[4] Bound, J., Brown, C., Duncan, G. and Willard L. Rodgers (1994), Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data, Journal of Labor Economics, 12(3), 345- 68.

[5] Bound, J. and Alan B. Krueger (1991), The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?, Journal of Labor Economics, 9(1), 1-24.

[6] Browning, M., Hansen, L. and J. Heckman (1999), Micro Data and General Equilibrium Models, Taylor, J.B., and M. Woodford (Eds.), Handbook of Macroeconomics, Elsevier Science, Amsterdam.

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

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

[9] Guvenen, F. (2009): An Empirical Investigation of Labor Income Processes," Review of Economic Dynamics, 12(1), 58-79.

[10] Guvenen, F. and Anthony Smith Jr. (2013), Inferring Labor Income Risk and Partial Insurance from Economic Choices, mimeo, Yale University.

[11] Heathcote, J., Storesletten, K. and G. Violante (2008), Insurance and Opportunities: A Welfare Analysis of Labor Market Risk, Journal of Monetary Economics, 55, 501-25.

[12] Heathcote, J., Storesletten, K. and G. Violante (2013), Consumption and Labor Supply with Partial Insurance: An Analytical Framework, American Economic Review, forthcoming.

[13] Heathcote, J., Perri F. and G. Violante (2010), Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006, Review of Economic Dynamics, 13, 15-51.

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

[15] Huggett, M. and J.C. Parra (2010), How Well Does the U.S. Social Insurance System Provide Social Insurance?, Journal of Political Economy, 118, 76-112.

[16] Huggett, M., Ventura, G. and A. Yaron (2006), Human Capital and Earnings Distribution Dynamics, Journal of Monetary Economics, 53, 265- 90.

[17] Huggett, M., Ventura, G. and A. Yaron (2011), Sources of Lifetime Inequality, American Economic Review, 101, 2923-2954.

[18] Kaplan, G. (2012), Inequality and the Lifecycle, Quantitative Economics, 3, 471- 525.

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

[20] 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.

[21] 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 Data

We use publicly available micro data from Heathcote, Perri and Violante (2010), hereafter HPV. Their dataset comes from the Consumer Expenditure Survey (CEX) quarterly Interview Survey and contains responses on quarterly household consumption expenditures, individual annual hours and individual annual wages for the 1980-2006 period. We focus on households with at least one working-age person, defined as being 25-60 years old. Since the CEX does not specify a household head, we use HPV's definition: if a working-age male is available, the head is the oldest working-age male, if not, the head is the oldest working-age female. Following HPV, we additionally restrict the sample to households for which the head reports at least 260 hours of work in the previous year.[14] Our quarterly measure of consumption is HPV's nondurable consumption.[15] Since we focus on annual flows, we measure annual consumption as the sum of quarterly flows, so our sample is restricted to households completing all four quarterly interviews of the CEX. Our annual consumption measure consists of the log of annual consumption expenditures deflated using the CPI-U and divided by the number of adult equivalents in the household (as calculated using an OECD scale described in HPV). Annual hours of work are measured as the product of records on "hours usually worked per week" and "number of weeks worked full or part time in last 12 months, including paid vacation and paid sick leave". Hourly wages are measured as an individual's total annual labor earnings divided by the individual's hours of work. Following HPV, we define individual labor earnings as the sum of wages and salaries plus two thirds of business and farm income earned.[16]

Our sample selection criteria are the same used to obtain Sample B in HPV.[17]

[14]A slight difference between our sample and HPV's is that they do not apply the hours restriction for plotting the profiles of consumption and wages. We apply the restriction to all variables since we also compute covariances. This change has almost no effect on the variance profiles of consumption and wages.

[15]This measure consists of the CPI-deflated sum of quarterly household expenditures on food and beverages (including food away from home and alcoholic beverages), tobacco, apparel and services, personal care, gasoline, public transportation, household operation, medical care, entertainment, reading material and education. Due to a methodological bias in the survey, food expenditure records in the years 1982 to 1987 are adjusted using a procedure described in HPV.

[16]For top-coded observations HPV impute mean earnings values by extrapolating a Pareto density fitted to the non-top-coded upper end of the empirical distribution.

[17]These consist of deleting a household if: (1) age of head or spouse is missing, (2) head is not working age, (3) either head or spouse has positive labor income but zero annual hours, (4) either head or spouse has an hourly wage below half of the prevailing federal

A-2 Estimating Life-Cycle Profiles

We construct life-cycle profiles of second moments controlling for year and cohort effects. For year effects, we use the following three steps. First, we split the dataset into age-year cells, compute the relevant variance (or covariance) within each cell and then collapse the dataset so there is a single observation per age-year cell. Following HPV, we put a CEX observation in the (a, y) cell if the interview was conducted during year $y = 1980, 1981, 1982, ..., 2006$, with reported head of household's age $j \in [a, a + 4]$, where a is allowed to take values $a = 25, 30, 35, 40, 45, 50, 55$. Second, we run ordinary-least-squares regressions of cell-level variances (or covariances) against a set of a and y dummy variables. The regression equation is additive in a and y dummies, without cross terms. Third, in Figure 1 we plot the vector $(d, \beta_{30} + d, \beta_{35} + d, ..., \beta_{55} + d)$, where the β_a are the estimated age coefficients and d is a vertical displacement defined as $d \equiv \widehat{var}(\log y_j | j \in [40, 44]) - \beta_{40}$, for a given variable $y = c, l, w$ and $\widehat{var}$ denotes the sample variance. For cohort effects we follow the same steps but replacing the age-year cells (a, y) with age-cohort cells (a, b) where $b = 1920, 1921, 1922, ..., 1982$.

A-3 Proof of Theorem 1

(i) By A1 and $u(c, l, z) = u_1(c, z) + u_2(l, z)$ the following two functions are well defined for $\lambda > 0$. In what follows it is understood that histories s^0 and s^j that appear in these two functions are compatible in that both share the same root term s^0.

$$c(s^j; \lambda) \equiv u_{1,c}(\cdot, z(s^j))^{-1} \left(\frac{\lambda}{\exp(\gamma(s^0))} \right)$$
$$l(s^j; \lambda) \equiv u_{2,l}(\cdot, z(s^j))^{-1} \left(\frac{-w(s^j)\lambda}{\exp(\gamma(s^0))} \right)$$

These functions are continuous in λ by the continuous differentiability of the period utility function and are strictly decreasing and increasing, respectively, in λ 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 constraint is strictly negative. The Intermediate

minimum wage, (5) reported consumption is implausible, (6) the households is flagged as an "incomplete income reporter".

Value Theorem then implies that there is a positive value λ^* at which the resource constraint holds with equality.

The candidate allocation $(c^*(s^j), l^*(s^j)) = (c(s^j;\lambda^*), l(s^j;\lambda^*))$ 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 for each period and delivers a finite value for the objective function in the planning problem by assumptions A2 and A3. 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. $\diamond$

A-4 STY Exercise

Storesletten, Telmer and Yaron (2001), henceforth STY, found that the slopes of the variance profiles for (log) consumption, wages and hours alone could not be produced by a Pareto efficient allocation within their model. Their model featured non separable preferences between consumption and labor but did not allow for preference shocks or preference heterogeneity. Thus, their model allowed for a different mechanism behind the rise in consumption dispersion compared to the model analyzed in this paper. The CEX data that we analyze has a dramatically flatter consumption dispersion profile compared to the profile analyzed by STY. Thus, one might ask whether their model continues to be incapable of producing the slopes in the variance profiles in light of more recent CEX data.

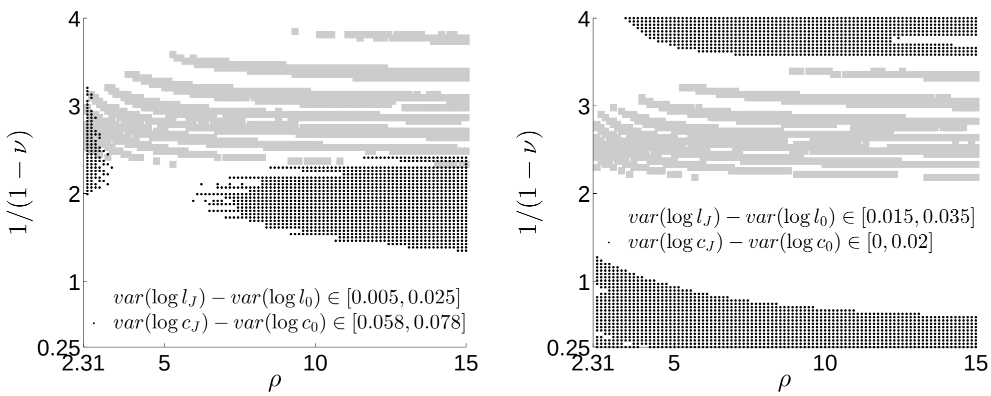
We now revisit the planning problem (Problem P1) from Section 2. We replace the period utility function used in Theorem 2 with the non separable utility function used by STY.

$$u(c,l) = \frac{1}{1-\rho}[\theta c^{\nu} + (1-\theta)(1-l)^{\nu}]^{\frac{1-\rho}{\nu}}$$

We follow STY and ask whether there are preference parameters (ρ, ν) so that the solution to the planning problem produces the observed rise in consumption dispersion in Figure 1 without an increase in labor dispersion beyond that in Figure 1, treating the rise in wage dispersion as exogenous data. Unlike the problem studied in Section 2, closed-form expressions describing the inequality implications of solutions to this planning problem are not available. For this reason, we put a fine grid on (ρ, ν) and compute, for each pair, a solution to the planning problem and the resulting cross-sectional inequality patterns. We employ STY's functional form for individual labor productivity.

Figure A-1: STY Exercise

(a) Cohort Effects (b) Time Effects



We compute the rise in the variance for 95×100 combinations with ρ ranging from 2.31 to 15 and ν ranging from -3 to 0.75. A black dot denotes that the consumption variance condition holds, whereas a gray block denotes that the hours variance condition holds. These numerical conditions are stated within panels (a) and (b).

The main result in STY was presented as a graph depicting two regions. The first region showed the combinations of the two "free" preference parameters (ρ, ν) for which the model produced the approximate empirical rise in the variance of log labor hours over the working lifetime. The second region showed the preference parameters (ρ, ν) for which the model produced the approximate empirical rise in the variance of log consumption over the working lifetime. As the two regions did not come close to intersecting, they concluded that the non separable preference mechanism was not a quantitatively viable explanation of the cross-sectional variance data.

Figure A-1 shows the results when one repeats the STY exercise using the data patterns in Figure 1. Contrary to STY, we find that there are preference parameters (ρ, ν) so that their model produces the approximate rise in consumption and hours dispersion in Figure 1, taking as given the rise in wage dispersion from Figure 1. This holds for the case of cohort effects but not for the case of time effects. Thus, their model does not completely fail this limited test. Of course, one could ask whether their model can produce the slopes in all the variance and covariance profiles in Figure 1.[18] We find that their model does not pass this more stringent test.

STY Exercise: Computation

We compute solutions to Problem P1 using the first-order conditions stated below.[19] This involves finding the Lagrange multiplier $\lambda > 0$ so that the resource constraint approximately holds (i.e. $\frac{1}{J}|\sum_j E[c_j - w_j l_j]| < 0.0001$), assuming no population growth and no mortality. After computing a solution to Problem P1, we calculate the variance of log consumption and work hours at each age. If the rise in consumption dispersion or hours dispersion fits the criteria specified in Figure A-1, then we shade the parameter combination (ρ, ν) appropriately. We repeat this procedure for each (ρ, ν) combination on a fine grid.

$$\exp(\gamma) u(c_j, 1 - l_j)^{\frac{1-\rho-\nu}{1-\rho}} \theta c_j^{\nu-1} = \lambda \tag{3}$$

$$\exp(\gamma) u(c_j, 1 - l_j)^{\frac{1-\rho-\nu}{1-\rho}} (1 - \theta)(1 - l_j)^{\nu-1} \geq \lambda w_j, \text{ with equality if } l_j > 0 \tag{4}$$

We make specific assumptions:

1. Put a rectangular grid on $(\rho, \nu) \in [2.31, 15] \times [-3, 0.75]$ with 95×100 combinations.

2. We allow 7 age groups (25 to 29, 30 to 34, 35 to 39, 40 to 44 , 45 to 50, 51 to 55, 56 to 60). For each age group we use 111×110 combinations of values of the exogenous shocks (i.e. Pareto weight and log labor productivity $(\gamma, \log w_j)$) at each age based on the joint Normal distribution implied by the labor productivity process and the parameter values in point 3 below.

3. Use STY's assumption on labor productivity.

$$\log w_j = d_j + a + x_j$$
$$x_j = x_{j-1} + e_j \text{ with } x_0 = 0,$$

[18] Here we pick preference parameters to minimize the equally-weighted squared distance between model-implied slopes and the empirical slopes in Figure 1.

[19] The procedure allows agents to work zero hours.

d_j is a deterministic mean, $a \sim N(0, \sigma_a^2)$ is a fixed effect, x_j is a random walk component with innovation $e_j \sim N(0, \sigma_e^2)$. $(\sigma_a, \sigma_e) = (0.5, 0.09486833)$ are set to match the level and slope of the variance profile of log wages. Set d_j to match mean log wages by age from an age-cohort dummy variable regression using HPV's data.

4. An agent's Pareto weight is set equal to the agent's fixed effect $\gamma = a$. The preference parameter $\theta = 0.36$. Both choices follow STY.

5. To compute the inequality implications, we need to deal with corner solutions for labor. We eliminate agents in the calculation of variances if a model agent's allocation fails an hours-of-work condition or a wage condition. The hours-of-work condition is $Tl_j \geq 260$, where $T = 12 \times 360$ denotes annual discretionary time (i.e. time not used to sleep or in personal maintenance). The minimum wage condition is $w_j > \frac{1}{2} \times 3.50$, where 3.50 comes from the average real federal minimum wage of the 1980-2006 period. This parallels the treatment in the data underlying Figure 1.

A-5 Heterogeneous Wage Growth

Data: We employ panel data from the Panel Study of Income Dynamics provided by HPV (2010). Their dataset is designed to match the variable definitions and sample selection criteria of the CEX data we describe in the Appendix A-1. In particular, we use the hourly wage data from Sample C of HPV (2010) and we refer the reader to their paper for further details. We apply two sample selection criteria in addition to the sample selection criteria employed in HPV's Sample C. First, we restrict the sample to years 1967-1996, which is the time period where the PSID was deployed annually so our sample is an unbalanced panel with $T = 1996 - 1967 + 1 = 30$ calendar years. Second, we restrict the sample to agents that report a valid hourly wage in at least 2, 10 or 20 sample years. We conduct a separate estimation using each of the three resulting samples.

Estimation Method: We use a common two-stage estimator. In the first stage we remove the deterministic component, d_{jt}, by regressing log wages for each sample year on a fourth degree polynomial of age. In the second stage we apply an Equally-Weighted Method of Moments estimator to the regression residuals of the first stage. We employ two sets of moment conditions. The first set is standard in the literature and equates the $\frac{T \times (T+1)}{2}$ unique elements of the empirical variance-covariance matrix of log wages across each pair of calendar years in the sample with a model-implied counterpart of this matrix.[20] The second set of moments consists of J additional moment conditions that equate the empirical variance of log wages of each of the J age groups in the dataset with a model-implied counterpart of that variance. These additional moments improve the fit of the estimated process for variances by age for some of the samples we consider. For each of our estimations we conduct a quasi-global numerical minimization as follows. We set the initial parameter values to those in Guvenen (2009) Table 1, row 4. Then we conduct several rounds of minimization, where each round consists of two steps. In the first step we evaluate 100,000 random variations of the best parameter vector. In the second step we conduct a local search using the Nelder-Mead Simplex algorithm, starting from an initial simplex containing the best vector from the first step. The algorithm stops when the improvement in the objective function between two rounds becomes negligible.

Results: Our estimates are displayed in Table 4. The estimation reveals that there is a slope heterogeneity component in all of the samples. This component, however, is not measured with great precision as can be seen by examining the standard errors. We attribute these large standard errors to two sources: measurement error in hours of work and the more inclusive nature of our sample selection criteria. We include an interview if the reported hours of work are above 260 while Guvenen (2009) only includes observations with at least 520 hours. We do this for consistency with the CEX sample employed throughout our paper.

We set the remaining parameters of the model in order to best match the levels and slopes of the variance-covariance profiles in Figure 1. That is, we repeat the estimation of Section 3.3 with three modifications. First, the evolution of the variance-covariance matrix of exogenous variables is as described in Section 3.4. Second, we set the parameters governing the wage process to equal the values from the first column of Table 4. Third, we interpret the estimated variance of the transitory shock, σ_ϵ, as the sum of the true variance of the transitory shock and the variance of measurement error in log wages from Table 2. As before, we impose positive semidefiniteness of the variance-covariance matrix of exogenous variables at every age j. Parameter estimates are presented in Table 5 and the model fit is displayed in Figure 6.

[20]See the appendix in Guvenen (2009) for explicit formulas for the estimator and the standard errors.

Table 3: Parameter Values

Case 1: $v_{w\gamma} = 0$

Parameter	Cohort Effects	Time Effects
$\frac{1}{\phi}$	0.04259 (0.00441)	0.08045 (0.00666)
g_z	69.48949 (13.84440)	20.54349 (3.71192)
g_w	0.18857 (0.00193)	0.23660 (0.00097)
g_γ	66.33080 (13.55054)	18.64862 (3.57476)
g_{wz}	0.27498 (0.04419)	0.13028 (0.02175)
$g_{z\gamma}$	-67.51648 (13.69767)	-19.12144 (3.64326)
$g_{w\gamma}$	-0.13149 (0.04385)	0.05977 (0.02190)
v_z	0.01611 (0.00066)	0.00390 (0.00029)
v_{wz}	0.01097 (0.00023)	0.00794 (0.00010)
v_w	0.00883 (0.00012)	0.00585 (0.00007)

Case 2: $v_{wz} = 0$

Parameter	Cohort Effects	Time Effects
$\frac{1}{\phi}$	0.03630 (0.00047)	0.03546 (0.00007)
g_z	95.02965 (2.46448)	100.00000 (0.01471)
g_w	0.18832 (0.00192)	0.23632 (0.00097)
g_γ	91.35358 (2.43088)	96.12177 (0.02905)
g_{wz}	0.51794 (0.01866)	0.58503 (0.01098)
$g_{z\gamma}$	-92.79795 (2.44745)	-97.58627 (0.01499)
$g_{w\gamma}$	-0.37289 (0.01892)	-0.39353 (0.01083)
v_z	0.01611 (0.00066)	0.00390 (0.00029)
$v_{w\gamma}$	0.01087 (0.00023)	0.00784 (0.00010)
v_w	0.00884 (0.00012)	0.00587 (0.00007)

Note: Standard errors are reported in parentheses next to each parameter value. In addition to non-negativity constraints for variances and the positive semidefiniteness of variance-covariance matrices, we impose the constraints $\rho = 2.5$ and $g_z \leq 100$ in the estimation. Standard errors are based on the procedure described in Table 1.

Table 4: Estimated Parameters of the Wage Process

Param.	$n_i \geq 2$	$n_i \geq 10$	$n_i \geq 20$
χ	0.9420 (0.0069)	0.9489 (0.0095)	0.9646 (0.0204)
σ_ε^2	0.0771 (0.0016)	0.0714 (0.002)	0.0706 (0.0036)
σ_η^2	0.0233 (0.0011)	0.0218 (0.0013)	0.0149 (0.0013)
σ_α^2	0.0931 (0.0067)	0.0764 (0.0118)	0.0214 (0.0202)
σ_β^2	0.00000433 (0.0000306)	0.00001003 (0.0000463)	0.0001195 (0.0000834)
$\sigma_{\alpha,\beta}$	0.0006 (0.003)	0.0009 (0.0029)	0.0016 (0.0034)
Obs.	65,611	53,038	27,671

Note: Results are based on PSID data from HPV (2010). Each of the three columns corresponds to a separate estimation. Each column is based on a different sample. The samples differ in the number of valid log wage observations, n_i, required for an individual to be included in the sample.

Table 5: Parameter Values

Parameter	Cohort Effects	Time Effects
$\frac{1}{\phi}$	0.03537 (0.01720)	0.03586 (0.01320)
g_z	99.95758 (33.56984)	97.77617 (29.84539)
g_γ	96.18886 (32.73626)	93.93733 (29.35454)
$g_{\alpha\gamma}$	0.00000 (0.00000)	0.00000 (0.00491)
$r_{\gamma\beta}$	0.52930 (0.20875)	0.38736 (0.14590)
$g_{z\gamma}$	-97.67955 (33.15308)	-95.38213 (29.59429)
$g_{z\alpha}$	0.56585 (0.12805)	0.61561 (0.13187)
v_z	0.01611 (0.00067)	0.00390 (0.00069)

Note: Standard errors are reported in parentheses next to each parameter value and are based on the procedure used in Table 1. The reported standard errors do not account for the sampling variation of the wage process estimates. In addition to non-negativity constraints for variances and the positive semidefiniteness of variance-covariance matrices, we impose the constraints $\rho = 2.5$ and $g_z \leq 100$ in the estimation.