Consumption and Labor Supply with Partial Insurance: An Analytical Framework*

Jonathan Heathcote
Federal Reserve Bank of Minneapolis and CEPR
heathcote@minneapolisfed.org

Kjetil Storesletten
University of Oslo, and CEPR
kjstore@econ.uio.no

Giovanni L. Violante
New York University, CEPR, and NBER
glv2@nyu.edu

This draft: April 2013

Abstract
This paper develops a model with partial insurance against idiosyncratic wage shocks to quantify risk sharing, and to decompose inequality into life-cycle shocks versus initial heterogeneity in preferences and productivity. Closed-form solutions are obtained for equilibrium allocations and for moments of the joint distribution of consumption, hours, and wages. We prove identification and demonstrate how labor supply data can inform the study of risk sharing. The model, estimated with data from the CEX and the PSID over the period 1967-2006, implies that: (i) 39% of permanent wage shocks pass through to consumption; (ii) the share of wage risk insured increased until the early 1980s and remained stable thereafter; (iii) life-cycle productivity shocks account for half of the cross-sectional variance of wages and earnings, but for much less of dispersion in consumption or hours worked.

*Heathcote and Violante's research is supported by a grant from the National Science Foundation (SES 0418029). Storesletten thanks the Research Council of Norway and the Centre of Equality, Social Organization, and Performance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

The purpose of this paper is to measure the degree of risk sharing achieved by US households. Quantifying existing risk sharing is a prerequisite for evaluating the welfare consequences of adjusting social insurance programs, or changing the progressivity of the tax system.

One approach to studying risk sharing is to build a structural equilibrium model, and to use it as an artificial laboratory to study the response of consumption to individual income fluctuations. A prominent example is the standard incomplete-markets model, where households self-insure against income fluctuations by borrowing and lending via a risk-free bond.

However, there are many other ways households can smooth shocks and share risk, including flexible labor supply, progressive taxation, social insurance programs, within-family transfers, informal networks, and default or bankruptcy (see Heathcote, Storesletten, and Violante (2009) for a survey). A problem with the structural approach is that the total amount of risk sharing achieved in equilibrium will be sensitive to the details of which risk-sharing mechanisms are introduced and how they are modelled, casting doubt on whether any particular formulation comes close to replicating the amount of risk sharing households can achieve in practice (see, for example, Kaplan and Violante, 2010). Thus, Deaton (1997) has argued that a more fruitful approach is to try to quantify directly the overall degree of risk sharing in the economy, while remaining agnostic about the exact details on how households achieve this outcome.[1] One influential recent example of this less structural approach is Blundell, Pistaferri, and Preston (2008), who estimate the degree to which permanent changes in earnings transmit to consumption in the United States.

In this paper, we take a fully structural approach to measuring risk-sharing that is nonetheless designed to address the Deaton critique. We start with a standard incomplete-markets model, and explicitly introduce two important smoothing mechanisms against idiosyncratic wage fluctuations: elastic labor supply and progressive taxation. The model also

[1]Deaton (1997, pp. 372-374) writes: “Saving is only one of the ways people can protect their consumption against fluctuations in their income. An alternative is to rely on other people, to share risk with friends and kin, with neighbors, or with other anonymous participants through private or government insurance schemes, or through participation in financial markets ... [T]he very multiplicity of existing mechanisms makes it likely that there is at least partial insurance through financial or social institutions, and that such risk sharing adds to the possibilities for autarkic consumption smoothing through intertemporal transfers of money or goods ... Although it is possible to examine the mechanisms, the insurance contracts, tithes and transfers, their multiplicity makes it attractive to look directly at the magnitude that is supposed to be smoothed, namely consumption.”

allows for insurance against a subset of wage fluctuations, as a flexible way to capture the presence of additional risk-sharing mechanisms. Inspired by Deaton, our focus will be on letting the data identify the extent of this residual insurance, rather than on specifying the details of how it is achieved.

The key advantage of retaining a structural approach is that it allows us to integrate evidence on risk sharing from data on hours worked and consumption in a theoretically consistent way. Most of the risk sharing literature to date has focussed on exploring comovement between household income and consumption (see e.g., Jappelli and Pistaferri, 2010), but data on individual labor supply turn out to be very informative about insurance against idiosyncratic shocks. The logic is simply that individuals should adjust hours worked more strongly in response to insurable versus uninsurable wage fluctuations, reflecting the absence of offsetting wealth effects in the former case.

Relative to the existing theoretical literature, the key innovation is that the framework developed here allows for two different types of shocks to individual hourly wages that are distinguished by their degree of insurability. As in standard incomplete markets models, no explicit insurance exists for the first type: these "uninsurable" shocks can only be smoothed via adjustments to own hours worked, via borrowing and lending in a riskless bond, or via government redistribution through progressive taxation. In contrast, the second type of shock can be fully insured, as in complete markets models. One motivation for this "insurable" component is that in reality some changes in individual wages are perfectly forecastable by agents and hence easily smoothed. In addition, there are certain shocks which can be insured within the family or for which existing institutions provide explicit insurance, such as unemployment or disability shocks. Since some but not all shocks are explicitly insurable, this is an economy with *partial insurance.*

In the equilibrium of the model, agents choose to not use the bond to smooth the uninsurable shock. This result extends the logic in Constantinides and Duffie (1996) to a much richer environment. Thanks to this result, and in sharp contrast to the standard incomplete markets model, equilibrium allocations of consumption and hours worked can be expressed in closed form, as log-linear functions of the two idiosyncratic wage components and an idiosyncratic preference shifter (we allow for heterogeneity in the relative tastes for consumption versus work).

These closed-form log-linear allocations make it possible to compute and interpret cross-sectional variances and covariances of the joint equilibrium distribution of wages, hours, and consumption. We use information contained in both the "macro facts" on the distributions of these variables in *levels* that have motivated recent macroeconomic investigations (e.g., Attanasio and Davis, 1996; Krueger and Perri, 2006; Heathcote, Storesletten, and Violante 2010b), and the "micro facts" on the distributions in *growth rates* that have been the primary focus of labor economists (e.g., Abowd and Card, 1989; Blundell, Pistaferri, and Preston, 2008). The analytical expressions for these cross-sectional moments allow us to formally prove identification of all the model's parameters – something that is usually impossible in large scale structural models – under mild data requirements that are satisfied in standard micro data sets. In fact, we prove that the model is fully identified given only panel data on wages and hours worked, i.e., without *any* consumption data. In light of the recent studies questioning the quality of self-reported consumption expenditures in the US (e.g., Attanasio, Battistin, and Ichimura, 2007; Aguiar and Bils, 2011), it is valuable to be able to assess whether estimates of risk sharing derived from wage and hours data alone are consistent with those that also use consumption moments.

Our baseline estimation uses data on wages and hours from the Panel Study of Income Dynamics (PSID) over the period 1967-2006, and consumption data from the Consumer Expenditure Survey (CEX) over the period 1980-2006. The estimated model replicates well the evolution of the empirical cross-sectional distribution over wages, hours worked and consumption, both over time and over the life cycle.

We use the model to derive quantitative answers to three central questions concerning risk-sharing in the US economy: (1) how effectively can households smooth idiosyncratic wage fluctuations?, (2) how has the extent of risk sharing changed over the last four decades, a period of sharply rising wage inequality?, and (3) what is the role of life-cycle shocks and initial heterogeneity in determining cross-sectional dispersion in economic outcomes?

First, we ask how much individual wage risk can be smoothed, and what are the relative contributions to smoothing of explicit insurance, labor supply adjustments, and progressive taxation. Blundell, Pistaferri, and Preston (2008) argue that a natural way to quantify consumption smoothing is to measure how much of a typical permanent income shock passes through to consumption. Our model suggests that this pass-through coefficient from individual wages to household consumption is around 40%, or equivalently that 60% of permanent wage fluctuations are effectively smoothed. Where does this smoothing come from? Half of it stems from directly insurable shocks, one-third reflects progressive taxation, and the rest reflects adjustments to labor supply.

An alternative metric for consumption smoothing, common in the literature, is the ratio of the within-cohort change in the variance of log consumption to the corresponding change in the variance of log income (e.g., Blundell and Preston, 1998; Storesletten, Telmer, and Yaron, 2004a). We demonstrate that these two measures of pass-through coincide only when earnings taxation is proportional and labor supply is absent as smoothing channel for uninsurable shocks (e.g., zero Frisch elasticity or balanced growth preferences). Our model also indicates that, for plausible parameter estimates, the ratio-of-variances statistic is always smaller than the pass-through coefficient.

Second, we ask how risk sharing has changed over time. We find that US households were effectively able to insure two thirds of the sharp increase in wage inequality over the past 40 years. In 1967 the insurable component of wages accounted for around 25% of the cross-sectional variance of log wages, while by the early 1980s this fraction had risen to around 45%. Since then, the variances of the insurable and uninsurable components of wages have risen at a similar rate, leaving the fraction of wage fluctuations insured relatively stable. Data on hours worked are an essential input for these estimates, since no consumption data is available prior to 1980, and it is the observed increase in the covariance between wages and hours that indicates an increase in the degree of risk sharing. Reassuringly, after 1980, we obtain very similar estimates for the relative importance of insurable and uninsurable shocks regardless of whether we use all available data, including consumption, or just data on earnings and hours worked.

Third, we use the estimated model to decompose inequality in the cross section into components reflecting life-cycle shocks versus initial heterogeneity in productivity and the disutility of work effort. This decomposition is unique and additive in our framework. Roughly half of the total cross-sectional variance in earnings reflects life-cycle shocks to productivity. In contrast, these shocks account for less than 20% of the cross-sectional variances of consumption and hours worked. Net of measurement error, the most important source of dispersion in consumption is initial heterogeneity in productivity. For hours worked, in contrast, it is initial heterogeneity in preferences.

The rest of the paper is organized as follows. Section 2 develops our framework, derives the equilibrium allocations, and explains how we obtain tractability. In Section 3, we derive closed-form expressions for the equilibrium cross-sectional moments. Section 4 proves how these moments allow us to identify all the structural parameters of the model, and describes the data and estimation algorithm. Section 5 lays out the results of the quantitative analysis.

Section 6 concludes.

2 Model economy

We first describe the model formally. Next, we discuss the key assumptions in detail.

Demographics We adopt the Yaari perpetual youth model: agents are born at age zero and survive from age a to age $a+1$ with constant probability $\delta < 1$. A new generation with mass $(1-\delta)$ enters the economy at each date t. Thus, the measure of agents of age a is $(1-\delta)\delta^a$, and the total population size is unity.

Preferences Lifetime utility for an agent born (i.e., entering the labor market) in cohort birth year b is given by

$$\mathbb{E}_b \sum_{t=b}^{\infty} (\beta\delta)^{t-b} u(c_t, h_t; \varphi), \tag{1}$$

where the expectation is taken over sequences of shocks defined below. Here c_t denotes consumption at date t for an agent of age $a = t-b$, while h_t is the corresponding value for hours worked. Agents discount the future at rate $\beta\delta$, where $\beta < 1$ is the pure discount factor. Period utility is

$$u(c_t, h_t; \varphi) = \frac{c_t^{1-\gamma} - 1}{1-\gamma} - \exp(\varphi) \frac{h_t^{1+\sigma}}{1+\sigma}. \tag{2}$$

The parameter γ is the inverse of the intertemporal elasticity of substitution for consumption, and σ governs the elasticity of labor supply.[2] The preference weight φ captures the strength of an individual's aversion to work.[3] The distribution of φ for the cohort with birth year t is denoted $F_{\varphi t}$, with cohort-specific variance $v_{\varphi t}$. We incorporate preference heterogeneity because, as we will show, it is important for explaining the observed cross-sectional joint distribution over wages, hours, and consumption.[4] In Section 2.3.2 we discuss how our results extend to alternative preference specifications.

[2]The parameter γ is also related to risk aversion. In particular, the coefficient of relative risk aversion is $1/(1/\gamma + 1/\sigma)$ (see Swanson, 2012). As we explain below, the most important role of γ in our model is that it determines the relative strength of income and substitution effects on hours worked.

[3]Note that preferences are defined over total hours per period, and model agents are implicitly indifferent between alternative ways to allocate hours with a period. Thus, the model cannot address the question of how total annual hours should be divided between hours worked per day (e.g., overtime), days worked per week (part-time work), and weeks worked by year (non-employment).

[4]It has long been recognized that a sizeable fraction of cross-sectional dispersion in hours worked is unrelated to dispersion in wages (e.g., Abowd and Card, 1989).

Idiosyncratic risk The population in the economy is partitioned into groups that we will refer to as "islands," where each island contains a continuum of individual agents. Agents face labor productivity shocks at the individual level, which are uncorrelated across members of each island, and shocks at the island level, which are common to all members of a given island, but uncorrelated across islands. Individual labor productivity w is given (in logs) by the sum of the island-level component, denoted α, and the (orthogonal) individual-level component, denoted ε:

$$\log w_t = \alpha_t + \varepsilon_t. \tag{3}$$

The market structure outlined below will assume differential trading opportunities between versus within islands, translating into differential insurance against shocks to α versus ε.

The island-level component α follows a random walk:

$$\alpha_t = \alpha_{t-1} + \omega_t,$$

where the innovation ω is drawn from the distribution $F_{\omega t}$ with variance $v_{\omega t}$ at time t. The individual-level component ε is itself the sum of two orthogonal random variables:

$$\varepsilon_t = \kappa_t + \theta_t.$$

Here θ is a transitory (independently distributed over time) shock drawn from $F_{\theta t}$ with variance $v_{\theta t}$, while κ is a permanent component that follows a second unit root process:

$$\kappa_t = \kappa_{t-1} + \eta_t,$$

where the innovation η is drawn from the distribution $F_{\eta t}$ with variance $v_{\eta t}$.[5]

Agents who enter the labor market at age $a = 0$ in year t draw initial realizations α^0 and κ^0 from distributions $F_{\alpha^0 t}$ and $F_{\kappa^0 t}$, with cohort-specific variances $v_{\alpha^0 t}$ and $v_{\kappa^0 t}$. The initial draws φ, α^0, and κ^0 are assumed to be uncorrelated.[6]

[5]The assumed statistical process for individual efficiency units – unit root plus independently distributed shocks – has a long tradition in the literature that estimates statistical models for individual wage dynamics (see, e.g., MaCurdy, 1982). The empirical autocovariance function for individual wages displays a sharp decline at the first lag, indicating the presence of a transitory component in wages. At the same time, within-cohort wage dispersion increases approximately linearly with age, suggesting the presence of permanent shocks.

[6]The initial draws (φ, α^0) could in principle be correlated if, for example, wages at labor market entry are a function of schooling, and schooling depends on the preference weight, φ. In a previous version of this paper, we allowed for correlation between α^0 and φ. The model was still tractable, but the estimated correlation coefficient was insignificantly different from zero.

A law of large numbers (e.g., Uhlig, 1996) can be applied twice so that individual-level ε shocks wash out within an island, and island-level α shocks induce no aggregate uncertainty in the economy as a whole (see Attanasio and Ríos-Rull (2000) for a similar structure).

Production Production of the final consumption good takes place through a constant returns to scale technology with labor as the only input. The economy-wide good and labor markets are frictionless and perfectly competitive. Hence, individual wages equal individual productivities (units of effective labor per hour worked).

Taxes and redistribution The government operates a progressive tax system that provides public insurance. Following Benabou (2002), an individual with a gross labor income $y_t = w_t h_t$ receives post-government earnings $\tilde{y}_t$ given by

$$\tilde{y}_t = \lambda \left(y_t\right)^{1-\tau}. \tag{4}$$

The fiscal parameters λ and τ are assumed constant over time. Loosely speaking, λ defines the level of taxation, while $\tau \geq 0$ defines the rate of progressivity built into the tax system. To see this, note that $\log(\tilde{y}_t) = \log(\lambda) + (1-\tau)\log(y_t)$, and thus $(1-\tau)$ defines the elasticity of after-tax earnings to pre-tax earnings. For $\tau = 0$ the system implies a flat tax $1-\lambda$ on labor income, while for $\tau > 0$ the tax system is progressive. The government uses aggregate net tax revenue to finance a non-valued public consumption good G_t, which adjusts to balance the government budget on a period-by-period basis. While this model of taxation is simple, it is sufficiently flexible to offer a reasonable approximation to the actual US tax system (see Section 4.3).

Market structure All assets in the economy are in zero net supply, and asset markets are competitive. At birth, each agent is endowed with zero financial wealth.[7] Individuals born in year b draw values for α^0 and φ before any markets open. They are then allocated to an island, which is defined by an *ex ante* unknown sequence $\{\omega_t\}_{t=b+1}^{\infty}$ that will apply to all island members. Within each island, agents trade a complete set of insurance contracts. In particular, in every period $t \geq b$, agents can purchase contracts indexed to $s_{t+1} = (\omega_{t+1}, \eta_{t+1}, \theta_{t+1})$.[8] Scope for insurance across islands is more limited: agents can only trade insurance contracts indexed to their individual-level shocks $(\eta_{t+1}, \theta_{t+1})$, but inter-island contracts contingent on the realization of the island-level shock ω_{t+1} are ruled out.

[7]It is straightforward to relax the assumption of zero initial *individual* financial wealth. The key requirement, as will become clear below, is that *average* initial wealth on each island is zero.

[8]New labor market entrants at date b can also purchase contracts indexed to (κ^0, θ_b).

Insurance contracts incorporate mortality risk: if an agent purchases one unit of insurance against any state s_{t+1}, the contract pays δ^{-1} units of consumption if the agent survives to the next period and s_{t+1} is realized, and 0 units otherwise.

Note that agents can effectively trade risk-free bonds freely within or across islands. In particular, purchasing δ units of insurance for every possible realization of the pair $(\eta_{t+1}, \theta_{t+1})$ delivers one unit of consumption risk-free in the next period.

Information Agents are assumed to take as given the sequences of distributions $\{F_{\varphi t}, F_{\alpha^0 t}, F_{\kappa^0 t}, F_{\omega t}, F_{\eta t}, F_{\theta t}\}$. Thus they have perfect foresight over future wage distributions.[9]

2.1 Agent's problem

Let $s^t = \{s_b, s_{b+1}, ..., s_t\}$ denote the individual history of the shocks for an agent from birth year b up to date t, where

$$s_j = \begin{cases} (b, \varphi, \alpha^0, \kappa^0, \theta_b) & \in \mathbb{S}_b = \mathbb{N} \times \mathbb{R}^4 \quad \text{for } j = b \\ (\omega_j, \eta_j, \theta_j) & \in \mathbb{S} = \mathbb{R}^3 \quad \text{for } j > b \end{cases} \tag{5}$$

with $s^t \in \mathbb{S}_b \times \mathbb{S}^{t-b}$.

Let $Q_t(S; s^t)$ denote the price of insurance claims purchased at date t from local (within-island) insurers by an agent with history s^t that deliver one unit of consumption at $t+1$ if and only $s_{t+1} \in S \subseteq \mathbb{S}$. Let $B_t(s_{t+1}; s^t)$ denote the quantity of the claim purchased that pays in individual state s_{t+1}. Recall that insurers can also offer contracts indexed to $(\eta_{t+1}, \theta_{t+1})$ to agents in other islands. Define $z_{t+1} \equiv (\eta_{t+1}, \theta_{t+1})$ where $z_{t+1} \in Z \subseteq \mathbb{Z} = \mathbb{R}^2$. Let $Q_t^*(Z; s^t)$ denote the price of insurance claims purchased at date t from outside (between-island) insurers by an agent with history s^t that deliver one unit of consumption at $t+1$ if and only $z_{t+1} \in Z$. Let $B_t^*(z_{t+1}; s^t)$ denote the quantity of the claim purchased from outside insurers that pays upon the realization z_{t+1}. The agent's budget constraint is given by

$$\begin{aligned} \lambda \left[w_t\left(s^t\right) h_t\left(s^t\right)\right]^{1-\tau} + d_t\left(s^t\right) \quad &= \quad c_t\left(s^t\right) + \int_{\mathbb{S}} Q_t\left(s_{t+1}; s^t\right) \, B_t\left(s_{t+1}; s^t\right) \, ds_{t+1} \\ &\quad + \int_{\mathbb{Z}} Q_t^*\left(z_{t+1}; s^t\right) \, B_t^*\left(z_{t+1}; s^t\right) \, dz_{t+1}, \end{aligned} \tag{6}$$

where realized wealth at node $s^t = (s^{t-1}, s_t)$ is given by

$$d_t(s^t) = \delta^{-1} \left[B_{t-1}(s_t; s^{t-1}) + B_{t-1}^*\left(z_t; s^{t-1}\right)\right].$$

[9] Alternatively, one could assume that the variances of these distributions themselves follow some stochastic process. The expression for the equilibrium interest rate would be affected, but equilibrium allocations would remain identical to those described below.

The problem for an agent entering the labor market at date b is to maximize (1) subject to a sequence of budget constraints of the form (6), and the wage process. In addition, agents face limits on borrowing that rule out Ponzi schemes, and non-negativity constraints on consumption and hours worked.

2.2 Competitive equilibrium

Given sequences $\{F_{\varphi t}, F_{\alpha^0 t}, F_{\kappa^0 t}; F_{\omega t}, F_{\eta t}, F_{\theta t}\}$, a competitive equilibrium is a set of allocations $\{c_t(s^t), h_t(s^t), d_t(s^t), B_t(\cdot; s^t), B_t^*(\cdot; s^t)\}$ and prices $\{Q_t(S; s^t), Q_t^*(Z; s^t)\}$ for all dates t, all histories $s^t \in \mathbb{S}_b \times \mathbb{S}^{t-b}$, and all $S \subseteq \mathbb{S}$, $Z \subseteq \mathbb{Z}$ such that (i) allocations maximize expected lifetime utility, (ii) insurance markets clear, and (iii) the economy-wide markets for the final good and labor services clear.

Proposition 1 [COMPETITIVE EQUILIBRIUM] *There exists a competitive equilibrium characterized as follows:*

(i) *There is no insurance trade between islands:* $B_t^*(Z; s^t) = 0$ *for all* Z *and all* s^t.

(ii) *Consumption and hours are given by*

$$\log c_t\left(s^t\right) = -\left(1-\tau\right)\widehat{\varphi} + \left(1-\tau\right)\left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)\alpha_t + \mathcal{C}_t^a \quad (7)$$

$$\log h_t\left(s^t\right) = -\widehat{\varphi} + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)\alpha_t + \frac{1}{\widehat{\sigma}}\varepsilon_t + \mathcal{H}_t^a, \quad (8)$$

where $a = t - b$ *is the age of the individual,* $\mathcal{C}_t^a$ *and* $\mathcal{H}_t^a$ *are age and date-specific constants (see Appendix A.1),* $1/\widehat{\sigma} \equiv \left(1-\tau\right)/\left(\sigma+\tau\right)$ *is a tax-modified Frisch elasticity, and* $\widehat{\varphi} \equiv \varphi/\left(\sigma+\gamma+\tau(1-\gamma)\right)$ *is a rescaled preference weight.*

(iii) *The prices of insurance claims are given by*

$$Q_t\left(S; s^t\right) = Q_t\left(S\right) = \beta \exp\left(-\gamma\,\Delta\mathcal{C}_{t+1}\right)\int_S \exp\left(-\gamma(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\omega_{t+1}\right) dF_{s,t+1} \quad (9)$$

$$Q_t^*\left(Z; s^t\right) = Q_t^*\left(Z\right) = \Pr\left(\left(\eta_{t+1}, \theta_{t+1}\right) \in Z\right) \times Q_t(\mathbb{S}),$$

where F_{st} *is the joint distribution over* (ω, η, θ) *at date* t, $Q_t(\mathbb{S})$ *is the price of a risk-free bond, and* $\Delta\mathcal{C}_{t+1} \equiv \mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a$ *is independent of age.*

Proof. See Appendix A.1. ■

Part (i) of Proposition 1 says that there is an equilibrium in which all trade takes place within islands. This result implies zero insurance against the α_t component of idiosyncratic wage risk, because shocks to α_t are common to all members of an island. In particular, there is no self-insurance, via non-contingent borrowing and lending, against these shocks. In contrast, there is perfect insurance, by assumption, against shocks to ε_t. Thus, in this equilibrium, there is a sharp dichotomy between one type of risk which is uninsured, and another that is fully insured. In what follows, we will use the label "uninsurable" to denote the ω shock and the initial draws α^0 and φ, and the label "insurable" to denote the (η, θ) shocks and the initial draw κ^0. When the variance of insurable shocks is zero, equilibrium allocations correspond to autarky. When the variance of uninsurable shocks is zero, there is complete insurance against idiosyncratic risk. In the general case, when both types of shocks have positive variance, insurance is *partial.*

Part (ii) characterizes equilibrium allocations for consumption and hours worked in closed form. These expressions indicate that the vector of cumulated values for the shocks $(\alpha_t, \varepsilon_t)$ together with φ and age a contain sufficient information to fully describe an individual's equilibrium choices at node s^t. The power of this result lies in the fact that these are all exogenous states. Crucially, individual wealth is a redundant state variable, in the sense that it is also only a function of $(a, \varphi, \alpha_t, \varepsilon_t)$. The expression for wealth d_t is in Appendix A.1. Note that no distributional assumptions for wage shocks or preference heterogeneity are required to deliver these functional forms for equilibrium allocations.[10]

Part (iii) describes the insurance prices supporting this equilibrium. The key result is that the prices of insurance contracts on the inter-island market are actuarially fair, in the sense that they are equal to event-specific probabilities times the risk-free bond price $Q_t(\mathbb{S})$ – the price at which all agents are indifferent between borrowing and lending on the margin. At these prices, agents have no incentive to buy insurance from or sell insurance to agents on other islands, thereby supporting the no-trade result in part (i).

The logic of the proof for Proposition 1 is as follows. We first guess that all insurance claims are traded within island and that there is no insurance trade between islands. Hence aggregate island-level net savings is zero on each island. Because insurance mar-

[10]The distributions only affect the separable constants $\mathcal{C}_t^a$ and $\mathcal{H}_t^a$. We implicitly assume that the distributions imply finite values for these constants. The absence of an explicit solution for $\mathcal{C}_t^a$ and $\mathcal{H}_t^a$ is no obstacle for the empirical analysis, since the constants can be modeled through age and time dummies in individual consumption and hours observations.

kets are complete within an island, we can solve for the island-specific allocations via a simple static equal-weight planner's problem.[11] We can use planner problems to solve for within-island allocations, notwithstanding the presence of progressive distortionary taxation at the economy-wide level, because each island planner controls a measure zero of aggregate resources and therefore takes the tax function as exogenous. With expressions for consumption and hours worked in hand, we use the agent's intertemporal first-order condition to compute the implied (potentially island-specific) insurance prices. Finally, we verify that agents on every island assign the same value to any insurance contract that can be traded, and thus that there are no gains from inter-island trade.

Interpreting equilibrium allocations The impact of the preference parameter φ on hours and consumption is readily interpreted: a stronger relative distaste for work (higher φ) reduces labor supply, which transmits to earnings and consumption.

Hours worked are increasing in the insurable component $\varepsilon_t = \kappa_t + \theta_t$, and the response of hours to shocks to ε_t is defined by the tax-modified Frisch elasticity, $1/\widehat{\sigma} \equiv (1-\tau)/(\sigma+\tau)$. Progressive taxation ($\tau > 0$) lowers the tax-modified Frisch elasticity because it reduces the return to increasing hours worked in response to a rise in pre-tax wages. While full insurance with respect to ε_t rules out any income effect on hours worked, uninsurable permanent shocks to α_t do have an income effect which is regulated by γ. If $\gamma > 1$, the income effect dominates the substitution effect, and hours worked decline in response to an increase in α_t. If $\gamma < 1$, the substitution effect dominates and hours increase.

Individual consumption is independent of ε_t, since these shocks are fully insured and utility is separable between consumption and hours. The response of consumption to uninsurable wage shocks depends on the response of hours worked and the progressivity of taxation. Stronger income effects (larger γ) reduce the pass-through from wage shocks to consumption, as does more progressive taxation (larger τ). Note that the expression for individual consumption is not what the permanent income hypothesis would imply. Consumption does follow a random walk, but some permanent shocks (innovations η_t) are insured and thus do not affect consumption. In other words, consumption in our model exhibits "excess smoothness" (as originally defined by Campbell and Deaton, 1989). It is precisely this feature of the data that has motivated a large amount of recent research aimed at developing "par-

[11]Within-island allocations can be determined using equal-weight island-level planning problems because we defined an island as a group of agents with the same birth date b, common initial conditions (φ, α^0), and a common sequence $\{\omega_s\}_{s=b+1}^{\infty}$.

tial insurance" models that lie in between the bond economy and complete markets (e.g., Krueger and Perri, 2006; Ales and Maziero, 2009; Attanasio and Pavoni, 2011;).

2.3 Tractability of the framework

With few exceptions, incomplete markets models do not admit an analytical solution and numerical methods are required to solve for equilibrium allocations.[12] In this section we explain how we retain tractability, and we relate this result to the existing literature. Readers who are more interested in the empirical application can skip directly to Section 3.

2.3.1 How we retain tractability

There are two keys to tractability in our framework: (i) individual wealth is a redundant state variable, and (ii) agents have access to perfect insurance against some shocks and no explicit insurance against others. To achieve this insurance dichotomy as an equilibrium outcome, the island-economy structure is important.

Why wealth is a redundant state The reason individual wealth is a redundant state variable is twofold. First, even though the within-island equilibrium wealth distribution is non-degenerate, allocations can be characterized without reference to it: full insurance within the island implies that within-island allocations can be derived by solving an island-level planner problem with an equal-weight welfare function corresponding to common initial asset positions for all agents, subject to an island-level resource constraint.

Second, the inter-island wealth distribution does not show up in allocations because, in equilibrium, this distribution remains degenerate at zero. This second argument can be explained in three simple steps. To understand why there is no asset trade between islands, it is sufficient to understand why there is no trade in a risk-free bond.[13] Let $r_{t+1} = -\log Q_t(\mathbb{S})$

[12]In standard (intractable) incomplete markets models, decision rules depend on wealth, and the distribution of wealth is endogenous and must be solved for numerically. The literature has followed three alternative routes to avoid this outcome. The first is to assume a statistical model for income risk (permanent, multiplicative shocks) such that the equilibrium wealth distribution remains degenerate at zero (Constantinides and Duffie, 1996). The second is to assume a preference specification – quadratic or in the constant absolute risk aversion (CARA) class – such that the precautionary motive for saving is either zero or independent of wealth (Caballero, 1990). The third is to allow agents to control the amount of idiosyncratic risk that they face such that equilibrium exposure to idiosyncratic risk is proportional to wealth, given CRRA preferences (Krebs, 2003; Angeletos, 2007). Krebs (2003) allows for human capital accumulation, so that agents can control the composition between (safe) physical and (risky) human wealth independently of total wealth by making savings choices in both assets. Angeletos (2007) models idiosyncratic risk to entrepreneurial business income rather than labor income. In his model agents control portfolio exposure to idiosyncratic risk by adjusting the quantity of entrepreneurial capital in total savings.

[13]Recall that inter-island insurance prices are simply event-specific probabilities times the bond price.

denote the equilibrium interest rate and $\rho = -\log \beta$ the discount rate. In the model, individuals have three saving motives: an intertemporal motive proportional to the gap between r_{t+1} and ρ, a smoothing motive linked to expected earnings growth over the life cycle, and a precautionary motive that is a function of the variance of uninsurable island-level shocks $v_{\omega,t+1}$. Importantly, each of these three factors applies with the same force on all islands. The strength of the intertemporal motive is given by the term $(r_{t+1} - \rho)/\gamma$, common across agents. All islands have the same smoothing motive, because island-level expected earnings growth is independent of age and of the current wage. The precautionary motive is the same because all agents face the same variance for the uninsurable component of wages. Consequently, there exists an economy-wide interest rate r_{t+1} at which, in equilibrium, the (negative) intertemporal motive exactly offsets the (negative) smoothing motive and the (positive) precautionary motive, and no agent wants to either borrow or lend across islands.

To gain more intuition, it is useful to make a specific distributional assumption. If each variable $x_t \in (\omega_t, \eta_t, \theta_t)$ is distributed Normally, $x_t \sim N(-v_{xt}/2, v_{xt})$, then asset prices can be derived in closed form. Focusing, for simplicity, on the special case $\sigma \to \infty$ (inelastic labor supply) and $\tau = 0$ (proportional taxation), we have

$$\frac{r_{t+1} - \rho}{\gamma} + (1 + \gamma) \frac{v_{\omega,t+1}}{2} = 0. \tag{10}$$

The first term measures the intertemporal motive to dis-save. The second term, capturing the precautionary motive for saving, is equal to half the variance of the island-level productivity shocks times the coefficient of relative prudence, $(1 + \gamma)$. The equilibrium interest rate is such that the two saving motives exactly offset.[14]

Insurance dichotomy Our model of risk and insurance (two types of shocks, one insurable and one uninsurable) stands in contrast to the standard approach (e.g., Huggett, 1994), in which there is a single shock to wages that can be partially smoothed. Our model is tractable, while the standard model is not. But which structure is most empirically relevant? The sharp insurability dichotomy in our model is certainly extreme, but it is broadly consistent with the idea that some wage changes are much more insurable than others. For example, as Low, Meghir, and Pistaferri (2010) emphasize, insurance against job loss and

[14]See eq. (A5) in the Technical Appendix A for the interest rate expression with σ finite and $\tau \neq 0$. If $\gamma > 1$, then hours respond negatively to uninsurable shocks (see eq. 8). In this case, a higher Frisch elasticity reduces the precautionary saving motive, since labor supply provides a useful hedge against risk. Tax progressivity ($\tau > 0$) reduces the precautionary saving motive.

severe health deterioration exists through explicit institutional arrangements, such as unemployment compensation and disability insurance. In addition, one should expect individuals to perfectly smooth forecastable wage changes, such as automatic raises linked to tenure. In contrast, no explicit insurance exists against many other shocks – such as unanticipated wage drops linked to long-lasting reductions in the demand for specific skills or occupations.

Note that while our description of the environment assumes that (i) all individual insurance arises from explicit markets and state-contingent financial income flows, and (ii) wage growth is unpredictable, one could generalize both assumptions. The same allocations for consumption and hours worked can be supported through a combination of non-market mechanisms, including public insurance programs, within-family state-contingent transfers, and spousal labor supply. Moreover, if agents could perfectly foresee future innovations (η_t, θ_t) , then trade in a non-contingent bond would suffice to allow them to perfectly smooth consumption in response to these wage changes. We use the label "insurable shocks" as a catchall for both insurable (through market and non-market mechanisms) and forecastable wage changes.[15] We will let the data discipline the overall amount of insurance individuals have access to, over and above progressive taxation and own labor supply, without digging further into its precise origins.

Island structure The island configuration allows to achieve the equilibrium outcome in which some shocks are perfectly insured while others remain uninsured. Because unrestricted contracts are only exchanged within the island, this partition prevents agents from pooling the island-level risk.[16]

The sorts of insurance contracts that be traded within and between islands are specified exogenously. Exploring whether differential information frictions are a viable microfoundation for this differential availability of insurance is of some importance, but it goes beyond the scope of this paper. As a starting point, one may assume that within-island information about shocks and insurance contracts is perfect, but that neither individual shocks nor individual insurance arrangements can be observed across islands (as in Cole and Kocherlakota, 2001, and Ales and Maziero, 2011). The first assumption allows for full insurance within islands. The second may make it impossible to improve insurance of island-level

[15] Cunha, Heckman, and Navarro (2005) and Guvenen and Smith (2010), among others, explain the difficulty in distinguishing, empirically, between insurable shocks and predictable changes to income.

[16] A similar modelling design is common in international economics, where perfect insurance is often assumed against idiosyncratic risk within a country, while only a bond can be traded internationally to smooth country-level shocks (see, for example, Baxter and Crucini, 1995).

shocks beyond what can be achieved through trade in a risk-free bond.

Finally, the reader might wonder what the empirical counterpart of an island is. For expositional simplicity, we have assumed that households are permanently assigned to an island and, therefore, always trade insurance contracts within the same set of agents, all of whom experience a common sequence of ω_t shocks. Under this implementation of the island structure, an island comprises of households whose consumption comoves closely over long periods of time. One particularly appealing empirical counterpart to a model island would then be a network of family members. Under such an interpretation, idiosyncratic risks within the family (model ε_t) would be perfectly pooled, while any common component to family wages (model α_t) would remain uninsured. Such a common component arises naturally if family members are concentrated within regions, occupations, or skill-levels and are therefore unable to diversify region- occupation- or skill-specific shocks.[17] However, it is important to note that identical equilibrium allocations arise under an alternative implementation of the island structure, according to which a risk-sharing group at date t is defined only by a common ω_{t+1} instead of a common sequence $\{\omega_{t+1}, \omega_{t+2}, ...\}$.[18] Under this implementation, the theory puts many fewer restrictions that can be tested empirically: an island is just a group of agents pooling a subset of idiosyncratic shocks at a point in time, whose consumption need not be correlated in the long run. In the special case in which insurable shocks are i.i.d. over time, the island structure can be dispensed with altogether.[19]

As we show in Sections 3 and 4, for identification and estimation of the model, it is enough to use economy-wide cross-sectional moments. Because these moments aggregate dispersion within and between groups, we do not need to determine empirical counterparts to model islands.

2.3.2 Relation to Constantinides and Duffie (1996)

Constantinides and Duffie (1996), henceforth CD, is an important forebear of our model. The key insight of CD is that a no-trade equilibrium exists when: (1) the exogenous process for disposable income is a multiplicative unit root with innovations drawn from a distribution

[17] Angelucci et al. (2012) provide some empirical evidence consistent with this view.

[18] To see this, note that our decentralization assumes trade in insurance contracts indexed only to one period ahead realizations for $(\omega_{t+1}, \eta_{t+1}, \theta_{t+1})$. Moreover, the only important restriction on the pattern of trade is that the set of agents trading these contracts will all draw the same (unknown) ω_{t+1} innovation.

[19] In particular, if $\kappa_t = 0$ for all t so that $\varepsilon_t = \theta_t$, then an alternative way to implement the equilibrium allocations described in the text is to assume that agents first observe the innovation ω_t, and then trade – economy-wide – insurance claims contingent only on the realization of the transitory component θ_t.

common to all agents, (2) preferences are in the power utility class, (3) assets are in zero net supply, and agents are endowed with zero initial wealth.[20] We extend CD's environment in four dimensions that are important for a quantitative study of risk sharing.

First, our primitive exogenous stochastic process is over hourly wages and also includes a transitory component beyond the unit root. Gross earnings are endogenous since individuals control their labor supply. Showing that the no-trade result extends to preferences defined over labor supply as well as consumption is important because, as will become clear shortly, data on hours worked are a rich source of information on the nature of risk and risk sharing. In Heathcote, Storesletten, and Violante (2011b) we generalize the preference class under which the no-trade result holds beyond our baseline specification (2). We provide a simple static sufficient condition that can be used to check whether there exists an equilibrium with no inter-island trade, for any particular utility function defined over consumption and hours worked. We use this condition to show that the no-trade result extends to Greenwood-Hercowitz-Huffman, Cobb-Douglas, and recursive Epstein-Zin preferences. These alternative specifications also deliver closed-form expressions for equilibrium allocations.

Second, we allow for progressive taxation, which allows us to quantify the role of the tax system in consumption smoothing.

Third, agents in our model differ with respect to preferences, in addition to productivity. This feature is important because we do not want to impose *a priori* that the entire cross-sectional dispersion in consumption and hours worked is driven by dispersion in wages.

Finally, and most importantly, in our economy some risks are insurable within islands, so our version of the no-trade result applies across groups rather than across individuals. Hence, our model allows for partial consumption insurance against disposable earnings shocks – a critical requirement for bringing the model to the data successfully (as shown by Blundell, Preston, and Pistaferri, 2008). In contrast, the most direct interpretation of the CD model is that theirs is a world with no risk sharing in which each individual consumes his or her endowment. An alternative interpretation is that their postulated endowment process is "post-trade" and incorporates non-modeled risk-sharing mechanisms against fundamental shocks. Relative to this alternative interpretation, the advantage of our setup is that we explicitly model and quantify the riks-sharing channels available to households: labor sup-

[20]Both CD's model and ours can have assets in positive net supply in a trivial case, namely when agents are endowed at birth with a unit of the market portfolio and pay a lump-sum tax each period equal to the dividend on the market portfolio each period. In equilibrium, agents never trade away from their initial holding of the market portfolio, rendering the allocations (7)-(8) unchanged.

ply (from wages to earnings), progressive taxation (from pre- to after-tax earnings), and additional insurance (from after-tax earnings to consumption).

3 Cross-sectional implications

The model has thus far abstracted from variation in household composition, while actual households in the data vary with respect to household size and the number of potential workers. Moreover, measurement error is pervasive in micro data. In this section, we first describe how to augment our theoretical allocations to address these two issues. Next, we use these augmented theoretical allocations to derive, and interpret, closed-form expressions for (co-)variances of the equilibrium cross-sectional joint distribution of consumption, hours, and wages – the key moments used for model identification and estimation.

3.1 Augmented theoretical allocations

Modeling household composition To address the first issue, we generalize the model to explicitly incorporate variation in household size. This extension delivers a theoretically coherent approach for controlling for household composition in the data.

Let g and k denote the number of adults (grown-ups) and children (kids) in a particular household. All members of a given household reside on the same island. Let $e\left(g,k\right)$ be a function that defines the economies of scale enjoyed by a household of type (g,k) such that effective per-person consumption is given by household consumption c divided by $e(g,k)$, where $e(1,0)$ is normalized to unity. Children receive no weight in household utility. Thus period utility for a household of type (φ,g,k) is given by

$$u\left(c,\{h_i\}_{i=1}^{g};\varphi,g,k\right)=\frac{g}{1-\gamma}\left(\frac{c}{e\left(g,k\right)}\right)^{1-\gamma}-\frac{\exp\left(\varphi\right)}{1+\sigma}\sum_{i=1}^{g}h_i^{1+\sigma}. \tag{11}$$

One could make alternative assumptions regarding whether agents can insure *ex ante* against the type (g,k) of household to which they are allocated. In Appendix A.2, we solve for allocations in the two polar cases where there is full insurance and no insurance against (g,k), respectively. The key difference between the two models is that the full insurance model implies that hours worked should be independent of household composition, while the no-insurance model implies that hours should vary systematically with household size (when $\gamma\neq 1$). The reason household type does not affect equilibrium hours in the insurable household composition model is that household type has no impact on productivity or the disutility of labor effort, and thus it would be inefficient for individuals in different-size households to work different numbers of hours.

Motivated by this distinction, we experimented with regressing log hours on household composition dummies. Conditional on annual hours being positive, household composition explains essentially none of the observed variation in hours worked on the intensive margin, which is evidence in favor of the insurable model of household composition.

In Appendix A.2 we show that with full insurance against household composition, total consumption is given by

$$\log c_t^a \left(s^t; g, k\right) = \log c_t^a \left(s^t; 1, 0\right) + D\left(g, k\right),$$

where $\log c_t^a\left(s^t; 1, 0\right)$, consumption for a single-adult household, is given by equation (7), and $D(g,k)$ is given by

$$D\left(g, k\right) \equiv \frac{1}{\gamma} \log g - \left(\frac{1-\gamma}{\gamma}\right) \log e\left(g, k\right). \tag{12}$$

From this expression it is clear that if $\gamma = 1$ or $e(g,k) = g$, then households are allocated consumption exactly in proportion to the number of adults g, so there are no transfers between households of different size. Suppose there are economies of scale from additional adults (so that $e(g,0) < g$ for $g > 1$). Then larger households are allocated less consumption per adult than smaller households if and only if $\gamma > 1$. On the one hand, economies of scale make it inexpensive to increase effective consumption $c/e(g,k)$ for large households — in the limit $\gamma \to 0$ this effect makes it efficient to allocate all consumption to the largest households. On the other hand, for $\gamma > 0$, economies of scale mean that for the same level of consumption per adult, larger households enjoy a lower marginal utility of consumption. If $\gamma > 1$ this second effect dominates.

With prior knowledge of the appropriate equivalence scale $e\left(g,k\right)$ and the risk aversion parameter γ, one could purge variation in household size from the data by applying eq. (12) directly. Instead we choose to be agnostic *ex ante* about the function $e\left(g,k\right)$ and simply regress log household consumption on a full set of composition dummies. In the same consumption regression, we also strip out the age/time dummies $\mathcal{C}_t^a$ (by including a quartic polynomial in age and a full set of year dummies), and run similar regressions (minus the composition dummies, as dictated by the theory) for individual wages and hours.[21]

[21] Note that the polynomial in age also eliminates life-cycle effects in wages, hours, and consumption that we do not model.

Measurement error We assume that consumption, earnings, and hours worked are measured with error and that this error is classical, i.e., i.i.d. over time and across agents. The log of the observed value for variable x_t is then $\log \hat{x}_t = \log x_t + \mu_t^x$, where measurement error μ_t^x has mean zero and variance $v_{\mu x}$. While we directly observe consumption, hours, and earnings, we compute hourly wages as earnings divided by hours. Hence measurement error in hourly wages reflects errors in both earnings and hours.

Augmented allocations Augmented log allocations at time t are therefore given by

$$\log \hat{w}_t = \alpha_t + \kappa_t + \theta_t + \mu_t^y - \mu_t^h \tag{13}$$

$$\log \hat{c}_t = -(1-\tau)\widehat{\varphi} + (1-\tau)\left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)\alpha_t + \mu_t^c \tag{14}$$

$$\log \hat{h}_t = -\widehat{\varphi} + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)\alpha_t + \frac{1}{\widehat{\sigma}}\varepsilon_t + \mu_t^h, \tag{15}$$

where, recall, $\widehat{\varphi}$ denotes the rescaled preference weight.

3.2 Interpreting cross-sectional variances and covariances

With these allocations in hand, we can express in closed form cross-sectional moments of the joint equilibrium distribution of wages, hours, and consumption. These theoretical moments represent an attractive feature of our framework, since they allow us to transparently interpret the dynamics of their empirical counterparts over the life cycle and over time.

We will focus on variances and covariances across all agents of age a at date t. These moments reflect dispersion both within and between islands. An important theoretical property of our framework (see Section 4.1) is that the information contained in these aggregate cross-sectional (co-)variances of wages, hours, and consumption is sufficient to identify all model parameters and to quantify risk sharing.[22]

We start from the moments in levels, which we call the "macro moments" and then move to those in differences, which we will refer to as the "micro moments."

Macro moments Let $var_t^a(\alpha)$ denote the within-cohort variance of cumulated permanent uninsurable shocks (up until) period t for agents of age a:

$$var_t^a(\alpha) = v_{\alpha^0,t-a} + \sum_{j=0}^{a-1} v_{\omega,t-j}. \tag{16}$$

[22]Note also that we do not need any data on wealth when estimating the model. Longitudinal wealth data could shed further light on how households smooth wage fluctuations (see, e.g., Krueger and Perri, 2010). In particular, wealth dynamics might help with the difficult task of distinguishing insurable shocks from predictable changes in wages.

Similarly, let $var_t^a(\widehat{\varphi}) = v_{\widehat{\varphi},t-a}$ denote the cohort $(t-a)$-specific variance of the rescaled preference weights, and let $var_t^a(\varepsilon) = v_{\kappa^0,t-a} + \sum_{j=0}^{a-1} v_{\eta,t-j} + v_{\theta t}$ be the variance of the insurable component of the wage for cohorts of age a in year t.

The macro moments for wages and hours for age group a at date t are, respectively,

$$var_t^a(\log \hat{w}) = var_t^a(\alpha) + var_t^a(\varepsilon) + v_{\mu y} + v_{\mu h} \tag{17}$$

$$var_t^a\left(\log \hat{h}\right) = var_t^a(\widehat{\varphi}) + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 var_t^a(\alpha) + \frac{1}{\widehat{\sigma}^2} var_t^a(\varepsilon) + v_{\mu h} \tag{18}$$

$$cov_t^a\left(\log \hat{w}, \log \hat{h}\right) = \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right) var_t^a(\alpha) + \frac{1}{\widehat{\sigma}} var_t^a(\varepsilon) - v_{\mu h}. \tag{19}$$

The variance of measured wages is the sum of variances of the orthogonal productivity components, plus the variances of measurement error in earnings and hours. The variance of hours has four components. First, the more heterogeneity in the taste for leisure φ, the larger is the cross-sectional dispersion in hours. Second, the variance of the uninsurable shock translates into hours dispersion proportionately to $1-\gamma$. As $\gamma \to 1$ (the log-consumption case), uninsurable shocks have no effect on hours. Third, the variance of the insurable shocks increases hours dispersion in proportion to the (squared) tax-modified Frisch elasticity. Finally, measurement error in hours contributes positively to observed dispersion.

The covariance between wages and hours has three components. The effect of uninsurable wage shocks on this covariance depends on the value for γ. If $\gamma > 1$, then uninsurable shocks decrease the wage-hours covariance, since strong income effects induce low wage (uninsured) workers to work longer hours. Insurable shocks, by contrast, make hours and wages move together. Measurement error in hours reduces the observed covariance between hours and wages (earnings divided by hours).

We now turn to the moments involving consumption:

$$var_t^a(\log \hat{c}) = (1-\tau)^2 var_t^a(\widehat{\varphi}) + (1-\tau)^2 \left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)^2 var_t^a(\alpha) + v_{\mu c} \tag{20}$$

$$cov_t^a\left(\log \hat{h}, \log \hat{c}\right) = (1-\tau)\, var_t^a(\widehat{\varphi}) + \frac{(1-\tau)(1+\widehat{\sigma})(1-\gamma)}{(\widehat{\sigma}+\gamma)^2} var_t^a(\alpha) \tag{21}$$

$$cov_t^a(\log \hat{w}, \log \hat{c}) = (1-\tau)\left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right) var_t^a(\alpha). \tag{22}$$

The variance of consumption is increasing in the variance of uninsurable preference heterogeneity and uninsurable wage shocks, as expected. Progressive taxation $(\tau > 0)$ reduces the variance of consumption for a given $var_t^a(\alpha)$. The role of labor supply depends on the value for γ: for $\gamma > 1$ a lower σ (higher Frisch) reduces consumption dispersion because labor supply offsets uninsurable wage shocks and dampens their impact on earnings.

The covariance between hours and consumption is increasing in the degree of preference heterogeneity, since individuals with higher φ work relatively few hours and thus earn and consume relatively less. The effect of uninsurable wage risk depends on the value of γ: when $\gamma > 1$, a positive uninsurable shock reduces hours worked but increases consumption.

The covariance between consumption and wages depends only on uninsurable wage shocks: fluctuations in uninsurable productivity affect both wages and consumption in the same direction. As expected, progressive taxation reduces this covariance.[23]

Dispersion over the life cycle Let $\Delta var_t^a (\log \hat{x}) = var_t^a (\log \hat{x}) - var_{t-1}^{a-1} (\log \hat{x})$ be the within-cohort change (i.e., between age $a-1$ in year $t-1$ and age a in year t) in the variance of $\log \hat{x}$. The model has sharp predictions for the life-cycle evolution of dispersion:

$$\Delta var_t^a (\log \hat{w}) = v_{\omega t} + v_{\eta t} + \Delta v_{\theta t} \tag{23}$$

$$\Delta var_t^a \left(\log \hat{h}\right) = \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 v_{\omega t} + \frac{1}{\widehat{\sigma}^2} (v_{\eta t} + \Delta v_{\theta t}) \tag{24}$$

$$\Delta cov_t^a \left(\log \hat{w}, \log \hat{h}\right) = \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right) v_{\omega t} + \frac{1}{\widehat{\sigma}} (v_{\eta t} + \Delta v_{\theta t}) \tag{25}$$

$$\Delta var_t^a (\log \hat{c}) = (1-\tau)^2 \left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)^2 v_{\omega t} \tag{26}$$

$$\Delta cov_t^a \left(\log \hat{h}, \log \hat{c}\right) = (1-\tau) \frac{(1-\gamma)(1+\widehat{\sigma})}{(\widehat{\sigma}+\gamma)^2} v_{\omega t} \tag{27}$$

$$\Delta cov_t^a (\log \hat{w}, \log \hat{c}) = (1-\tau) \left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right) v_{\omega t}. \tag{28}$$

None of these moments involve measurement error, reflecting our assumption that the variance of measurement error is independent of age and time. Moreover, because all shocks in our economy are either permanent or i.i.d., all of these moments are independent of age.

The rise in wage inequality over the life cycle is determined by the variance of the innovations to the permanent insurable and uninsurable components, and by the change in the variance of the transitory insurable component. Wage dispersion will increase over the life

[23] Since we have filtered out differences in mean values for allocations across age groups, the expressions for dispersion in the entire cross section are identical to those above, but without the age a superscripts. This follows from the variance decomposition $var_t (x) = \mathbb{E} [var_t^a (x)] + var_t [\mathbb{E} (x|a)]$, where the second term is zero if we abstract from the terms $\mathcal{C}_t^a$ and $\mathcal{H}_t^a$ in the allocations. Thus, for example, $var_t (\log \hat{w}) = var_t(\alpha) + var_t(\varepsilon) + v_{\mu y} + v_{\mu h}$, where $var_t(\alpha) = (1-\delta) \sum_{a=0}^{\infty} \delta^a var_t^a(\alpha)$ is the unconditional cross-sectional variance of the uninsurable component of log wages, and $var_t(\varepsilon)$ is the corresponding variance for the insurable component of wages.

cycle as permanent shocks cumulate. The model suggests that the variance of hours should be increasing over the life cycle for the same reasons as wages, though with different weights on the insurable and uninsurable permanent variances. In the log-consumption utility case ($\gamma = 1$), only the former matters for hours.

Whether the covariance between wages and hours rises or falls over the life cycle depends on risk aversion and the relative size of permanent and transitory innovations. When $\gamma > 1$, the cumulation of permanent uninsurable shocks pushes the covariance down as individuals age, while the cumulation of permanent insurable shocks pulls the covariance up.

The change in the variance of consumption over the life cycle is determined by the variance of uninsurable productivity shocks. The uninsurable-wage-shock coefficient for consumption is exactly one when $\tau = 0$ and either $\gamma = 1$ or $\sigma \to \infty$.

When $\gamma > 1$, hours move up in response to a negative uninsurable wage shock, while consumption moves down, driving the consumption-hours covariance down over the life cycle as $var_t^a(\alpha)$ rises with age. Finally, the model predicts that the covariance between consumption and wages will increase over the life cycle, in proportion to $v_{\omega t}$.

Micro moments Micro moments are computed as variances and covariances of individual changes in log wages and log hours between $t-1$ and t.[24] Let $\Delta \log \hat{x}_t \equiv \log \hat{x}_t - \log \hat{x}_{t-1}$ denote the observed individual growth rate for variable $\hat{x}$, and let $var_t^a(\Delta \log \hat{x})$ be its cross-sectional variance, for the set of individuals of age a at date t for whom variable $\hat{x}$ is observed at both $t-1$ and t :

$$var_t^a(\Delta \log \hat{w}) = v_{\omega t} + v_{\eta t} + v_{\theta t} + v_{\theta,t-1} + 2v_{\mu y} + 2v_{\mu h} \tag{29}$$

$$var_t^a\left(\Delta \log \hat{h}\right) = \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 v_{\omega t} + \frac{1}{\widehat{\sigma}^2}\left(v_{\eta t} + v_{\theta t} + v_{\theta,t-1}\right) + 2v_{\mu h} \tag{30}$$

$$cov_t^a\left(\Delta \log \hat{w}, \Delta \log \hat{h}\right) = \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right) v_{\omega t} + \frac{1}{\widehat{\sigma}}\left(v_{\eta t} + v_{\theta t} + v_{\theta,t-1}\right) - 2v_{\mu h} \tag{31}$$

Again, the model implies that the variances and covariances of individual growth rates should be invariant to age and thus common across cohorts. Similar expressions obtain for second differences in wages and hours. For example, the variance of wage growth over a

[24]Given the specification of the stochastic process for shocks and measurement error, in the model covariances of the individual changes are all zero beyond lag one. Moreover, we omit moments involving changes in consumption, since we do not use the longitudinal dimension of CEX. The panel aspect of CEX is quite weak. It consists of two, generally noisy, observations spaced nine months apart. See Davis (2003) for a discussion.

two-year horizon is

$$var_t^a\left(\Delta^2 \log \hat{w}\right) = v_{\omega t} + v_{\omega,t-1} + v_{\eta t} + v_{\eta,t-1} + v_{\theta t} + v_{\theta,t-2} + 2v_{\mu y} + 2v_{\mu h}. \tag{32}$$

As we shall see, such moments are especially useful for exploiting the PSID data in the years when the survey was conducted biannually.

Finally, note that all of our cross-sectional moments are the sum of additively separable terms capturing the roles of preference heterogeneity, insurable productivity shocks, uninsurable productivity shocks, and measurement error. This implies that (co-)variance decompositions are always unique, in sharp contrast to the existing literature (e.g., Keane and Wolpin, 1997; Storesletten, Telmer, and Yaron, 2004a; Heathcote, Storesletten, and Violante 2010b), where decompositions must be obtained by simulation, and where the sequence in which various model ingredients are added or removed typically affects their measured contribution to moments of interest. In Section 5.4 we document our decompositions in detail.

4 Identification, data, and estimation

In this section, we first exploit the closed-form cross-sectional moments to prove identification of the model parameters. Next, we describe the data used for the structural estimation, and finally we discuss our estimation method. We estimate all structural parameters except δ and τ, which are set exogenously. Both macro and micro moments contain valuable information about parameters, and both are used to identify and estimate the model.

4.1 Identification

Typically, identification in estimated structural equilibrium models is discussed only at an informal level, because the mapping from parameters to equilibrium moments can at most be weakly illuminated by numerical experimentation. In contrast, our closed-form expressions for equilibrium allocations deliver explicit analytical links between structural parameters and equilibrium moments, enabling us to prove identification formally and lending transparency to the empirical analysis. This is one of the key payoffs from the tractability of our framework.

The conditions for identification depend on data availability. We therefore consider an array of different scenarios. Our baseline scenario (Proposition 2 below) is that one has access to an unbalanced panel on wages and hours (e.g., the PSID) and a repeated cross section on wages, hours, and consumption (e.g., the CEX). Next, we consider several variants encompassing alternative data structures.

Proposition 2 [IDENTIFICATION] *With an unbalanced panel on wages and hours and a repeated cross section on consumption, wages, and hours from $t = 1, ..., T$, the parameters $\{\sigma, \gamma, v_{\mu h}, v_{\mu y}, v_{\mu c}\}$ as well as the sequences $\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\}_{t=1}^{T}$, $\{v_{\kappa^0 t}, v_{\theta t}\}_{t=1}^{T-1}$, $\{v_{\omega t}\}_{t=2}^{T}$ and $\{v_{\eta t}\}_{t=2}^{T-1}$ are identified. The sums $v_{\eta T} + v_{\theta T}$ and $v_{\kappa^0 T} + v_{\theta T}$ are also identified.*

Proof. See Appendix A.3. ■

We now consider two alternative data structures that reflect additional limitations of available survey data for the United States. The first constraint is that consumption data in the CEX are available only from 1980, whereas the PSID starts in 1967. The second limitation is that, starting in 1996, the PSID becomes biannual. Since we estimate the model by combining the PSID and the CEX, these next two corollaries are important for us.

COROLLARY 2.1 [LIMITED CONSUMPTION DATA] *Suppose available data comprise an unbalanced panel on wages and hours from $t = 1, ..., T$ and a repeated cross section on consumption, wages, and hours for at least two years $\hat{t}$ and $\hat{t} + 1$, where $1 \leq \hat{t} < T$. Then, parameter identification is exactly as in Proposition 2.*

COROLLARY 2.2 [BIANNUAL PANEL DATA] *Suppose available data comprise an unbalanced panel on wages and hours and a repeated cross section on wages, hours, and consumption, where the cross-sectional data on consumption are annual for all years $t = 1, ..., T$, while the panel data on wages and hours are annual only until year $\hat{t}$ and biannual thereafter, i.e., data are available for the years $t = 1, 2, ..., \hat{t}$ and $t = \hat{t} + 2, \hat{t} + 4, ..., T - 2, T$. Then, one can identify $\{\sigma, \gamma, v_{\mu h}, v_{\mu y}, v_{\mu c}\}$, the sequences $\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\}_{t=1}^{T}$, $\{v_{\omega t}\}_{t=2}^{T}$, $\{v_{\theta t}, v_{\kappa^0 t}\}_{t=1}^{\hat{t}}$, $\{v_{\eta t}\}_{t=2}^{\hat{t}}$, and $\{v_{\theta t}, v_{\kappa^0 t}, v_{\eta, t-1} + v_{\eta t}\}$ for the years $t = \hat{t} + 2, \hat{t} + 4, ..., T - 2$, as well as the sums $\{v_{\eta, T-1} + v_{\eta, T} + v_{\theta, T}\}$ and $\{v_{\kappa^0, T} + v_{\theta, T}\}$.*

These two corollaries are proved in Appendix A.3. It is also straightforward to prove that, up to the composition of insurable shocks (i.e., the split between $v_{\theta t}$, $v_{\eta t}$, and $v_{\kappa^0 t}$), the model is also identified with *only* cross-sectional data on consumption, hours, and wages –for example, with data from the CEX alone.[25]

[25] To see this, note that Step A of the proof of Proposition 2 identifies σ, γ, $\{v_{\omega t}\}_{t=2}^{T}$, and $\{v_{\eta t} + \Delta v_{\theta t}\}_{t=2}^{T}$. Following Step C of the same proof, one identifies $\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\}_{t=1}^{T}$ and $\{v_{\kappa^0 t} + v_{\theta t}\}_{t=1}^{T}$. Measurement error $\{v_{\mu y}, v_{\mu h}, v_{\mu c}\}$ is identified following Step D.

Polynomial model for the variances Small sample sizes and data quality issues might preclude precise point estimates of year-specific shock variances. One way to reduce the information needed in estimation is to restrict the time series for the variances to follow time polynomials. In the baseline estimation, we follow this approach and model the time paths for the variances of insurable and uninsurable innovations $\{v_{\eta t}, v_{\omega t}\}$ as fourth-order time polynomials. This choice allows us to estimate a more parsimonious model (the number of parameters is reduced from 232 to 164) that can still capture the low-frequency movements in insurable and uninsurable wage risk in which we are interested.[26] Moreover, this restriction improves overall identification, as we demonstrate in the following corollary to Proposition 2 (proved in Appendix A.3).

Corollary 2.3 [time-polynomials for $(v_{\eta t}, v_{\omega t})$] *Suppose the sequences* $\{v_{\eta t}, v_{\omega t}\}_{t=1}^{T}$ *are modelled as time-polynomials of order* $T-3$ *or lower. Then, with an unbalanced panel on wages and hours, and a repeated cross section on consumption, wages, and hours from* $t = 1, ..., T$, *the parameters* $\{\sigma, \gamma, v_{\mu h}, v_{\mu y}, v_{\mu c}\}$ *as well as all the entire sequences* $\{v_{\widehat{\varphi} t}, v_{\alpha^0 t}, v_{\kappa^0 t}, v_{\theta t}, v_{\omega t}, v_{\eta t}\}_{t=1}^{T}$ *are identified.*

Analogous modifications on identification can be easily shown for the alternative data structures corresponding to Corollaries 2.1 and 2.2.

4.1.1 Identification via labor supply

It is well understood in the literature that consumption data can be used to differentiate between insurable and uninsurable shocks (see, e.g., Attanasio and Davis, 1996; Blundell and Preston, 1998; Guvenen and Smith, 2010). Proposition 2 and its corollaries expand this earlier research by introducing data on hours worked alongside consumption to obtain sharper identification. We now prove that, under a weak additional restriction on measurement error, the whole model can be identified *without using any consumption data.*

Proposition 3 [identification with no consumption data] *With an unbalanced panel on wages and hours from* $t = 1, ..., T$, *and an external estimate of measurement error in earnings* $v_{\mu y}$, *all the parameters listed in Proposition 2 are identified.*

[26] We chose to restrict only $v_{\eta t}$ and $v_{\omega t}$ to follow time polynomials because (as we explain in Section 5) those variances, when unconstrained, were by far the most volatile and least precisely estimated. In some years, point estimates hit the zero lower bound, suggesting a practical identification problem.

Proof. See Technical Appendix C.1.[27] ■

Why are data on labor supply informative about risk sharing and preference parameters? At a basic level, the logic is that theory has sharply different implications for the response of hours to uninsurable versus insurable shocks, just as for consumption. Households adjust hours worked more strongly in response to latter type of wage fluctuations, because of the absence of offsetting wealth effects. Moreover, the magnitudes of these responses are mediated by preference parameters.

4.2 Data

Our data are drawn from two surveys, the *Michigan Panel Study of Income Dynamics* (PSID), and the *Consumer Expenditure Survey* (CEX). We use PSID data for interview years 1968-2007 (which refer to calendar years 1967-2006). After the 1997 interview, the PSID becomes biannual, so we only have data for survey years 1968-1997, 1999, 2001, 2003, 2005, and 2007. We use CEX data from the quarterly Interview Surveys. Consistent and continuous data over time are available annually since 1980, hence we restrict attention to the 1980-2006 surveys.[28]

Since we jointly use both PSID and CEX data, we apply the same sample selection criteria to both datasets. Namely, we exclude badly incomplete or highly implausible observations.[29] We use an imputation procedure to adjust for top-coding based on the Pareto distribution. We then select households in which the male is between the ages of 25 and 59, and works at least 260 hours in the year.[30] In both datasets, the hourly wage is computed as annual pre-tax labor earnings divided by annual hours worked.[31] To avoid severe selection issues,

[27] Proposition 3 has two immediate implications. First, with an unbalanced panel, only a very short longitudinal dimension is required: all parameters are identified with a three-year panel. Second, the model could alternatively be estimated with longitudinal data on wages and hours for a single cohort. Therefore, besides the PSID, the model can be estimated on the SIPP or the NLSY. With a two-year panel (for example, the rotating panel of the CPS) all parameters are identified, except for $v_{\eta t}$.

[28] In the PSID, we exclude all PSID oversamples (SEO, Latino) so we do not need sample weights, while for the CEX computations use sample weights throughout.

[29] We drop records if 1) there is no information on age for either the head or the spouse, 2) if either the head or spouse has positive labor income but zero annual hours, and 3) if either the head or spouse has an hourly wage less than half of the corresponding federal minimum wage in that year. In the CEX, we drop households that report implausibly low quarterly consumption expenditures (less than $100, in 2000 dollars). In order to reduce measurement error, we also exclude CEX households flagged as "incomplete income reporters."

[30] The resulting unbalanced panel from the PSID comprises 2,930 individuals and 93,153 person-year observations. The resulting repeated cross sections from the CEX have a total of 87,966 household-year observations (on average, 3,258 households per year).

[31] Labor earnings are defined in both surveys as the sum of all income from wages, salaries, commissions,

we use wages and hours for males only. Our measure of household consumption includes expenditures on nondurables, services, small durables, and an estimate of the service flow from vehicles and housing. All nominal variables are deflated using the Consumer Price Index (CPI-U). Our PSID and CEX samples are updated versions of those constructed by Heathcote, Perri, and Violante (2010). We refer to that paper for a detailed description of these two surveys, the sample selection, and exact variable definitions.

As discussed in Section 3.1, we regress individual log wages, individual log hours, and household log consumption on year dummies, a quartic in age, and (for consumption) household composition dummies.

We then use the residuals from these regressions to construct variances and covariances in levels and differences for all available age/year cells constructed by grouping observations in any given year into 31 five-year overlapping age classes (27-57).[32] From the PSID data we construct (i) 1,085 age/year covariances corresponding to 31 age groups over 35 years (1967-1996, 1998, 2000, 2002, 2004, 2006) for each of the three moments in levels involving wages and hours; (ii) 899 age/year covariances corresponding to 31 age groups over 29 years for each of the three moments in first differences; and (iii) 1,203 age/year covariances corresponding to 31 age groups over 33 years for each of the three moments in second differences. From the CEX data, we construct 837 age/year covariances corresponding to 31 age groups over 27 years (1980-2006) for each of the three moments in levels involving consumption.

4.3 Estimation method

The structural estimation of the model uses the minimum distance estimator introduced by Chamberlain (1984), which minimizes a weighted squared sum of the differences between each moment in the model and its data counterpart. Let $\mathbf{m}(\Lambda)$ denote the $(J \times 1)$ vector of theoretical covariances, and Λ denote the $(N \times 1)$ vector of parameter values to estimate. Correspondingly, we define $\hat{\mathbf{m}}$ as the vector of empirical covariances. The estimator solves the following minimization problem:

$$\min_{\Lambda} [\hat{\mathbf{m}} - \mathbf{m}(\Lambda)]' \mathcal{W} [\hat{\mathbf{m}} - \mathbf{m}(\Lambda)], \tag{33}$$

bonuses, and overtime, and the labor component of self-employment income.

[32]For example, the variance of log wages for the youngest age group (age class 27) at date t is constructed with all wage observations for individuals aged 25-29 at date t, the variance of log wages for the next age group (age class 28) at date t is constructed with all wage observations for individuals aged 26-30 at date t, and similarly for all other age groups until the oldest one (age class 57). Since the number of observations in many one-year age cells is very small, this procedure reduces sampling variation. We apply the same procedure to construct the model analogue of these moments.

where $\mathcal{W}$ is a $(J \times J)$ weighting matrix. Standard asymptotic theory implies that the estimator $\widehat{\Lambda}$ is consistent and asymptotically Normal. Due to the small sample size, we make two choices: (i) we use an identity matrix for $\mathcal{W}$;[33] (ii) we compute 90–10 confidence intervals through a block-bootstrap procedure based on 500 replications.[34]

The discussion of identification in Section 4.1 indicates that, absent additional assumptions, one cannot identify some of the time-varying parameters in the missing PSID survey years. We describe the minor technical identifying assumptions needed to overcome this issue in Appendix A.4. Moreover, we assume that prior to 1967 the variances of all shocks were equal, in each year, to their respective values in 1967.[35] Overall, the estimation uses $J = 11,532$ moment conditions for $N = 164$ parameters.

Parameters set outside the model We set $\delta = 0.996$ to match the annualized probability of surviving from age 25 to age 60 for US men.[36] To estimate the progressivity parameter τ, for each household in our PSID sample we compute after-tax income as income minus all federal and state taxes (calculated using the NBER's TAXSIM program) plus social security benefits. We exclude state-contingent government transfers in the form of cash (e.g., UI benefits and TANF) or kind (e.g., food stamps and Medicaid) since, as discussed earlier, this type of social assistance is subsumed in our estimate of insurance with respect to ε shocks.[37] From eq. (4), a consistent estimate of $1 - \tau$ can be obtained by regressing log household after-tax income on log household pre-tax income, including a constant in the regression. The ordinary least squares estimate of this coefficient implies $\tau = 0.185$ ($s.e. = 0.001$). The associated R^2 measure of fit is 0.92, which demonstrates that our functional form provides a good approximation to the actual US tax system.

[33] The bulk of the literature follows this strategy, in light of the Monte Carlo simulations of Altonji and Segal (1996) who argue that in common applications there is a substantial small sample bias when using the optimal weighting matrix characterized by Chamberlain (1984).

[34] Bootstrap samples are drawn at the household level with each sample containing the same number of observations as the original sample. The implied confidence intervals thus account for arbitrary serial correlation, heteroscedasticity, and estimation error induced by the first-stage regression of individual observations on age, time, and household type.

[35] Alternatively, we could have treated the cumulative variances of the insurable and uninsurable components for the cohorts alive in 1967, i.e., $\{v_{\kappa^a,1967}, v_{\alpha^a,1967}\}_{a=27}^{57}$, as parameters to be estimated. When pursuing this alternative estimation strategy, we found the results to be virtually identical to those under the baseline "steady-state" identification scheme.

[36] The survival rate δ does not appear in any of the age/year moments we use to estimate the model, and hence its calibration has no bearing on the parameter estimates. We use δ only to construct the aggregate cross-sectional variances and covariances plotted to measure the fit of the model against the data. The fit is extremely robust to varying δ within a plausible range.

[37] Since state income taxes from TAXSIM are only available from 1978, we exclude years 1967-1977 in this calculation. See Appendix B in Heathcote, Perri, and Violante (2010) for details.

Table 1: Baseline Parameter Estimates

Preference Elasticities			Life-Cycle Shocks		
σ	γ		$\overline{v_\omega}$	$\overline{v_\eta}$	$\overline{v_\theta}$
2.165	1.713		0.0056	0.0044	0.043
(0.173)	(0.054)		(0.0008)	(0.0012)	(0.005)
Initial Heterogeneity			**Measurement Error**		
$\overline{v_{\alpha^0}}$	$\overline{v_{\kappa^0}}$	$\overline{v_{\widehat{\varphi}}}$	$v_{\mu y}$	$v_{\mu h}$	$v_{\mu c}$
0.102	0.047	0.054	0.000	0.036	0.041
(0.030)	(0.023)	(0.016)	(0.000)	(0.006)	(0.002)

Notes: Bars denote sample averages. Bootstrapped standard errors based on 500 replications are shown in parentheses.

5 Results

Table 1 reports parameter estimates. Our estimates for the two preference elasticity parameters are $\gamma = 1.71$ and $\sigma = 2.16$. In both cases the confidence intervals are narrow. Given our assumed value for the tax progressivity parameter τ, the implied tax-modified Frisch elasticity with respect to pre-tax wages is $1/\hat{\sigma} = (1-\tau)/(\sigma+\tau) = 0.35$, a value that is broadly consistent with the microeconomic evidence (see, e.g., Keane, 2011).

The average estimated values for the variances of uninsurable and insurable permanent wage shocks ($\overline{v_\omega}$ and $\overline{v_\eta}$) and corresponding cohort effects ($\overline{v_{\alpha^0}}$ and $\overline{v_{\kappa^0}}$) indicate that almost 45% of permanent life-cycle wage innovations are insurable, while around 30% of initial wage variation at labor market entry is insurable.[38] The estimated average transitory wage variance is $\overline{v_\theta} = 0.043$, an order of magnitude larger than the variance of permanent shocks. The entire time series for the variances are reported in Table E in the Technical Appendix. Our estimates for the variances of measurement error in log hours worked, individual earnings, and household consumption are, respectively, 0.036, 0, and 0.041.[39]

[38] The total variance of permanent wage innovations is 0.01, in line with existing estimates. For example, Low, Meghir, and Pistaferri (2010) estimate a variance of permanent wage shocks of 0.011.

[39] The estimate of zero measurement error in earnings might seem surprising. However, Gottschalk and Huynh (2010) find that the cross-sectional variance of true earnings is greater than the variance of measured earnings in survey data. They argue that this reflects a non-classical structure for measurement error in earnings.

5.1 Life-cycle fit

Figures 1 and 2 compare the evolution of model and data along the life-cycle dimension and show that the model-implied moments align closely with their empirical counterparts from the PSID and the CEX. In particular, the model-implied moments almost always lie within the 90-10 confidence intervals around the empirical moments. With the help of these figures, we offer some economic intuition relating the life-cycle profiles for inequality to the parameter estimates described above. We then demonstrate that each feature of the baseline model plays an important role in accounting for the empirical moments by estimating a set of restricted models.

Understanding parameter estimates In both US and model-simulated data, the variance of log wages increases by around 37 log points, approximately linearly, between ages 27 and 57. In contrast, the variance of log consumption grows much less, by about 10 log points over the life cycle. The much steeper life-cycle increase in wage dispersion relative to consumption dispersion explains why almost half of permanent shocks to wages are estimated to be insurable.

The fact that the empirical profile for the variance of log hours is fairly flat, notwithstanding the fact that dispersion in wages increases sharply as permanent shocks cumulate, points to a relatively low Frisch elasticity of labor supply. However, we show below that the model fits poorly if we impose exogenously a *zero* Frisch elasticity.

The point estimate for γ exceeds one because the covariance between wages and hours is negative, indicating significant wealth effects from uninsurable shocks to wages (recall that insurable wage shocks push this covariance up).[40] The framework allows for one alternative way to generate a negative wage-hours covariance, namely measurement error in hours. However, the estimation procedure does not attribute the low covariance entirely to measurement error because this would translate into an excessively high variance for the growth of individual hours.[41]

[40] In a similar spirit, Chetty (2006) argues that existing empirical evidence on the response of hours to permanent shocks to wages can be used to bound estimates for risk aversion. An advantage of our fully structural approach is that we can identify γ in an environment with a mix of uninsurable and insurable permanent wage shocks.

[41] Figure 1 indicates that the estimated model exaggerates the increase in the correlation between wages and hours observed over the life cycle. A larger value for γ would improve the model's fit in this dimension, by amplifying the offsetting effect on hours or permanent uninsurable wage shocks. However, a larger value for γ would also steepen the age decline in the theoretical correlation between hours and consumption. See equations (19) and (21). Thus the estimated value for γ reflects a compromise in an attempt to reconcile various conflicting moments.

Table 2: Parameter Estimates for Alternative Models

	Baseline	(2)	(3)	(4)	(5)	(6)
σ	2.165	2.682	2.251	∞*	1.642	1.637
γ	1.713	1.483	1.849	1.713*	1.705	2.108
$v_{\mu h}$	0.036	0.037	0.036	0.033	0.039	0.037
$v_{\mu y}$	0.000	0.000	0.000	0.000	0.000	0*
$v_{\mu c}$	0.041	0.040	0.038	0.016	0.084	0.041
$\overline{v}_{\widehat{\varphi}}$	0.054	0.050	0.055	0.058	0*	0.036
$\overline{v}_{\alpha^0}$	0.102	0.156	0.070	0.086	0.089	0.085
$\overline{v}_{\omega}$	0.0056	0*	0.0093	0.0059	0.0064	0.0063
$\overline{v}_{\kappa^0}$	0.047	0.014	0.082	0.051	0.065	0.067
$\overline{v}_{\eta}$	0.0044	0.0081	0*	0.0056	0.0035	0.0031
$\overline{v}_{\theta}$	0.043	0.043	0.045	0.043	0.039	0.042
$1/\widehat{\sigma}$	0.347	0.284	0.334	0*	0.446	0.447
SSR	10.204	11.213	11.314	13.012	14.114	–

Notes: Externally set values are followed by an asterisk. The baseline estimates are reproduced from Table 1. Other columns: (2) complete markets for all shocks ($v_{\omega t} = 0$), (3) no private insurance against permanent shocks ($v_{\eta t} = 0$), (4) inelastic labor supply ($\sigma \to \infty$), (5) no preference heterogeneity ($v_{\varphi t} = 0$), and (6) baseline model without using CEX consumption data (Section 5.5). Values for $1/\widehat{\sigma}$ are implied by the other parameter estimates. The sum of squared residuals SSR is reported only where comparable with the baseline.

Figure 2 shows that the model also accounts well for the life-cycle moments in first and second differences. For example, the top left and bottom left panels plot the cross-sectional variances of annual and bi-annual log wage growth. The first differences apply to the period 1967-1996, while the second differences refer to 1967-2006. How does the model discriminate between transitory insurable shocks and measurement error? Equations (29)-(32) illustrate that if moments in first (and second) differences were driven primarily by measurement error in hours, then the correlation between hours and wage growth would be close to minus one. A substantial amount of true transitory wage variation is needed to raise this correlation to the level observed in the data. Finally, note that the variance of biannual wage and hours growth (the bottom panels) is not much larger than the variance of annual growth, which helps explain why the estimated variances for permanent shocks are small relative to transitory shocks.

Alternative models: What goes wrong? To better understand why each model element is needed to account for the observed cross-sectional moments, we now discuss a range of experiments in which we shut down one model element at a time, and re-estimate the model. See Table 2 for the parameter estimates of these alternative models.

We first consider two alternative insurance market structures. In the first, we assume perfect insurance against permanent life-cycle shocks, by imposing $v_{\omega t} = 0$. In the second, we make the opposite assumption, namely that there is no explicit insurance against permanent life-cycle shocks, by imposing the restriction $v_{\eta t} = 0$. This economy captures the spirit of the permanent income hypothesis (PIH), according to which transitory shocks are largely insurable, while permanent shocks are uninsurable.

The estimated "complete markets" model ($v_{\omega t} = 0$) features almost twice as large an average variance for permanent insurable shocks $\overline{v_\eta}$ relative to the baseline model. Absent changes in other parameter values, this would imply too much dispersion in hours worked and too little dispersion in consumption: thus, the estimation also delivers a larger estimate for σ (a lower Frisch) and a higher estimate for $\overline{v_{\alpha^0}}$ (more uninsurable wage dispersion at labor market entry). However, absent permanent uninsurable shocks, the model has no way to generate the observed rise in consumption dispersion over the life-cycle. Another indication that this model exaggerates insurance against life-cycle shocks is that it generates much too large an increase in the correlation between wages and hours over the life cycle.

The estimated PIH model ($v_{\eta t} = 0$) delivers similar parameter estimates to the baseline model, with the exception that the average variance of permanent uninsurable shocks $\overline{v_\omega}$ rises from 0.0056 to 0.0093. Perhaps surprisingly, the estimated model replicates fairly closely the empirical life-cycle profile for the variance of log consumption, because uninsurable wage shocks are partially smoothed via labor supply and progressive taxation. However, the model now generates a counterfactual decline over the life cycle in the correlation between wages and hours worked. Recall that permanent uninsurable shocks drive this correlation down, while permanent insurable shocks (shut off in this experiment) drive the correlation up. Consequently, the estimated model also delivers a life-cycle increase in the variance of earnings that is much too small.

We next experiment with shutting off flexible labor supply by setting $\sigma = \infty$ in the baseline model.[42] With inelastic labor supply, measurement error is the only source of variance in the growth of individual hours. However, with a zero Frisch elasticity, measurement error in hours implies a negative correlation between wages and hours worked, while this correlation

[42]Technically, we set $\sigma = 500$. With σ large but finite, the model can still generate dispersion in hours through preference heterogeneity. Given a Frisch elasticity near zero, our identification strategy for γ (based on cross-sectional moments involving hours) fails. Thus we set γ equal to its value in the baseline model.

is close to zero in the data. The estimation compromises, delivering too little variation over time in individual hours and a counterfactually negative wage-hours correlation. In addition, the model with inelastic hours generates too much comovement between hours worked and consumption because it rules out income effects as a force to offset preference heterogeneity. We conclude that allowing for elastic labor supply is essential in accounting for all moments involving hours worked.

In our last experiment, we eliminate preference heterogeneity by imposing $v_{\widehat{\varphi}t} = 0$. In our baseline model, preference heterogeneity is required to replicate the positive empirical correlation between hours worked and consumption. Absent preference variation, the model generates a counterfactual negative correlation since, with $\gamma > 1$, individuals with a higher uninsurable wage component enjoy more consumption but work fewer hours – see equations (7) and (8). Preference heterogeneity also plays an important role in generating cross-sectional dispersion in hours worked and consumption, and when it is shut down the estimation looks for alternative ways to replicate these moments. In particular, it assigns larger values for the variance of measurement error in consumption and delivers a higher Frisch elasticity.

We conclude this section by highlighting two key messages from this exploration of alternative models. First, the overall model fit worsens dramatically in each restricted version of the baseline model we estimate (see the sum of squared residuals in Table 2), indicating that each model element plays an important quantitative role in accounting for observed dynamics of inequality. In particular, the data – and especially the moments involving hours worked – speak strongly to the existence of risk-sharing mechanisms that allow households to insure a fraction (but only a fraction) of permanent idiosyncratic fluctuations in wages. They also speak strongly to the existence of two fundamental drivers of dispersion in hours worked: a positive elasticity in response to wage fluctuations and a second source of dispersion in hours that is unrelated to wages.

Second, it is important to estimate the scope for risk sharing and preference parameters jointly. The logic is simply that both matter for the dynamics of consumption and labor supply. If we use more restricted models for risk sharing (by imposing too much or too little insurance), the estimation contorts estimates for preference elasticities or for preference heterogeneity in order to try to match the same moments involving consumption and hours. If we restrict the model for preferences (by imposing inelastic hours or an absence of preference heterogeneity), the model delivers the wrong estimate for the fraction of wage risk that is insurable.

5.2 Insurance and inequality over the life cycle

We now turn to the first of our motivating questions: How effectively can households smooth idiosyncratic wage fluctuations via insurance arrangements, labor supply adjustments, and progressive taxation?

Pass-through coefficients There are three reasons for incomplete pass-through from changes in wages to changes in consumption. First, shocks to wages that are insurable will not be reflected in changes in consumption. Second, labor supply decisions determine how uninsurable wage shocks transmit to earnings. Third, the progressive tax system dampens the response of consumption to fluctuations in earnings.

Let $\phi_t^{w,c}$ denote the pass-through coefficient from wages to consumption, defined as the OLS coefficient from a panel regression of model-simulated changes in log consumption between $t-1$ and t on permanent (uninsurable or insurable) changes in log individual wages.[43] We focus here on permanent shocks, because transitory shocks are fully insurable in our framework. The elasticity of consumption with respect to an uninsurable permanent innovation ω_t is $(1+\widehat{\sigma})\,/\,(\widehat{\sigma}+\gamma)\cdot(1-\tau)$ (see eq. 7), while consumption does not respond to permanent insurable innovations η_t. Thus $\phi_t^{w,c}$ is given by

$$\underbrace{\phi_t^{w,c}}_{0.386} = \underbrace{\frac{v_{\omega t}}{v_{\omega t}+v_{\eta t}}}_{0.560} \cdot \underbrace{\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}}_{0.845} \cdot \underbrace{(1-\tau)}_{0.815}. \tag{34}$$

Plugging in the estimated values for γ, σ, $\overline{v_\omega}$, and $\overline{v_\eta}$ from Table 1 along with $\tau = 0.185$ gives an average pass-through coefficient of $\bar{\phi}^{w,c} = 0.386$. Thus, on average, less than 40% of permanent wage shocks transmit to consumption.

The roles of explicit insurance, labor supply, and progressive taxation in delivering consumption smoothing against permanent wage fluctuations are captured, respectively, by the three terms in the expression for $\phi_t^{w,c}$. Evaluated at the sample-average parameter estimates, 44% of permanent wage shocks are explicitly insured. Recall that we have remained agnostic on the sources of this insurance: state-contingent private or public transfers, spousal labor supply, and perfect smoothing of forecastable wage changes are among the most plausible

[43] According to the model of the household described in Section 3.1, the household composition dummy $D(g,k)$ drops out when looking at the growth rate of log consumption. This implies that $\phi_t^{w,c}$ can be interpreted either as measuring pass-through to raw household consumption, or as pass-through to equivalized consumption.

candidates. Of the non-insured component of wages, 15.5% of fluctuations are smoothed through individual labor supply, reflecting the fact that our estimate for γ is larger than one (see eq. 8). Of the component transmitted to earnings, 18.5% of fluctuations are smoothed through progressive taxation. We conclude that all three channels play important roles in mediating the response of consumption to permanent wage shocks. Explicit insurance is the most important of these channels, followed by progressive taxation.[44]

While the primitive shocks in our model are shocks to wages, we can also compute a pass-through coefficient from pre-tax individual earnings to consumption:

$$\phi_t^{y,c} = \frac{v_{\omega t}}{v_{\omega t} + \left(\frac{\hat{\sigma}+\gamma}{\hat{\sigma}}\right)^2 v_{\eta t}} \cdot (1-\tau) \tag{35}$$

which implies an average value of $\bar{\phi}^{y,c} = 0.272$. Blundell, Pistaferri, and Preston (2008, Table 7) estimate a quantitatively similar pass-through coefficient of 0.225 from permanent shocks to male earnings to non-durable consumption on US data. They conclude that the bulk of permanent individual income risk is insurable. Our framework suggests that one has to be cautious with this interpretation, for two reasons. First, earnings are endogenous in the model, and the pass-through from the primitive wage shocks to consumption is 42 percent larger than the one for earnings. Second, because labor supply adjustments tend to amplify insurable wage shocks and dampen uninsurable wage shocks, pass-through from earnings to consumption can be low even if the underlying shocks are mostly uninsurable in nature. To see this, consider the extreme case in which preferences are linear in hours worked ($\sigma = 0$) and taxation is linear ($\tau = 0$). The pass-through coefficient from earnings to consumption $\phi_t^{y,c}$ would mistakenly suggest perfect risk-sharing, i.e., $\lim_{\sigma \to 0} \phi_t^{y,c} = 0$, irrespective of the size of $v_{\omega t}$ and $v_{\eta t}$, whereas $\phi_t^{w,c}$ would correctly indicate some degree of transmission of wage shocks to consumption.[45]

Growth in life-cycle variances An alternative, and more common, metric for quantifying the extent of smoothing against life-cycle income fluctuations is to compare the

[44] An alternative way to gauge the roles of these different smoothing mechanisms is to shut them off one at a time, and then compute by how much the implied pass-through coefficient would increase, holding constant other parameter values. We implement this by setting, respectively, $\overline{v_\eta} = 0$, $\sigma \to \infty$, and $\tau = 0$, in which cases $\bar{\phi}^{w,c}$ rises from 0.39 to, respectively, 0.69, 0.46, and 0.47. In this second calculation, the ranking of smoothing channels is thus the same as in the first one.

[45] We can also define a pass-through coefficient from permanent wage changes to pre-tax earnings:

$$\phi_t^{w,y} = \frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma} \cdot \frac{v_{\omega t}}{v_{\omega t} + v_{\eta t}} + \frac{1+\widehat{\sigma}}{\widehat{\sigma}} \cdot \frac{v_{\eta t}}{v_{\omega t} + v_{\eta t}}.$$

In our model, $\phi_t^{w,c} = \phi_t^{w,y} \cdot \phi_t^{y,c}$ if and only if either (i) $v_{\eta t} = 0$, (ii) $\gamma = 0$, or (iii) $\sigma \to \infty$.

within-cohort life-cycle growth in the variances of consumption on the one hand, and wages or earnings on the other (see, e.g., Blundell and Preston, 1998; Storesletten, Telmer, and Yaron, 2004a; Huggett, Ventura, and Yaron, 2011). The analytical expressions for these moments are in equations (23) and (26).

Our framework uncovers a useful relationship between (i) the ratio of life-cycle growth in the variance of consumption to growth in the variance of wages, and (ii) the pass-through coefficient described above. Assuming $\Delta v_{\theta t} = 0$, we obtain

$$\frac{\Delta var_t^a\left(\log \hat{c}\right)}{\Delta var_t^a\left(\log \hat{w}\right)} = (1-\tau)^2\left(\frac{1+\hat{\sigma}}{\hat{\sigma}+\gamma}\right)^2\frac{v_{\omega t}}{v_{\omega t}+v_{\eta t}} = (1-\tau)\left(\frac{1+\hat{\sigma}}{\hat{\sigma}+\gamma}\right)\cdot\phi_t^{\omega,c}. \tag{36}$$

This relation reveals that these two alternative measures of smoothing coincide exactly if and only if progressive taxation and labor supply are both absent as smoothing mechanisms, i.e., when either (i) $\tau = 0$ and $\sigma \to \infty$, or (ii) $\tau = 0$ and $\gamma = 1$. In the latter case, even though labor supply is elastic, it is not used to smooth uninsurable shocks to wages.

If $\tau > 0$ or if $\gamma > 1$ (and $\sigma < \infty$), then smoothing provided through taxation and/or labor supply shows up more strongly in the ratio $\Delta var_t^a\left(\log \hat{c}\right)/\Delta var_t^a\left(\log \hat{w}\right)$ than in the pass-through coefficient $\phi_t^{\omega,c}$. At our baseline parameter values, the life-cycle increase in the variance of log consumption is only 25% of the corresponding increase in the variance of log wages, even though around 40% of permanent wage shocks transmit to consumption.

5.3 Insurance and inequality over time

Insurability over time Table 2 in the Technical Appendix contains the complete set of year-by-year estimates for all the time-varying parameters of the model. Figure 3 summarizes what these estimates imply for changes over time in the structure of relative wages. Panel A shows that the variance of the total uninsurable component (α_t) declines slightly in the 1970s and then rises in the remainder of the sample period. This pattern broadly accords with the fall of the skill premium in the late 1960s to mid-1970s, and the subsequent increase in the 1980s and beyond. Under this interpretation, "skill-biased demand shifts" represent an important source of uninsurable wage shocks.[46] The total cross-sectional variance of the permanent insurable component of wages (κ_t) is generally increasing throughout the first two decades, but declines somewhat in the 1990s (Panel B). The variance of transitory insurable

[46]This interpretation is consistent with Attanasio and Davis (1996) and with Heathcote, Storesletten, and Violante (2010b), who, in the context of an augmented version of the standard incomplete-markets model, show that skill-biased demand shifts are the main driver of the rise in consumption inequality.

shocks (θ_t) plotted in Panel C grows steadily throughout the sample, consistent with Moffitt and Gottschalk's (2002) estimates for earnings dynamics.[47]

Combining these estimates allows us to address the second of our motivating questions: What fraction of the observed rise in wage dispersion over our sample period was insurable for US households? Panel D of Figure 3 indicates that in the late 1960s the insurable component of wages accounted for around one third of the cross-sectional variance of log wages, while by the early 1980s this fraction was around 50%. Since then, the variances of the two components of wages have risen at a similar rate, leaving the fraction of wage fluctuations insured relatively stable.

Finally, note that the "cohort" components in Panels A, B, and D are rather steady over time, indicating that the bulk of the dynamics in cross-sectional wage dispersion reflect changes in the variances of life-cycle shocks and not cohort effects.

Time series fit The variance of log male wages increases by around 15 log points over the sample period, with especially rapid growth in the 1980s. How do the moments involving consumption and hours account for the partition of this increase, described in Figure 3, between insurable and uninsurable risk? Figure 4 plots the evolution over time of these moments, alongside the corresponding values for the estimated model.

Over the first half of the sample, we see a sharp rise in the wage-hours correlation (Panel D). The model interprets this as indicating a rise in the variance of the insurable wage component and a fall in the variance of the uninsurable component. The latter translates into a theoretical prediction of modestly declining consumption inequality before 1980, when our CEX sample begins. This pattern for consumption inequality parallels the dynamics of the skill premium over the period.[48]

After 1980 consumption data are available and further inform the estimation. The variance of log consumption grows by only about 5 log points between 1980 and 2006, in line with earlier estimates by Krueger and Perri (2006). This rise, paired with the one in the

[47]Around 1992-1993, this variance displays a spike. This estimated higher volatility may be linked to the fact that survey year 1993 was the first year of computer-assisted telephone interviewing in the PSID. In the previous version of the paper (Heathcote et al., 2012), we allowed for a temporary increase in measurement error in 1992. Except for a slightly smaller variance of the transitory shock in 1992, this extension was inconsequential for the rest of the estimation, and hence in the current version we have omitted it.

[48]It is also broadly consistent with evidence from Slesnick (2001, chapter 6), who, notwithstanding data comparability issues, uses CEX data pre-1980 in order to construct a longer series for US consumption dispersion. Guvenen and Smith (2010, Figure A.1) impute nondurable consumption into the PSID from the CEX going back to 1967 and also uncover a decline in the variance of log consumption over the first decade of their sample.

wage-consumption correlation, calls for an increase in uninsurable wage dispersion and a slowdown in the rise of insurable wage dispersion. This pattern is also consistent with the end of growth in the empirical wage-hours correlation. The increase in the variance of consumption over time is small relative to the increase in uninsurable wage dispersion (see Figure 3) because, as with the life-cycle dimension, labor supply and progressive taxation mitigate the impact of uninsurable wage dispersion on consumption dispersion.

Larger uninsurable wage shocks tend to drive the consumption-hours correlation down over time. To offset this force and replicate the roughly flat pattern for the correlation in the data, the estimation calls for a modest increase over time in preference dispersion (see Table 2 in the Technical Appendix).

Figure 5 shows the time series plots for moments in first and second differences. Recall that these moments are driven primarily by measurement error and transitory wage shocks, given the relatively small estimated variances for the innovations to permanent shocks. Thus we can point to the rise in the variance of wage growth over time as the source of the corresponding rise in the estimated variance of transitory shocks (Panel C of Figure 3). These larger transitory shocks, in turn, account for the model-predicted increase in the correlation between wage and hours growth.

5.4 Inequality decomposition

We now turn to the third motivating question of our paper. Is observed cross-sectional inequality primarily the result of life-cycle shocks, initial heterogeneity in productivity and preferences, or simply measurement error? Given parameter estimates and the moment expressions in eqs. (17)-(22), variance decompositions are unique and easy to compute.

Cross-sectional average In Table 3, we report the average contribution of each component across the entire 1967–2006 period.[49]

Interestingly, initial heterogeneity explains between 40% and 50% of the observed variance for all variables. However, the source of this inequality at labor market entry varies. Preference heterogeneity is dominant in accounting for dispersion in hours worked, whereas heterogeneity in productivity (mostly uninsurable) is paramount for wages, earnings, and consumption. Measurement error also plays a large role, accounting for one-third of the observed variance for both hours and consumption. The flipside of the finding that initial

[49] These values are computed by taking survival-probability-weighted averages across within-age-group values for dispersion at each date, and then computing a simple average across the years in our sample.

Table 3: Decomposition of Cross-Sectional Inequality

	Total Variance	Percent Contribution to Total Variance					
		Initial Heterogeneity			Life-Cycle Shocks		Measurement
		Prefs.	Unins.	Ins.	Unins.	Ins.	Error
$var(\log \hat{w})$	0.351	0.0	31.5	10.0	17.1	31.3	10.1
$var(\log \hat{h})$	0.107	48.9	2.2	3.2	1.2	9.8	34.7
$var(\log \hat{y})$	0.432	11.7	22.8	12.5	10.4	43.7	0.0
$var(\log \hat{c})$	0.159	20.0	32.6	0.0	17.8	0.0	29.6

heterogeneity and measurement error account for a large share of dispersion in consumption and hours worked is that life-cycle shocks to wages contribute relatively little to dispersion in these variables. Instead, life-cycle shocks explain half of the cross-sectional variation in wages and earnings.

We conclude that there is no simple answer to the question: What determines measured cross-sectional inequality among households? The answer depends on the variable of interest: for hours it is mostly preference heterogeneity and measurement error; for wages and earnings it is dispersion in productivity, predominantly over the life cycle; while for consumption it is a mix of all these factors.

Lifetime earnings An alternative way to measure the relative roles of initial conditions versus life-cycle shocks is in terms of their contributions to discounted lifetime pre-tax earnings. Storesletten, Telmer, and Yaron (1994) conclude that roughly half of the variance of lifetime earnings is attributable to variation in initial conditions. Huggett, Ventura, and Yaron (2011, Table 5) estimate that heterogeneity in initial conditions accounts for 62% of the variance of lifetime earnings.[50] We have simulated a distribution for discounted lifetime earnings in our model, making the following two assumptions: 1) earnings are discounted at an annual rate of 4.2% over a 38 year working life (as in Huggett et al.), and 2) all wage innovations and initial conditions are log-normally distributed. Our estimates imply that initial conditions (i.e., dispersion in κ^0, α^0 and $\widehat{\varphi}$) account for 63% of the variance of lifetime earnings, which is very similar to the Huggett et al. estimate.[51]

[50] In an important early contribution, Keane and Wolpin (1991) estimated this fraction to be 91%. However, their estimate is a loose upper bound, because their model assumes i.i.d. wage shocks. That assumption, made for computational reasons, is clearly counterfactual.

[51] Huggett, Ventura and Yaron (2011) omit preference heterogeneity from their model, which we estimate to be an important determinant of inequality. On the other hand, an important initial condition in their model (but not ours) is idiosyncratic learning ability.

5.5 Estimation without consumption data

Section 3 documents that moments involving labor supply are informative about risk sharing. Proposition 3 proves that the model is in fact identified without any data on consumption. In this section, we exploit this identification result and re-estimate the model using only data on wages and hours from the PSID.[52] One motivation for this exercise is that there is some debate about how much consumption inequality has risen over time in the United States (e.g., Attanasio, Battistin, and Ichimura, 2007; Aguiar and Bils, 2011). A second motivation is that the literature on risk sharing to date focuses almost exclusively on moments involving consumption, and we would like to know whether moments involving labor supply tell a comparable story in terms of the fraction of idiosyncratic risk that households can insure.

When we estimate the model without CEX data, we find that the estimations with and without consumption data deliver very similar dynamics for the insurability of wage risk. Panel A of Figure 6 shows that the insurable fraction of total cross-sectional wage dispersion, as estimated without consumption data, is very close to the corresponding fraction in the baseline when consumption moments are used. Moreover, the estimated pass-through coefficient $\bar{\phi}^{w,c}$ is essentially unchanged relative to the baseline case (0.41 compared to 0.39).

The main difference relative to the baseline estimates is that estimated preference heterogeneity is now much smaller (see column (6) in Table 2). Figure 6 shows that lower preference heterogeneity translates to predicted levels for the variance of log consumption (Panel B) and the consumption-hours correlation (Panel D) that are much too low relative to their empirical counterparts. We conclude that it is consumption moments, and especially the positive covariance between consumption and hours, that offer the strongest evidence of extensive preference heterogeneity. Because the model without consumption data estimates a smaller role for preference heterogeneity, it calls for a higher Frisch elasticity of labor supply ($1/\widehat{\sigma}$ is now 0.45) in order to replicate observed hours dispersion. Although the model is estimated without consumption data, it replicates the dynamics of consumption moments remarkably well, subject to the caveat about levels discussed above (see also Figures E1 and E2 in the Technical Appendix).[53]

Taken together, these results indicate that moments involving labor supply and moments

[52] The identification proof of Proposition 3 is up to an external estimate for measurement error in earnings. We therefore impose the baseline estimate $v_{\mu y} = 0$.

[53] We experimented with estimating the model without consumption data while imposing the baseline estimates for γ, σ, and v_{φ}. In this case, the no-consumption-data model consumption moments are virtually indistinguishable from those of the baseline model.

involving consumption paint a very consistent picture with respect to how much insurance households achieve against idiosyncratic risk. This finding is reassuring from the standpoint of theory and strengthens the case for using labor supply moments in future studies of risk sharing – especially given the high quality and long panel dimension of existing datasets that record hours worked.

We have conducted our analysis within a simple static model of labor supply in which wages are exogenous. While this is a natural starting point to explore how micro data on labor supply can inform the study of risk sharing, one interesting direction for future work would be to consider dynamic models in which current hours worked affect future wages. In a stripped down version of our framework, we have experimented with introducing learning by doing along the lines of Imai and Keane (2004). In one parametric special case, insurable and uninsurable shocks turn out to have exactly the same effects on labor supply and consumption as in the benchmark specification. Thus, the identification of model parameters, including the degree of insurance, remains valid. The only difference relative to the benchmark model is that wages now have an endogenous component. In particular, transitory insurable shocks have a permanent effect on wages because working more hours today raises future productivity. This analysis suggests that introducing learning by doing is one way to micro-found the existence of permanent insurable shocks.[54]

5.6 Robustness and statistical fit

We now examine the robustness of our estimates with respect to (i) the statistical model for the variances of the innovations η and ω, and (ii) the choice of the weighting matrix used in estimation.

We begin by estimating a version of our model where $v_{\eta t}$ and $v_{\omega t}$ follow unrestricted time sequences instead of fourth-order polynomials of time. The results, reported in column (2) of Table 4, show that parameter estimates are remarkably similar across the two versions of the model. Figure E3 in the Technical Appendix compares the two time series for the variances. The polynomials capture the main low-frequency dynamics of the two series while avoiding the abrupt fluctuations from one year to the next and the many zero boundary values that are features of the unconstrained sequences.

The first alternative weighing scheme that we explore is one where, rather than giving each moment equal weight, we weigh each moment by its number of observations – a scheme

[54] More details of this extension are available upon request.

Table 4: Parameter Estimates: Robustness on Baseline Model Estimates

	Baseline	(2)	(3)	(4)	(5)	(6)
σ	2.165	2.139	2.029	2.443	1.981	2.077
γ	1.713	1.691	1.715	1.745	1.610	1.747
$v_{\mu h}$	0.036	0.036	0.037	0.027	0.037	0.038
$v_{\mu y}$	0.000	0.000	0.000	0.000	0.000	0.000
$v_{\mu c}$	0.041	0.041	0.042	0.033	0.039	0.043
$\overline{v}_{\widehat{\varphi}}$	0.054	0.054	0.054	0.052	0.048	0.063
$\overline{v}_{\alpha^0}$	0.102	0.102	0.088	0.106	0.105	0.083
$\overline{v}_{\omega}$	0.0056	0.0057	0.0070	0.0056	0.0055	0.0064
$\overline{v}_{\kappa^0}$	0.047	0.047	0.053	0.047	0.048	0.056
$\overline{v}_{\eta}$	0.0044	0.0044	0.0045	0.0051	0.0043	0.0056
$\overline{v}_{\theta}$	0.043	0.042	0.040	0.039	0.042	0.045
$1/\widehat{\sigma}$	0.347	0.351	0.368	0.310	0.376	0.364
p- value of OID test	0.99	–	–	–	–	0.99

Notes: The baseline estimates are reproduced from Table 1. Other columns: (2) unrestricted sequences for $v_{\eta t}$ and $v_{\omega t}$, (3) each moment weighted by its number of observations, (4) each moment weighted by the inverse of the corresponding element on the diagonal of the fourth moment matrix, (5) weighting scheme that realigns (absolute) values of moments. Variables with bars (e.g., $\overline{v}_{\theta}$) denote average estimates over the sample period, (6) minimum distance estimation on collapsed set of moments. The OID tests for the models in columns (1) and (6) have χ^2 distributions with $11,368$ and 580 degrees of freedom, respectively.

that puts more weight on the PSID moments and on the moments in levels. The second alternative is a weighting matrix with elements given by the inverse of the diagonal of the fourth-moment matrix described in the Technical Appendix.[55] The third alternative weighing matrix divides every variance at age/year (a, t) by its sample average value, and every covariance between pairs of variables (x, y) at age/year (a, t) by the product of the sample average of the standard deviations of x and y.[56] This addresses a potential concern that under our baseline weighting scheme (the identity matrix) moments whose values are large on average (e.g., the variance of log wages) will receive more weight than moments whose values are closer to zero on average (e.g., the variance of changes in log hours), since

[55] The square roots of the elements in this matrix provide the standard errors of the corresponding elements in the vector of empirical moments. Hence, this weighting matrix gives more emphasis to moments measured more precisely. As discussed in Blundell et al. (2008), this method avoids the pitfalls of using the full optimal weighting matrix described by Altonji and Segal (1996), which are primarily related to the terms outside the main diagonal.

[56] Effectively, we match correlations instead of covariances. Dividing the covariances at (a, t) by their sample average is not feasible because some of them are too close to zero.

the estimation algorithm minimizes the sum of squared residuals between empirical and theoretical moments.

Estimation results under these three alternative schemes are reported in columns (3)-(5) of Table 4. Point estimates are always within two standard deviations of the benchmark, and often much closer. Also the time paths of all the variances are very similar to those of the baseline model.

The last row of Table 4 presents a test of the overidentifying restrictions (OID), a χ^2 statistic with degrees of freedom equal to the number of moments in excess of the number of parameters. When the test is performed on the baseline set of moments (and hence, with 11,368 degrees of freedom), the p-value is near 1 indicating that the structural model cannot be rejected. In the context of dynamic panel data models, Bowsher (2002) reports severe loss of power of OID tests when the number of overidentifying restrictions is large relative to the number of observations used to calculate each moment.[57] To strengthen test power, one has to reduce the number of restrictions. We therefore collapse our full set of age/year moments into unconditional moments by age and by year.[58] When the model is re-estimated on this smaller set of moments we achieve very similar point estimates for all parameters (column (6) of Table 4). More importantly, the OID test performed on this subset of restrictions still returns a large p-value, above 0.99. This suggests that the model cannot be rejected and that it fits the data quite well in a purely statistical sense. The Technical Appendix contains a detailed description of how we compute the test statistics.

6 Conclusion

In this paper, we have developed a novel theoretical framework to analyze consumption and labor supply in the presence of idiosyncratic labor income shocks. A distinguishing feature of the model is that it can be solved analytically. Tractability is achieved by extending the environment of Constantinides and Duffie (1996) to incorporate flexible labor supply, partially insurable wage risk, progressive taxation, and heterogeneity in the taste for leisure.

[57]Monte Carlo simulations in Bowsher (2002) show that inference becomes misleading as soon as the number of overidentifying restrictions approaches the number of observations. In our case, the average number of observations used in PSID is 285 and in CEX is 179.

[58]See Roodman (2009) for a discussion of this "collapsed instruments" technique and a list of applications. After collapsing the moments, we end up with 580 overidentifying restrictions, and an average number of observations per moment of 8,291 in PSID and 5,318 in CEX. Note that Roodman (2009) also reports that unbiasedness of the estimates is not affected by this instrument proliferation; if anything estimates are slightly more precise.

From the closed-form equilibrium allocations, it is straightforward to derive expressions for the cross-sectional (co-)variances of wages, hours, and consumption. These expressions allow, in turn, a formal identification proof and facilitate the estimation of the structural parameters. We used this framework (i) to measure the extent to which US households can insure against wage risk, (ii) to quantify how risk sharing has changed over the past 40 years –a period of sharp widening in the wage distribution, and (iii) to decompose the sources of cross-sectional inequality in wages, hours and consumption.

This paper takes a first step toward showing how labor supply can help identify insurance, an exercise usually done with consumption data only. The framework could be extended to incorporate a participation decision along the extensive margin. For example, with a minimum requirement on hours worked per period, equilibrium labor supply allocations would feature a threshold such that low wage workers do not work. Combining evidence on both the extensive and intensive margins would, in principle, bring even more information to bear on the nature of risk and insurance. Future work should also study how labor supply data can inform the study of risk-sharing in dynamic models for labor supply, where current hours affect future wages thanks to some form of human capital accumulation.

The theoretical framework can be extended to shed light on a range of macroeconomic questions where heterogeneity and risk are central to the analysis. In Heathcote, Storesletten, and Violante (2010a), we use a version of the model to explore the optimal degree of progressivity in the tax schedule, focusing on how the optimal degree of public redistribution varies with the fraction of wage risk that can be insured privately, the desire for public goods, and the elasticity of labor supply. In Heathcote, Storesletten, and Violante (2011a), we extend the model to incorporate an education choice, and quantify the welfare effect of the observed increase in the college premium, alongside the observed rise in wage risk within education groups. Finally, it is also possible to introduce aggregate shocks that are correlated with the variance of idiosyncratic risk, as in Constantinides and Duffie (1996) and Storesletten, Telmer, and Yaron (2004b), and non time-separable Epstein-Zin preferences. Such an extended setup is a natural environment for studying asset pricing, and the welfare costs of business cycles.

Many of these issues have been extensively explored using conventional incomplete-markets models and numerical solution methods. The reason to revisit them is that our framework remains tractable when extended along these dimensions, making the economic forces at play transparent and readily quantifiable.

References

Abowd, J., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

Aguiar, M., and M. Bils (2011): "Has Consumption Inequality Mirrored Income Inequality?," NBER Working Paper 16807.

Ales, L., and P. Maziero (2009): "Accounting for Private Information," Mimeo.

Ales, L., and P. Maziero (2011): "Non-Exclusive Dynamic Contracts: Competition and the Limits of Insurance," Mimeo.

Altonji, J., and L. M. Segal (1996): "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 14(3), 353–366.

Angeletos, G.-M. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10(1), 1–30.

Angelucci, M., G. de Giorgio, and I. Rasul (2012): "Resource Pooling Within Family Networks: Insurance and Investment," Mimeo, University of Michigan.

Attanasio, O., E. Battistin, and H. Ichimura (2007): "What Really Happened to Consumption Inequality in the United States?," in *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, ed. by E. R. Berndt, and C. R. Hulten, vol. 67 of *Studies in Income and Wealth*, chap. 17, pp. 515–544. University of Chicago Press.

Attanasio, O., and S. J. Davis (1996): "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy*, 104(6), 1227–1262.

Attanasio, O., and N. Pavoni (2011): "Risk Sharing in Private Information Models with Asset Accumulation: Explaining the Excess Smoothness of Consumption," *Econometrica*, 79(4), 1027–1068.

Attanasio, O., and J.-V. Ríos-Rull (2000): "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?," *European Economic Review*, 44(7), 1225–1258.

Baxter, M., and M. J. Crucini (1995): "Business Cycles and the Asset Structure of Foreign Trade," *International Economic Review*, 36(4), 821–854.

Benabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5), 1887–1921.

Blundell, R., and I. Preston (1998): "Consumption Inequality and Income Uncertainty," *Quarterly Journal of Economics*, 113(2), 603–640.

Bowsher, C. (2002): "On Testing Overidentifying Restrictions in Dynamic Panel Data Models," *Economics Letters*, 77, 211–220.

Caballero, R. (1990): "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics*, 25(1), 113–136.

Campbell, J. Y., and A. Deaton (1989): "Why Is Consumption So Smooth?," *Review of Economic Studies*, 56(3), 357–373.

Chamberlain, G. (1984): "Panel Data," in *Handbook of Econometrics*, ed. by Z. Griliches, and M. Intriligator, vol. 3, pp. 1247–1318. New York: North-Holland.

Chetty, R. (2006): "A New Method of Estimating Risk Aversion," *American Economic Review*, 96(5), 1821–1834.

Cole, H., and N. Kocherlakota (2001): "Efficient Allocations with Hidden Income and Hidden Storage," *Review of Economic Studies*, 68(3), 523–542.

Constantinides, G. M., and D. Duffie (1996): "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 104(2), 219–240.

Cunha, F., J. J. Heckman, and S. Navarro (2005): "Separating Uncertainty from Heterogeneity in Life-Cycle Earnings," *Oxford Economic Papers*, 57(2), 191–261.

Davis, S. J. (2003): "Comment on Krueger and Perri: On the Welfare Consequences of the Increase in Inequality in the United States," *NBER Macroeconomics Annual*, 18, 121–132.

Deaton, A. (1997): *The Analysis of Household Surveys*. Johns Hopkins University Press.

Gottschalk, P., and M. Huynh (2010): "Are Earnings Inequality and Mobility Overstated? The Impact of Nonclassical Measurement Error," *Review of Economics and Statistics*, 92(2), 302–315.

Guvenen, F., and A. A. Smith (2010): "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach," Mimeo.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13(1), 15–51.

Heathcote, J., K. Storesletten, and G. L. Violante (2009): "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 1(1), 319–354.

——— (2010a): "Redistributive Taxation in a Partial-Insurance Economy," Mimeo.

——— (2010b): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

——— (2011a): "From Wages to Welfare: Decomposing Gains and Losses from Rising Inequality," Mimeo.

——— (2011b): "A Note on the Partial Insurance Framework with General Preferences," Mimeo.

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101(7), 2923–2954.

Imai, S., and M. P. Keane (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45(2), 601–641.

Jappelli, T., and L. Pistaferri (2010): "The Consumption Response to Income Changes," *Annual Review of Economics*, 2(1), 479–506.

Kaplan, G., and G. L. Violante (2010): "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2(4), 53–87.

Keane, M. P. (2011): "Labor Supply and Taxes: A Survey," *Journal of Economic Literature*, 49(4), 961–1075.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

Krebs, T. (2003): "Human Capital Risk and Economic Growth," *Quarterly Journal of Economics*, 118(2), 709–744.

Krueger, D., and F. Perri (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73(1), 163–193.

——— (2010): "How do households respond to income shocks?," Mimeo.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life-cycle," *American Economic Review*, 100(4), 1432–1467.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18(1), 83–114.

Moffitt, R., and P. Gottschalk (2002): "Trends in the Transitory Variance of Earnings in the United States," *Economic Journal*, 112(3), C68–C73.

Roodman, D. M. (2009): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.

Slesnick, D. T. (2001): *Consumption and Social Welfare: Living Standards and Their Distribution in the United States*. Cambridge University Press.

Storesletten, K., C. I. Telmer, and A. Yaron (2004a): "Consumption and Risk Sharing Over the Life Cycle," *Journal of Monetary Economics*, 51(3), 609–633.

——— (2004b): "Cyclical Dynamics in Idiosyncratic Labor-Market Risk," *Journal of Political Economy*, 112(3), 695–717.

Swanson, E. T. (2012): "Risk Aversion and the Labor Margin in Dynamic Equilibrium Models," *American Economic Review*, 102(4), 1663–1691.

Uhlig, H. (1996): "A Law of Large Numbers for Large Economies," *Journal of Economic Theory*, 8(1), 41–50.

A Appendix for Publication

A.1 Proof of Proposition 1

The proof is in two parts. In the first we describe a planner's problem and show that the solution to this problem is the allocations for consumption and hours described in Proposition 1, part (ii). In the second, we decentralize these allocations in a competitive equilibrium and show that the asset prices described in Proposition 1, part (iii), and the no-inter-island-trade result described in part (i) form part of this decentralization. In what follows, we omit some technical details the proof. See Technical Appendix A for a complete derivation.

Planner's allocations: We first solve for equilibrium allocations for consumption and hours worked by solving a set of static planning problems. Each island-level planner maximizes equally weighted period utility for a set of agents that share a common age a, a common preference weight φ, and a common wage component α_t. Let $x_t = (a, \varphi, \alpha_t)$ denote these island-level components of the individual state. Each island-level planner controls a set of agents with the age-specific population distributions for the wage components $F^a_{\kappa,t}$ and $F_{\theta,t}$. Let $F^a_{\varepsilon,t}$ denote the implied age-specific distribution over $\kappa_t + \theta_t$. The planner's problem on an island defined by x_t is to choose functions $c_t(x_t, \varepsilon_t)$, $h_t(x_t, \varepsilon_t)$ to solve

$$\max_{\{c_t(x_t,\cdot),h_t(x_t,\cdot)\}} \int \left[\frac{c_t(x_t,\varepsilon_t)^{1-\gamma} - 1}{1-\gamma} - \exp(\varphi) \frac{h_t(x_t,\varepsilon_t)^{1+\sigma}}{1+\sigma} \right] dF^a_{\varepsilon,t}$$

subject to the island-level resource constraint

$$\int \Big[\lambda \left(\exp\left(\alpha_t + \varepsilon_t\right) h_t(x_t,\varepsilon_t)\right)^{1-\tau} - c_t(x_t,\varepsilon_t) \Big] dF^a_{\varepsilon,t} = 0. \tag{1}$$

Combine the first-order conditions with respect to c_t and h_t to get

$$h_t(x_t,\varepsilon_t) = \left((1-\tau)\,\lambda\right)^{\frac{1}{\sigma+\tau}} c_t(x_t,\varepsilon_t)^{-\frac{\gamma}{\sigma+\tau}} \exp\left((\alpha_t+\varepsilon_t) \left(\frac{1-\tau}{\sigma+\tau} \right) - \frac{1}{\sigma+\tau} - \frac{\varphi}{\sigma+\tau} \right). \tag{2}$$

Substituting (2) into (1), using the definition for the tax-modified Frisch elasticity $\widehat{\sigma} = (\sigma+\tau)/(1-\tau)$, and rearranging yields the expressions for c_t and h_t in eq. (7)-(8), where C^a_t and H^a_t are constants common to all agents of age a in year t given by

$$\begin{aligned}
\mathcal{C}^a_t &= \frac{1}{\widehat{\sigma}+\gamma} \left((1+\widehat{\sigma})\log\lambda + \log(1-\tau)\right) + \mathcal{M}^a_t \\
\mathcal{H}^a_t &\equiv \frac{1}{(1-\tau)\left(\widehat{\sigma}+\gamma\right)} \left((1-\gamma)\log\lambda + \log(1-\tau)\right) - \frac{\gamma}{\widehat{\sigma}(1-\tau)} \mathcal{M}^a_t \\
\mathcal{M}^a_t &= \frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma} \log \int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}} \varepsilon_t \right) dF^a_{\varepsilon,t}.
\end{aligned}$$

Decentralization (prices): To decentralize the solution to the above planner's problem, we start by conjecturing prices in this equilibrium. Pre-tax wages equal individual labor productivity, $w(x_t, \varepsilon_t) = \exp(\alpha_t + \varepsilon_t)$. At this wage, the intratemporal first-order condition from the agent's problem described in Section 2.1 is identical to the intra-temporal first order condition for the planner described in eq. (2). Thus at competitive wages and the conjectured allocations (eqs. 7 and 8), agents are optimizing on the intra-temporal margin. At first blush this might seem surprising, given the presence of progressive earnings taxation in the economy. Recall, however, that individual agents (in the competitive equilibrium) and island-level planners (in the problem described above) are both atomistic and take the tax/transfer system parameters as exogenous.

To conjecture equilibrium prices for intertemporal insurance claims, it is convenient to revert to history-dependent notation and write $c_t(s^t)$ rather than $c_t(x_t, \varepsilon_t)$. We begin with the price of within-island insurance $Q_t\left(S; s^t\right)$. The intertemporal first-order condition from the agent's problem (Section 2.1) defines the price at which an agent of age a with history s^t is willing, on the margin, to buy or sell a set of insurance contracts $B_t(S; s^t)$ that pay δ^{-1} units of consumption if and only if $s_{t+1} = (\omega_{t+1}, \eta_{t+1}, \theta_{t+1}) \in S \subseteq S$. This price is simply the average marginal rate of substitution in those states. Substituting in the expression for consumption (7) yields the expression for $Q_t\left(S; s^t\right)$ in eq. (9) of Proposition 1. Thus the prices $Q_t(S; s^t)$ are consistent with optimization on the consumer side.

Note that $Q_t\left(S; s^t\right) = Q_t(S)$: insurance prices are independent of the individual history s^t and age a. From eq. (9) there are two pieces to this result. First, $F_{s,t+1}$, the joint distribution over $s_{t+1} = (\omega_{t+1}, \eta_{t+1}, \theta_{t+1})$ at $t+1$, is independent of s^t and thus the second term in eq. (9) is independent of s^t. Second, insurance prices are also independent of age a, because the growth in average consumption $\exp\left(\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a\right)$ is independent of age, reflecting the permanent-transitory model for individual productivity dynamics. Note also that due to full insurance against $(\eta_{t+1}, \theta_{t+1})$, the price of insurance against η_{t+1} and θ_{t+1} simply reflects probabilities, while the price of insurance against ω_{t+1} also reflects the conditional marginal rate of substitution, with insurance against low ω_{t+1} realizations being more expensive than equally likely high ω_{t+1} realizations.

We now turn to the price function for insurance claims traded across islands. Because any contract that can be traded between islands can also be traded within an island, the inter-island price for a claim that pays δ^{-1} units of consumption iff $s_{t+1} \in Z$ must, by arbitrage, equal the corresponding within-island price, for any Z. This implies $Q_t^*\left(Z; s^t\right) = \Pr\left((\eta_{t+1}, \theta_{t+1}) \in Z\right) \times Q_t(S) = Q_t^*(Z)$, where $Q_t(\mathbb{S})$ is the price of insurance against all states (i.e., a risk-free bond). Thus these prices are just probabilities times $Q_t(\mathbb{S})$.

Decentralization (asset purchases): We now derive asset purchases, $B_t(s_{t+1}; s^t)$ and $B_t^*(\eta_{t+1}, \theta_{t+1}; s^t)$ and verify that agents' budget constraints are satisfied in equilibrium.

Given that any available inter-island insurance contract can be purchased at the same price on the within-island market, $B_t^*(\eta_{t+1}, \theta_{t+1}; s^t) = 0$ for all $(\eta_{t+1}, \theta_{t+1})$ is consistent with individual optimization (Proposition 1, part (iii)). Thus, agents are optimizing by purchasing all their insurance on the island on which they are located. At the same time, because $Q_t^*\left(Z; s^t\right) = Q_t^*(Z)$, no agent has an incentive to try to sell insurance to an agent located on another island. To understand this, note that the price at which one agent (say agent i_1) with history $s_{i_1}^t$ is willing to buy, on the margin, a set of claims that pay if and only if $(\eta_{t+1}, \theta_{t+1}) \in Z$ is the probability of that event times agent i_1's expected marginal rate of substitution, i.e., $\Pr\left((\eta_{t+1}, \theta_{t+1}) \in Z\right) \times Q_t\left(\mathbb{S}; s_{i_1}^t\right)$. The price at which a second agent on a different island (agent i_2 with history $s_{i_2}^t$) is willing to sell this insurance to agent i_1 is the same probability times agent i_2's expected marginal rate of substitution, $\Pr\left((\eta_{t+1}, \theta_{t+1}) \in Z\right) \times Q_t\left(\mathbb{S}; s_{i_2}^t\right)$. If agents i_1 and i_2 did not share the same marginal rate of substitution (i.e., if $Q_t\left(\mathbb{S}; s_{i_1}^t\right) \neq Q_t\left(\mathbb{S}; s_{i_2}^t\right)$), then there could be no equilibrium without inter-island trade, because any such equilibrium would feature unexploited gains from trade. Thus, $Q_t(S, s^t) = Q_t(S)$ is the crucial result supporting an absence of inter-island trade.

Finally, we now derive an expression for purchases of state-contingent claims, $B_t(s_{t+1}; s^t)$, and verify budget balance. Given $B_t^*(Z; s^t) = 0\, \forall Z, \forall s^t$, realized wealth at s^t implicitly defines insurance purchases $B_{t-1}(s_t; s^{t-1}) = \delta d_t(s^t)$. Since insurance payouts must deliver the discounted present value of lifetime differences between consumption and after-tax earnings, the realized wealth must be

$$d_t(s^t) = T_t\left(s^t\right) + \mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty} \frac{(\beta\delta)^j\, c_{t+j}(s^{t+j})^{-\gamma}}{c_t(s^t)^{-\gamma}} T_{t+j}\left(s^{t+j}\right)\right],$$

where $T_{t+j}\left(s^{t+j}\right) \equiv c_{t+j}\left(s^{t+j}\right) - \lambda\left(w\left(s^{t+j}\right)h_{t+j}\left(s^{t+j}\right)\right)^{1-\tau}$ is the net transfer in period $t+j$.

Given this guess for $d_t(s^t)$, it is straightforward to verify that the agent's budget constraint is satisfied (see Technical Appendix A for a complete derivation).

A.2 The household model of Section 3

Full insurance against (g,k): Assume that utility for individual i in a household of g adult workers and k children is

$$u(c,h_i,g,k) = \frac{1}{1-\gamma}\left(\frac{c}{e(g,k)}\right)^{1-\gamma} - \frac{\exp(\varphi)}{1+\sigma}h_i^{1+\sigma},$$

where c is household consumption and h_i is agent i's hours worked. The household attaches equal weights to all adults and no weight to the children.

As in Section A.1, let $x_t = (a,\varphi,\alpha_t)$ denote the island-level components of the individual state. The planner can insure against realizations of ε_t, g, and k. The planner problem is to choose functions $c_t(x_t,g,k)$ and $h_{it}(x_t,\varepsilon_t,g,k)$ for $i=1,...,g$ to solve

$$\max_{\{c_t(x_t,\cdot),h_{it}(x_t,\cdot)\}} \int\left[\frac{g}{1-\gamma}\left(\frac{c_t(x_t,g,k)}{e(g,k)}\right)^{1-\gamma} - \sum_{i=1}^{g}\int\frac{\exp(\varphi)}{1+\sigma}h_{it}(x_t,\varepsilon_t,g,k)^{1+\sigma}\,dF_{\varepsilon t}^a\right]dF_t(g,k), \quad (3)$$

subject to the island-level after-tax resource constraint

$$\int\left[c_t(x_t,g,k) - \sum_{i=1}^{g}\int\lambda\left[\exp(\alpha_t+\varepsilon_t)\,h_{it}(x_t,\varepsilon_t,g,k)\right]^{1-\tau}\,dF_{\varepsilon t}^a\right]dF_t(g,k) = 0, \quad (4)$$

where eqs. (3)-(4) incorporate the within-island distribution $F_t(g,k)$ of household workers and children, and where, based on the result in Section A.1, we have already let consumption be independent of ε_t.

The first-order condition w.r.t. c_t implies that consumption for a (g,k) household is

$$c_t(x_t,g,k) = c_t(x_t,1,0)\left(\frac{g}{e(g,k)^{1-\gamma}}\right)^{1/\gamma}. \quad (5)$$

Combine the first-order conditions w.r.t. c_t and h_{it} with eqs. (4)-(5) to derive an expression for $c_t(x_t,1,0)$. Define $D(g,k) \equiv (\log g)/\gamma - (1-\gamma)/\gamma\,\log(e(g,k))$. Then use eq. (5) and the definitions for $\widehat{\sigma}$ and $\widehat{\varphi}$ to derive the equilibrium allocations for household consumption and individual hours,

$$\log c_t(x_t,g,k) = D(g,k) - (1-\tau)\widehat{\varphi} + (1-\tau)\left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)\alpha_t + \mathcal{C}_t^a \quad (6)$$

$$\log h_{it}(x_t,\varepsilon_t,g,k) = -\widehat{\varphi} + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)\alpha_t + \frac{1}{\widehat{\sigma}}\varepsilon_{it} + \mathcal{H}_t^a, \quad (7)$$

where expressions for C_t^a and H_t^a are in Technical Appendix B.1. Note that hours do not depend on (g,k)

No insurance against (g,k): Consider now the model without insurance against household type. In this model, there is no within-island variation in household composition (g,k). Thus the island-level components of the individual state are $x_t = (a, \varphi, \alpha_t, g, k)$, and the island planner problem corresponding to the competitive equilibrium is to choose a number $c_t(x_t)$ and functions $h_{it}(x_t, \varepsilon_t)$ for $i = 1, ..., g$ to solve

$$\begin{aligned}
&\max_{c_t(x_t),\{h_{it}(x_t,,\cdot)\}} \left\{ \frac{g}{1-\gamma}\left(\frac{c_t(x_t)}{e(g,k)}\right)^{1-\gamma} - \sum_{i=1}^{g}\int \frac{\exp(\varphi)}{1+\sigma} h_{it}(x_t,\varepsilon_t)^{1+\sigma}\, dF^a_{\varepsilon t}\right\} \\
\text{s.t.} \quad & c_t(x_t) - \sum_{i=1}^{g}\int \lambda\left[\exp(\alpha_t+\varepsilon_{it})\, h_{it}(x_t,\varepsilon_t)\right]^{1-\tau} dF^a_{\varepsilon t} \;=\; 0.
\end{aligned}$$

In Technical Appendix B.2 we derive the following allocations:

$$\begin{aligned}
\log c_t(x_t) &= D^c(g,k) - (1-\tau)\,\widehat{\varphi} + (1-\tau)\left(\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\right)\alpha_t + \mathcal{C}^a_t \\
\log h_{it}(x_t,\varepsilon_t) &= D^h(g,k) - \widehat{\varphi} + \frac{1-\gamma}{\widehat{\sigma}+\gamma}\alpha_t + \frac{\kappa_{it}+\theta_{it}}{\widehat{\sigma}} + \mathcal{H}^a_t,
\end{aligned}$$

where the equivalization dummies are $D^c(g,k) = ((1+\widehat{\sigma})\log g - (1-\gamma)\log e(g,k))/(\widehat{\sigma}+\gamma)$ and $D^h(g,k) = (D^c(g,k) - \log g)/(1-\tau)$.

A.3 Proofs of Identification

A.3.1 Proof of Proposition 2

The proof is organized in four recursive steps.

Step A. The four (sets of) parameters $\widehat{\sigma}$, γ, $\{v_{\eta t} + \Delta v_{\theta t}\}_{t=2}^{T}$, $\{v_{\omega t}\}_{t=2}^{T}$ are identified from within-cohort changes in the macro moments, $\Delta var^a_t(\log \hat{w})$, $\Delta var^a_t(\log \hat{h})$, $\Delta var^a_t(\log \hat{c})$, and $\Delta cov^a_t(\log \hat{w}, \log \hat{c})$, all available from $t = 2, ..., T$. These parameters are identified recursively as follows. Each element of the sequence $\{v_{\omega t}\}_{t=2}^{T}$ is identified by:

$$\Delta cov^a_t(\log \hat{w}, \log \hat{c})^2 / \Delta var^a_t(\log \hat{c}) = v_{\omega t}.$$

Given $v_{\omega t}$, each element of the sequence $\{v_{\eta t} + \Delta v_{\theta t}\}_{t=2}^{T}$ is identified by

$$\Delta var^a_t(\log \hat{w}) = v_{\omega t} + (v_{\eta t} + \Delta v_{\theta t}).$$

Given $v_{\omega t}$ and $v_{\eta t} + \Delta v_{\theta t}$, the tax-modified Frisch elasticity $\widehat{\sigma}$ is identified by

$$\Delta var^a_t(\log \hat{h}) = [\Delta cov^a_t(\log \hat{h}, \log \hat{c})/\Delta cov^a_t(\log \hat{w}, \log \hat{c})]^2 v_{\omega t} + 1/\widehat{\sigma}^2\,(v_{\eta t} + \Delta v_{\theta t}).$$

Given $\widehat{\sigma}$, the parameter γ is identified by

$$\Delta cov^a_t(\log \hat{h}, \log \hat{c})/\Delta cov^a_t(\log \hat{w}, \log \hat{c}) = (1-\gamma)/(\widehat{\sigma}+\gamma).$$

Step B. Since $\widehat{\sigma}$ is known, the variances of transitory insurable shocks $\{v_{\theta t}\}_{t=1}^{T-1}$ are identified from the difference between the dispersion in growth rates ("micro moments") and the growth rate of within-cohort dispersion ("macro moments") available from $t = 2, ..., T$:

$$cov^a_t(\Delta \log \hat{w}, \Delta \log \hat{h}) + var^a_t(\Delta \log \hat{h}) - \Delta cov^a_t(\log \hat{w}, \log \hat{h}) - \Delta var^a_t(\log \hat{h}) = 2(1+\widehat{\sigma})/\widehat{\sigma}^2\, v_{\theta,t-1}.$$

Combining the sequence $\{v_{\theta t}\}_{t=1}^{T-1}$ with $\{v_{\eta t}+\Delta v_{\theta t}\}_{t=2}^{T}$ identifies $\{v_{\eta t}\}_{t=2}^{T-1}$. Substituting the value for $v_{\theta,T-1}$ into $(v_{\eta T}+\Delta v_{\theta T})$ from Step A identifies $(v_{\eta T}+v_{\theta T})$.

Step C. Since $\widehat{\sigma}$ and γ are known, the following moments, available for all $t=1,...,T$ and evaluated for the youngest age group, identify the cohort effects sequence $\left\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\right\}_{t=1}^{T}$:

$$\begin{aligned} cov_t^0(\log\hat{w},\log\hat{c}) &= (1-\tau)(1+\widehat{\sigma})/(\widehat{\sigma}+\gamma)\ v_{\alpha^0 t} \\ cov_t^0(\log\hat{h},\log\hat{c}) &= (1-\tau)v_{\widehat{\varphi}t}+(1-\tau)(1+\widehat{\sigma})(1-\gamma)/(\widehat{\sigma}+\gamma)^2\ v_{\alpha^0 t}. \end{aligned}$$

Then $\{v_{\kappa^0 t}\}_{t=1}^{T-1}$ and $(v_{\kappa^0 T}+v_{\theta T})$ are identified from

$$cov_t^0(\log\hat{w},\log\hat{h})+var_t^0(\log\hat{h})=v_{\widehat{\varphi}t}+(1-\gamma)(1+\widehat{\sigma})/(\widehat{\sigma}+\gamma)^2\ v_{\alpha^0 t}+(1+\widehat{\sigma})/\widehat{\sigma}^2\ (v_{\kappa^0 t}+v_{\theta t}).$$

Step D. Finally, the variances of measurement error $\{v_{\mu y}, v_{\mu h}, v_{\mu c}\}$ are identified from the following moments in levels, for example those corresponding to the youngest age group:

$$\begin{aligned} cov_t^0(\log\hat{w},\log\hat{h}) &= (1-\gamma)/(\widehat{\sigma}+\gamma)\ v_{\alpha^0 t}+1/\widehat{\sigma}\ (v_{\kappa^0 t}+v_{\theta t})-v_{\mu h} \\ var_t^0(\log\hat{w}) &= v_{\alpha^0 t}+(v_{\kappa^0 t}+v_{\theta t})+v_{\mu y}+v_{\mu h} \\ var_t^0(\log\hat{c}) &= (1-\tau)^2 v_{\widehat{\varphi}t}+(1-\tau)^2(1+\widehat{\sigma})^2/(\widehat{\sigma}+\gamma)^2\ v_{\alpha^0 t}+v_{\mu c}. \end{aligned}$$

A.3.2 Proof of Corollary 2.1

At dates $t=\hat{t},\hat{t}+1$, the data availability is the same as in Proposition 2, and hence one can identify $v_{\mu y}$. Applying Proposition 3 to dates other than $(\hat{t},\hat{t}+1)$ when only wage and hours data are available, the whole model is then identified.

A.3.3 Proof of Corollary 2.2

From Proposition 2 we identify the parameters $\{\widehat{\sigma},\gamma,v_{\mu h},v_{\mu y},v_{\mu c}\}$, the sequences $\left\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\right\}_{t=1}^{\hat{t}}$, $\{v_{\kappa^0 t}, v_{\theta t}\}_{t=1}^{\hat{t}-1}$, $\{v_{\omega t}, v_{\eta t}\}_{t=2}^{\hat{t}}$, and the sums $v_{\eta,\hat{t}}+v_{\theta,\hat{t}}$ and $v_{\kappa^0\hat{t}}+v_{\theta\hat{t}}$.

From the cross-sectional moment $\Delta var_t^a(\log\hat{c})=(1-\tau)^2(1+\widehat{\sigma})^2/(\widehat{\sigma}+\gamma)^2\ v_{\omega t}$, which is available every year, we can identify $\{v_{\omega t}\}_{t=\hat{t}+1}^{T}$. We identify the cohort effects $\left\{v_{\widehat{\varphi}t}, v_{\alpha^0 t}\right\}_{t=\hat{t}+1}^{T}$ from the moments, available in every year,

$$\begin{aligned} cov_t^0(\log\hat{w},\log\hat{c}) &= (1-\tau)(1+\widehat{\sigma})/(\widehat{\sigma}+\gamma)\ v_{\alpha^0 t} \\ cov_t^0(\log\hat{h},\log\hat{c}) &= (1-\tau)v_{\widehat{\varphi}t}+(1-\tau)(1+\widehat{\sigma})(1-\gamma)/(\widehat{\sigma}+\gamma)^2\ v_{\alpha^0 t}. \end{aligned}$$

By combining the moments

$$cov_t^a(\Delta^2\log\hat{w},\Delta^2\log\hat{h})+var_t^a(\Delta^2\log\hat{h})-\Delta^2 cov_t^a(\log\hat{w},\log\hat{h})-\Delta^2 var_t^a(\log\hat{h})=2(1+\widehat{\sigma})/\widehat{\sigma}^2\ v_{\theta,t-2},$$

we identify $\{v_{\theta t}\}$ for the biannual years $t=\hat{t},\hat{t}+2,\hat{t}+4,...,T-2$. Note that, since $v_{\theta,\hat{t}}$ is identified, so are $v_{\eta,\hat{t}}$ and $v_{\kappa^0\hat{t}}$. From $\Delta^2 var_t^a(\log\hat{w})=v_{\omega t}+v_{\omega,t-1}+(v_{\eta t}+v_{\eta,t-1}+v_{\theta t}-v_{\theta,t-2})$, available for $t=\hat{t},\hat{t}+2,...,T$, we can identify the sum $\left\{v_{\eta t}+v_{\eta,t-1}+\Delta^2 v_{\theta t}\right\}$. This, together with the sequence $\{v_{\theta,t}\}$, available for $t=\hat{t},\hat{t}+2,...,T$, allows us to identify $\{v_{\eta t}+v_{\eta,t-1}\}$ for the biannual years $t=\hat{t},\hat{t}+2,\hat{t}+4,...,T-2$, as well as $\{v_{\eta T}+v_{\eta,T-1}+v_{\theta T}\}$. Finally, consider the moment

$$var_t^0(\log\hat{w})=v_{\alpha^0 t}+(v_{\kappa^0 t}+v_{\theta t})+v_{\mu y}+v_{\mu h}.$$

This moment is available for the biannual years and identifies $\{v_{\kappa^0 t}\}$ for $t=\hat{t},\hat{t}+2,\hat{t}+4,...,T-2$ and $v_{\kappa^0,T}+v_{\theta T}$.

A.3.4 Proof of Corollary 2.3

Step A of Proposition 2 shows that $T-1$ realizations of $v_{\omega t}$ are identified, and hence one can uniquely identify all the coefficients of a time polynomial of order $T-2$ or lower and recover the entire time-series $\{v_{\omega t}\}_{t=1}^{T}$. From Step B of of Proposition 2, $T-2$ realizations of $v_{\eta t}$ are identified, and in the same vein one can uniquely identify all the coefficients of a time polynomial of order $T-3$ or lower and recover the entire time-series $\{v_{\eta t}\}_{t=1}^{T}$. Then, from Step A and Step B, it can be seen that the whole sequence $\{v_{\theta t}\}_{t=1}^{T}$ is identified. As a result, from Step C, one can identify the entire time-series $\{v_{\kappa^0 t}\}_{t=1}^{T}$. The rest of the parameter vector is identified exactly as in Proposition 2.

A.4 Additional identifying assumptions

When we model $v_{\eta t}$ and $v_{\omega t}$ as time-polynomials, we make the following two additional assumptions to complete identification in the missing PSID years:

1. For $t = \hat{t}+1, \hat{t}+3, ..., T-1$, assume $v_{\kappa^0,t} = \frac{v_{\kappa^0,t-1}+v_{\kappa^0,t+1}}{2}$. Given this "smooth cohort effects" assumption, the moment

$$var_t^1(\log \hat{w}) - var_t^0(\log \hat{w}) = \left(v_{\alpha^0,t-1} + v_{\omega t}\right) + \left(v_{\kappa^0,t-1} + v_{\eta t}\right) - v_{\alpha^0 t} - v_{\kappa^0 t} \qquad \text{(D8)}$$

   for $t = \hat{t}+2, \hat{t}+4, ..., T$ identifies the corresponding values for $v_{\eta,t}$. Given that $\{v_{\eta,t-1} + v_{\eta t}\}$ is already identified for these years from Corollary 2.2 and Assumption 1, the corresponding values for $v_{\eta,t-1}$ are also identified.

2. For $t = \hat{t}+1, \hat{t}+3, ..., T-1$, assume $v_{\theta t} = \frac{v_{\theta,t-1}+v_{\theta,t+1}}{2}$.

When we estimate the model where variances are allowed to vary freely year by year, we make the following three additional identifying assumptions, beyond 1. and 2. above, to complete identification at endpoints:

3. Assume $v_{\kappa^0,T} = v_{\kappa^0,T-2}$. Given that $\left\{v_{\kappa^0,T} + v_{\theta,T}\right\}$ and $\{v_{\eta,T-1} + v_{\eta,T} + v_{\theta,T}\}$ are already identified from Corollary 2.2, this assumption identifies $v_{\theta,T}$ and $\{v_{\eta,T-1} + v_{\eta,T}\}$.

4. Assume $v_{\omega 1} = v_{\omega 2}$.

5. Assume $v_{\eta 1} = v_{\eta 2}$.

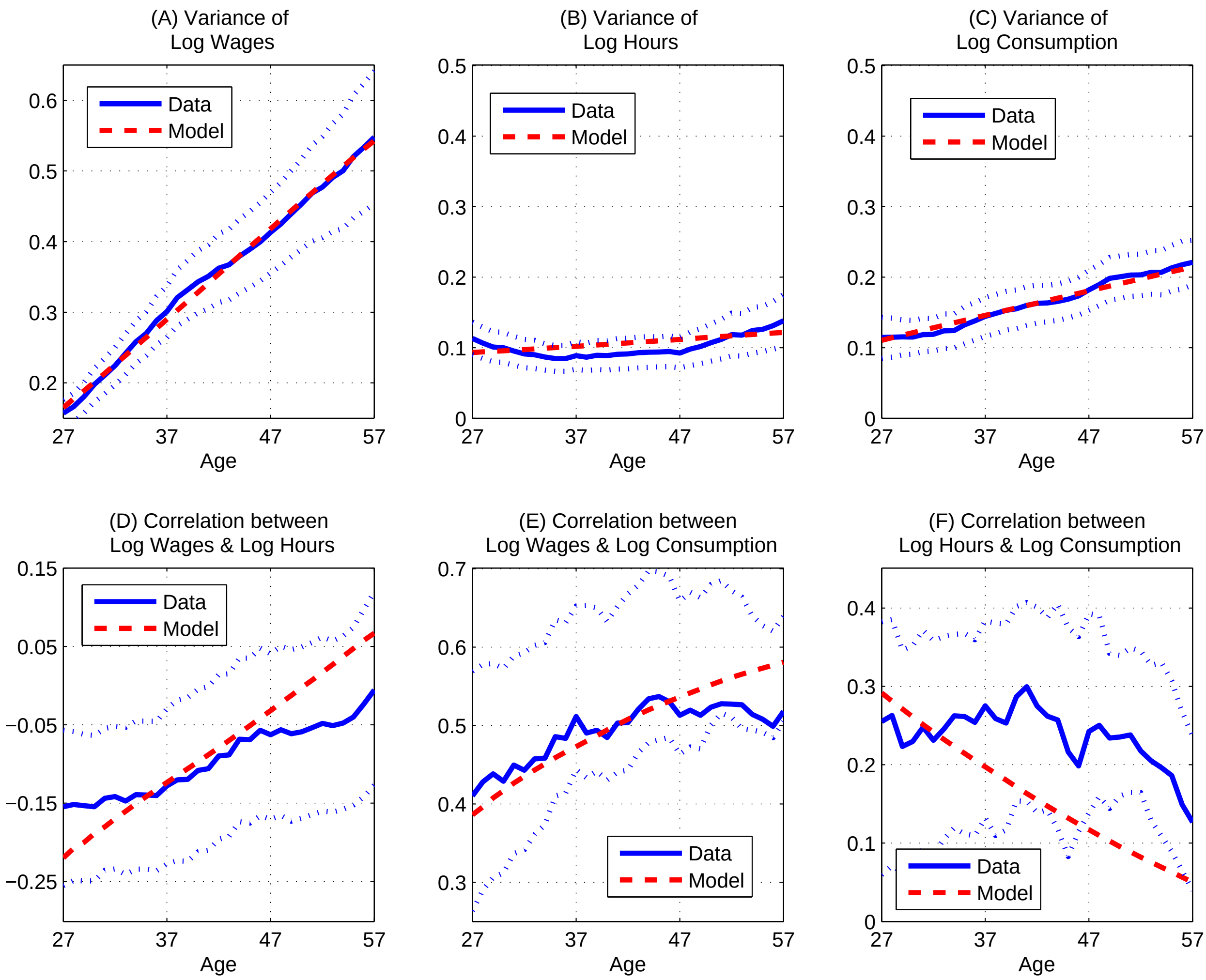

Figure 1: Data and model fit for moments in levels along the age dimension. These plots are constructed by regressing observations for all (age a, year t) cells on a set of age and cohort dummies. The plots show the estimated age coefficients. For the variances of wages and hours and for the wage-hours correlation, we use the entire 1967-2006 sample period. For the moments involving consumption, we use the 1980-2006 sample for which consumption data are available. The same regression procedure for constructing the age-profiles is applied to the data and to the model-generated moments. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

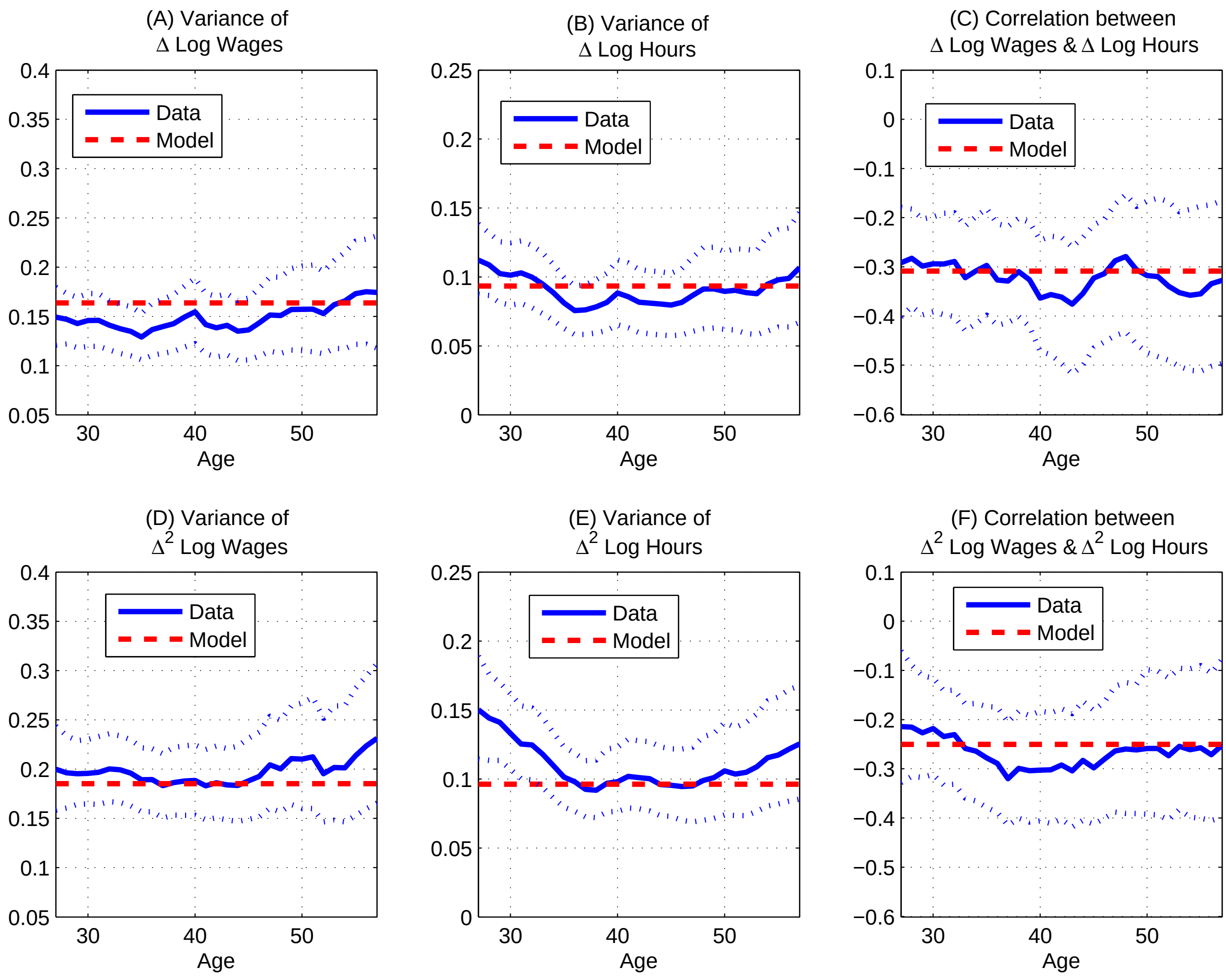

Figure 2: Data and model fit for moments in differences along the age dimension. Panels in the upper row show first differences for the years 1967-1996. Panels in the lower row show second differences for the years 1967-2006. These plots are constructed by taking the average across time for each age group a: we do not control for cohort effects in constructing these plots, because differencing already eliminates cohort effects from the theoretical moments. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

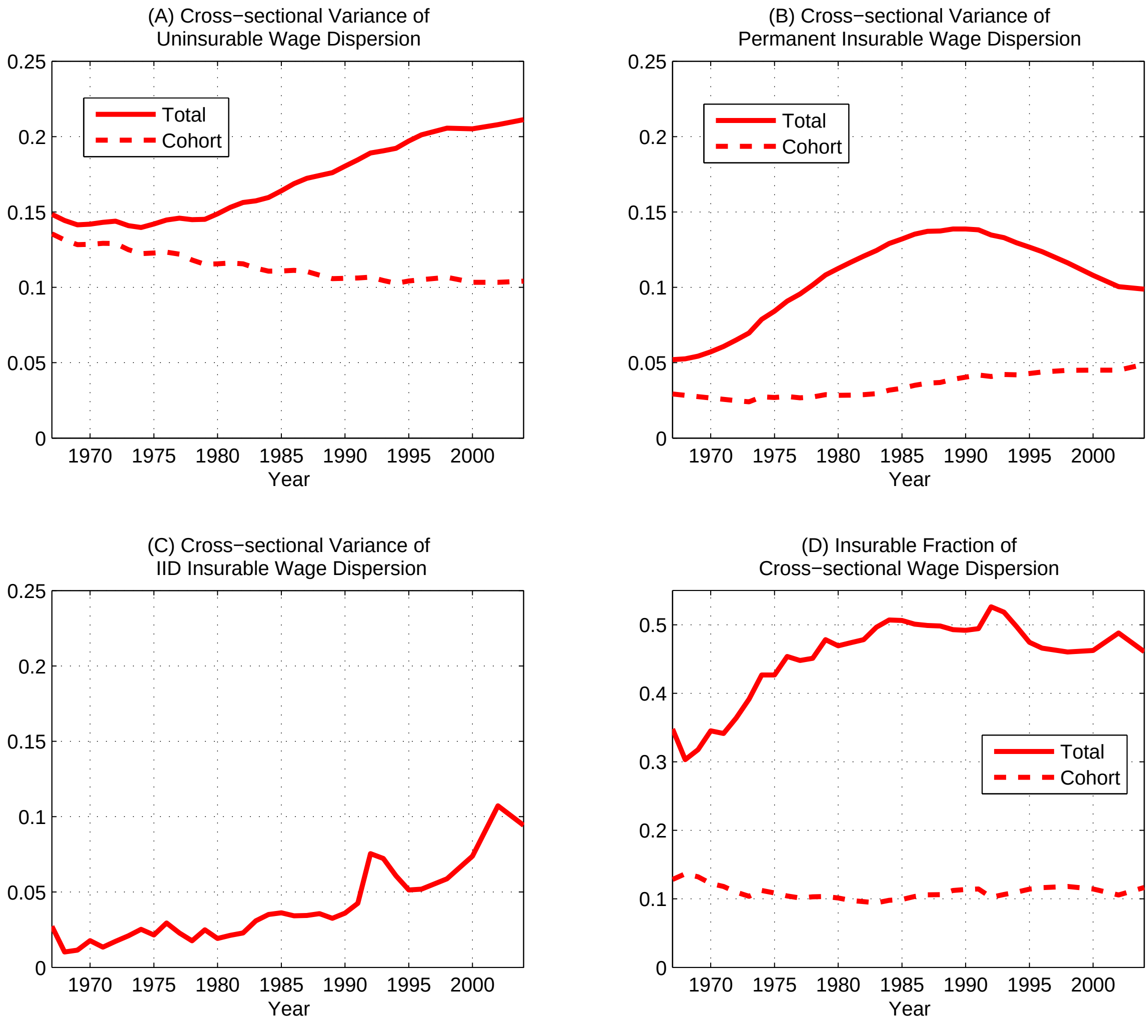

Figure 3: Panel A plots the cross-sectional variance for the uninsurable component of wages α_t (series labeled "Total") and the cross-sectional variance for the cohort-specific initial-age uninsurable component α_t^0 (series labeled "Cohort"). Panel B plots the corresponding series for the insurable component: κ_t (series labeled "Total") and κ_t^0 (series labeled "Cohort"). Panel C plots $v_{\theta t}$. In Panel D the "Total" line is the ratio of the sum of the "Total" series in Panels B and C to the total cross-sectional variance of wages (the sum of the "Total" series in Panels A, B and C). The "Cohort" line in Panel D is the ratio of the "Cohort" series in Panel B to the sum of "Cohort" lines in Panels A and B.

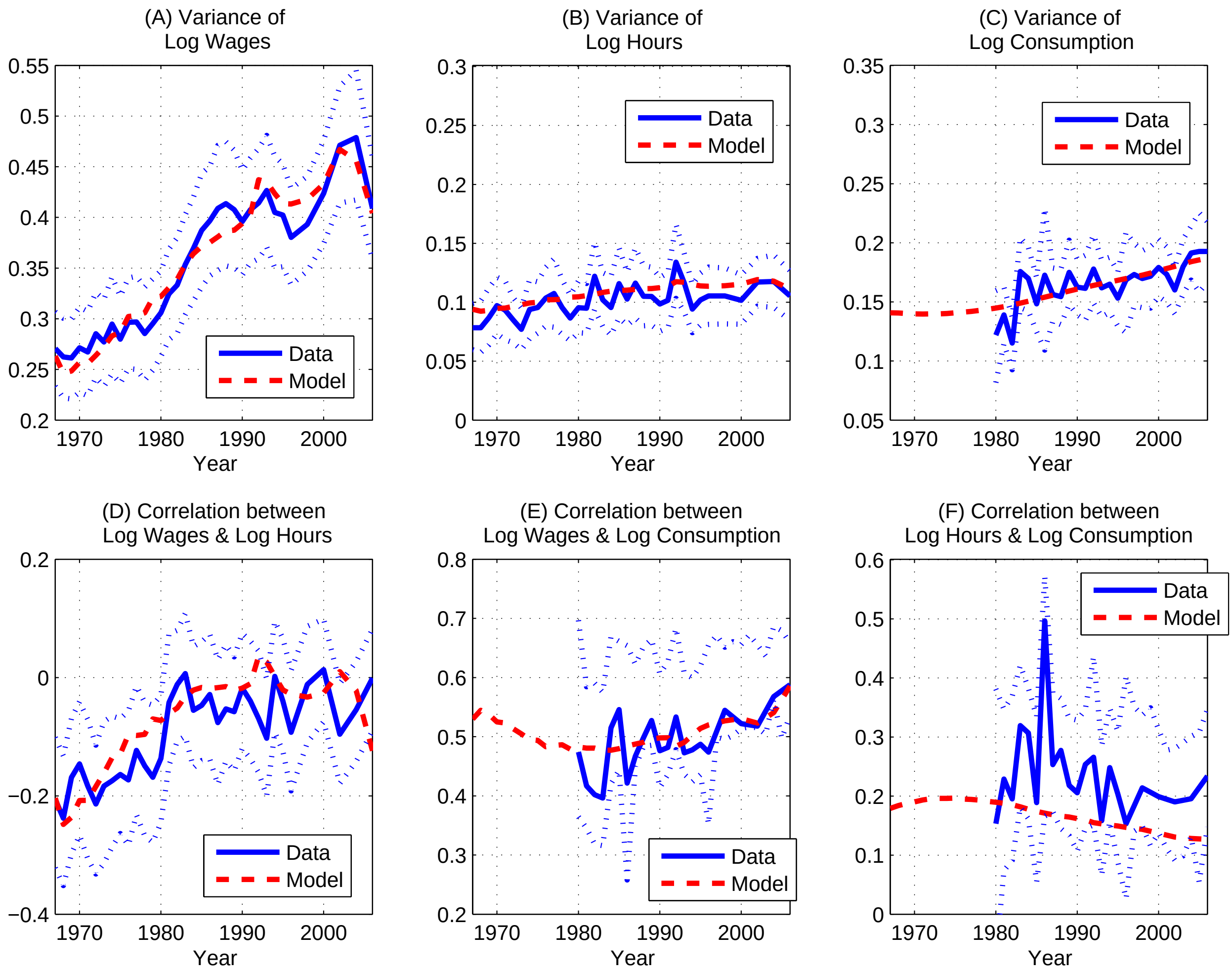

Figure 4: Data and model fit for moments in levels along the time dimension. These plots are constructed by aggregating across age groups within a given year by weighting each age group by its survival probability to account for mortality. We use the same weights in both model and data. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

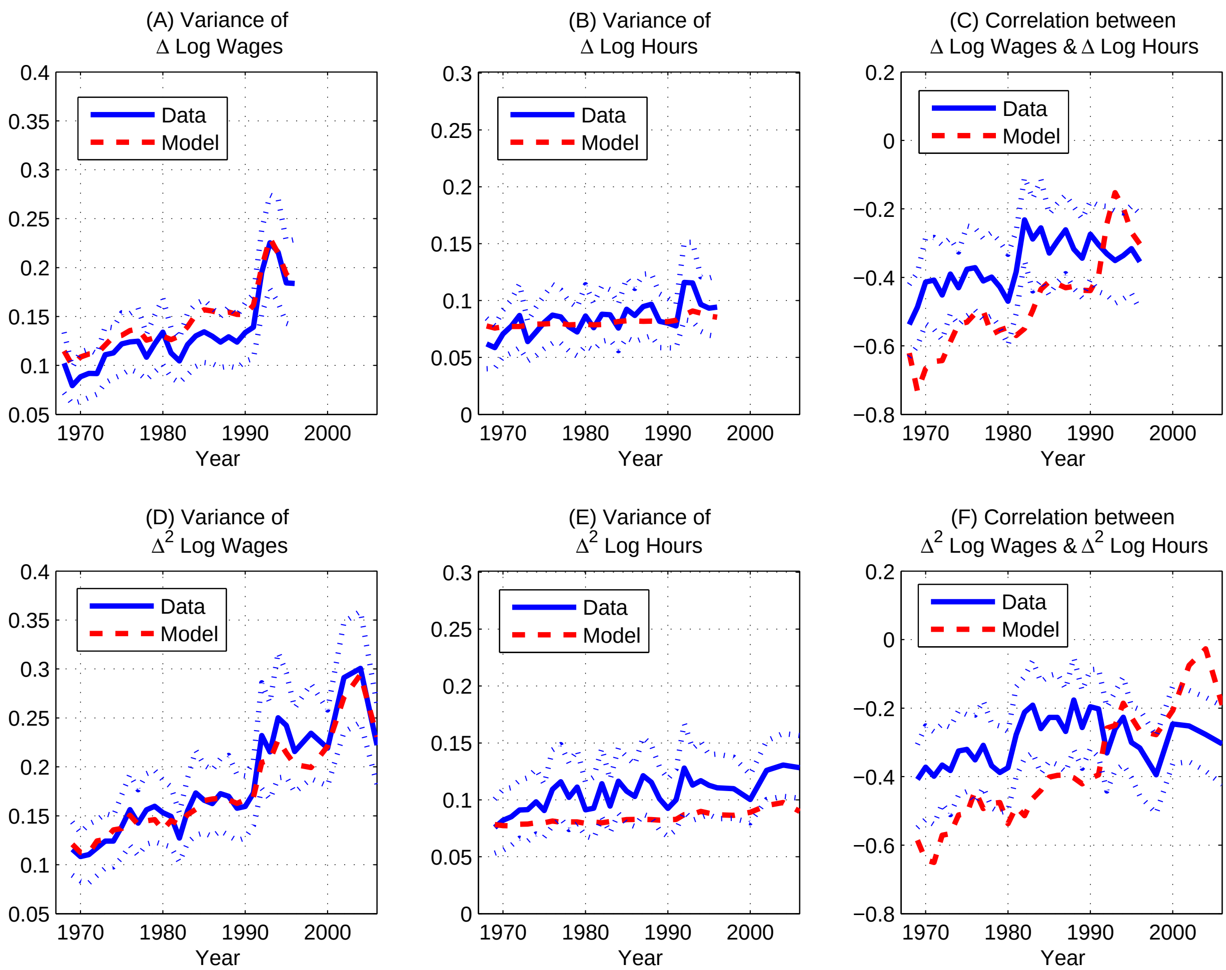

Figure 5: Data and model fit for moments in differences along the time dimension. Panels in the upper row show first differences for the years 1967-1996. Panels in the lower row show second differences for the years 1967-2006. These plots are constructed by aggregating across age groups within a given year by weighting each age group by its survival probability to account for mortality. We use the same weights in both model and data. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

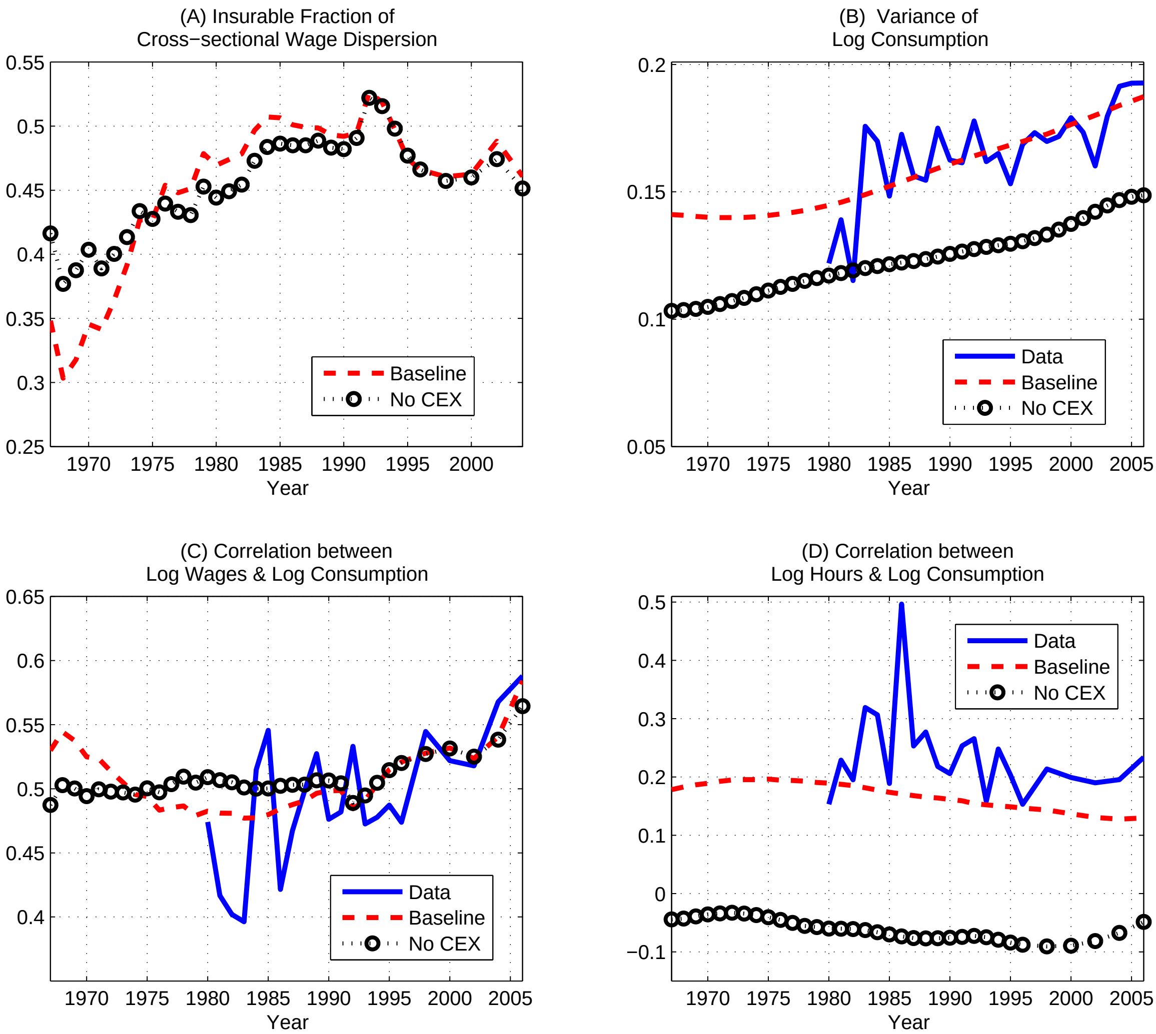

Figure 6: Data, baseline model estimated on the PSID and the CEX, and model estimated without CEX data (series labeled "No CEX"). Plots in Panel A are constructed as the line labeled "Total"in Panel A of Figure 3, and plots in Panels B-D as in Figure 4.

Technical Appendix for "Consumption and Labor Supply with Partial Insurance: An Analytical Framework"

Jonathan Heathcote
Federal Reserve Bank of Minneapolis and CEPR
heathcote@minneapolisfed.org

Kjetil Storesletten
Federal Reserve Bank of Minneapolis, University of Oslo, and CEPR
kjstore@econ.uio.no

Giovanni L. Violante
New York University, CEPR, and NBER
glv2@nyu.edu

April 2013

This Technical Appendix is organized as follows. Section A contains an extended proof of Proposition 1 (existence of no-trade competitive equilibrium and characterization of equilibrium allocations, asset prices and asset purchases). Section B develops in detail the two household models discussed in Section 3.1 of the paper which provide a foundation for equivalizing the data. Section C contains identification proofs for Proposition 3 (when consumption data are not available) and an extension for the case where data are biannual. Section D describes the construction of the overidentifying restriction test statistic.

A Extended Proof of Proposition 1

The proof is in two parts. In the first part we describe a planner's problem, and show that the allocations for consumption and hours described in Proposition 1, part (ii) are the solution to this problem. In the second part, we decentralize these allocations in a competitive equilibrium, and show that the asset prices described in Proposition 1, part (iii) and the no-inter-island-trade result described in part (i) form part of this decentralization.

Planner's Problem (allocations): We first solve for equilibrium allocations for consumption and hours worked by solving a set of static planning problems. Each island-level planner maximizes equally weighted period utility for a set of agents that share a common age a, a common preference weight φ, and a common wage component α_t. Let $x_t = (a, \varphi, \alpha_t)$ denote these island-level components of the individual state. Each island-level planner controls a set of agents with the age-specific population distributions $F^a_{\kappa,t}$ and $F_{\theta,t}$. Let $F^a_{\varepsilon,t}$ denote the implied age-specific distribution over $\varepsilon_t = \kappa_t + \theta_t$. The planner's problem on an island defined by x_t is to choose functions $c_t(x_t, \varepsilon_t)$, $h_t(x_t, \varepsilon_t)$ to solve

$$\max_{\{c_t(x_t,\cdot),h_t(x_t,\cdot)\}} \int \left[\frac{c_t(x_t, \varepsilon_t)^{1-\gamma} - 1}{1-\gamma} - \exp(\varphi) \frac{h_t(x_t, \varepsilon_t)^{1+\sigma}}{1+\sigma} \right] dF^a_{\varepsilon,t}$$

subject to the island-level resource constraint

$$\int \left[\lambda \left(\exp(\alpha_t + \varepsilon_t) h_t(x_t, \varepsilon_t) \right)^{1-\tau} - c_t(x_t, \varepsilon_t) \right] dF^a_{\varepsilon,t} = 0.$$

The first-order conditions with respect to $c_t(x_t, \varepsilon_t)$ and $h_t(x_t, \varepsilon_t)$ are, respectively,

$$\begin{aligned} c_t(x_t, \varepsilon_t)^{-\gamma} &= \chi_t(x_t), \\ \exp(\varphi) h_t(x_t, \varepsilon_t)^{\sigma} &= \chi_t(x_t) \lambda \exp(\alpha_t(1-\tau)) \exp(\varepsilon_t(1-\tau)) (1-\tau) h_t(x_t, \varepsilon_t)^{-\tau}, \end{aligned}$$

where $\chi_t(x_t)$ is the multiplier on the date t resource constraint. Note that $c_t(x_t, \varepsilon_t) = \chi_t(x_t)^{-\frac{1}{\gamma}}$, and thus does consumption does not depend on ε_t. Combining the two FOCs gives

$$h_t(x_t, \varepsilon_t) = ((1-\tau)\lambda)^{\frac{1}{\sigma+\tau}} c_t(x_t)^{-\frac{\gamma}{\sigma+\tau}} \exp\left(\frac{1-\tau}{\sigma+\tau} (\alpha_t + \varepsilon_t) - \frac{1}{\sigma+\tau} - \frac{\varphi}{\sigma+\tau} \right). \qquad \text{(A1)}$$

Substituting (A1) into the resource constraint gives

$$\begin{aligned} c_t(x_t, \varepsilon_t) &= \lambda((1-\tau)\lambda)^{\frac{1-\tau}{\sigma+\tau}} \exp(\alpha_t(1-\tau)) c_t(x_t, \varepsilon_t)^{-\frac{\gamma(1-\tau)}{\sigma+\tau}} \exp\left(-\frac{1-\tau}{\sigma+\tau} \right) \\ &\times \exp\left(-\frac{\varphi(1-\tau)}{\sigma+\tau} + \alpha_t \frac{(1-\tau)^2}{\sigma+\tau} \right) \int \exp((1-\tau)\varepsilon_t) \exp\left(\frac{(1-\tau)^2}{\sigma+\tau} \varepsilon_t \right) dF^a_{\varepsilon t}. \end{aligned}$$

Taking logs and simplifying yields

$$\begin{aligned} &\log c_t(x_t, \varepsilon_t) \\ &= \frac{1+\sigma}{\sigma+\tau+\gamma(1-\tau)} \log \lambda + \frac{1-\tau}{\sigma+\tau+\gamma(1-\tau)} \log(1-\tau) - \frac{1-\tau}{\sigma+\tau+\gamma(1-\tau)} \varphi \\ &+ \frac{(1-\tau)(1+\sigma)}{\sigma+\tau+\gamma(1-\tau)} \alpha_t + \frac{\sigma+\tau}{\sigma+\tau+\gamma(1-\tau)} \log \int \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau} \varepsilon_t \right) dF^a_{\varepsilon t}. \end{aligned}$$

By using the definition for the tax-modified Frisch elasticity $\widehat{\sigma} = (\sigma + \tau)/(1-\tau)$, the above expression simplifies to:

$$\log c_t(x_t, \varepsilon_t) = -\frac{\varphi}{\widehat{\sigma}+\gamma} + \frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}+\gamma}\alpha_t + \mathcal{C}_t^a \tag{A2}$$

which is the expression in Proposition 1, part (ii), where $\mathcal{C}_t^a$ is a constant common to all agents of age a in year t given by

$$\begin{aligned}
\mathcal{C}_t^a &= \frac{1}{\widehat{\sigma}+\gamma}\left((1+\widehat{\sigma})\log\lambda + \log(1-\tau)\right) + \mathcal{M}_t^a, \\
\mathcal{M}_t^a &= \frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\varepsilon_t\right) dF_{\varepsilon,t}^a.
\end{aligned}$$

Note that if we were to assume, for example, that $\log\varepsilon_t^a \sim N\left(-\frac{v_{\varepsilon t}^a}{2}, v_{\varepsilon t}^a\right)$, then we could solve out the integral in the expression for $\mathcal{M}_t^a$:

$$\mathcal{M}_t^a = \frac{1}{\widehat{\sigma}+\gamma}\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\left(1-\tau\left(1+\widehat{\sigma}\right)\right)\frac{v_{\varepsilon t}}{2}\right).$$

We now substitute the expression for $\log c_t(x_t, \varepsilon_t)$ in (A2) into (A1) to solve for $\log h_t(x_t, \varepsilon_t)$:

$$\log h_t(x_t, \varepsilon_t) = -\frac{1}{(1-\tau)(\widehat{\sigma}+\gamma)}\varphi + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)\alpha_t + \frac{1}{\widehat{\sigma}}\varepsilon_t + \mathcal{H}_t^a$$

which is the expression in Proposition 1, part (ii), where

$$\mathcal{H}_t^a \equiv \frac{1}{(1-\tau)(\widehat{\sigma}+\gamma)}\left((1-\gamma)\log\lambda + \log(1-\tau)\right) - \frac{\gamma}{\widehat{\sigma}(1-\tau)}\mathcal{M}_t^a.$$

Decentralization (prices): We now turn to the second part of the proof of Proposition 1, namely the decentralization of the solution to the above planner's problem. We begin by conjecturing prices in this equilibrium. We set pre-tax wages equal to individual labor productivity:

$$w_t(x_t, \varepsilon_t) = \exp(\alpha_t + \varepsilon_t).$$

At this wage, the intratemporal FOC from the agent's problem (2.1) described in the main text is identical to the intratemporal FOC for the planner described in eq.($A1$). Thus at competitive wages and the conjectured allocations (eqs. 7 and 8) agents are optimizing on the intratemporal margin. At first blush this might seem surprising, given the presence of progressive earnings taxation in the economy. Recall, however, that individual agents (in the competitive equilibrium) and island-level planners (in the problem described above) are atomistic and hence both take the tax system parameters as exogenous.

We next conjecture equilibrium prices for intertemporal insurance claims. At this point it is convenient to revert to history-dependent notation, so we will write $c_t(s^t)$ rather than $c_t(x_t, \varepsilon_t)$. We begin with the price of within-island insurance $Q_t(S; s^t)$. The intertemporal FOC from the agent's problem (2.1) defines the price at which an agent of age a with history s^t is willing, on the margin, to buy or sell a set of insurance contracts $B_t(S; s^t)$ that pay δ^{-1} units of consumption if and only if $s_{t+1} = (\omega_{t+1}, \eta_{t+1}, \theta_{t+1}) \in S \subseteq \mathbb{S}$. This price is simply the average marginal rate of substitution in those states:[2]

$$Q_t\left(S; s^t\right) = \beta\delta\delta^{-1}\int_S \frac{c_{t+1}(s^t, s_{t+1})^{-\gamma}}{c_t(s^t)^{-\gamma}} dF_{s,t+1}. \tag{A3}$$

Substituting in the expression for consumption (A2) we have

$$Q_t\left(S; s^t\right) = \beta \exp\left(-\gamma\left(\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a\right)\right)\int_S \exp\left(-\gamma(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\omega_{t+1}\right) dF_{s,t+1}, \tag{A4}$$

which is the expression in Proposition 1, part (iii), where $\mathcal{C}_t^a$ is defined above, and

$$\begin{aligned}\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a &= \frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma}\left[\log\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\varepsilon_{t+1}\right) dF_{\varepsilon,t+1}^{a+1} - \log\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\varepsilon_t\right) dF_{\varepsilon,t}^{a}\right] \\ &= \frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log\left(\frac{\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\eta_{t+1}\right) dF_{\eta,t+1}\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\theta_{t+1}\right) dF_{\theta,t+1}}{\int \exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\theta_t\right) dF_{\theta,t}}\right)\end{aligned}$$

is independent of a. Thus the prices $Q_t(S; s^t)$ are consistent with optimization on the consumer side.

Note that $Q_t(S; s^t) = Q_t(S)$: insurance prices are independent of the individual history s^t and age a. From eq. (A4) there are two pieces to this result. First, $F_{s,t+1}$, the joint distribution over $s_{t+1} = (\omega_{t+1}, \eta_{t+1}, \theta_{t+1})$ at $t+1$, is independent of s^t and thus the second term in eq. (A4) is independent of s^t. Second, insurance prices are also independent of age a, because while average consumption $\mathcal{C}_t^a$ is age-dependent, *growth* in average consumption $\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a$ is independent of age, reflecting the permanent-transitory model for individual productivity dynamics. Note also that the price of insurance against η_{t+1} and θ_{t+1} simply reflects probabilities, while the price of insurance against ω_{t+1} also reflects the conditional marginal rate of substitution, with insurance against low ω_{t+1} realizations being more expensive than equally likely high ω_{t+1} realizations. This asymmetry reflects the fact that η_{t+1}

[2] Note that the agent effectively discounts at rate $\beta\delta$, while mortality insurance ensures payment of δ^{-1} units of consumption in the event that the agent survives to the next period and $s_{t+1} \in S$.

and θ_{t+1} are perfectly insured in equilibrium, while ω_{t+1} remains uninsured. The price of a risk-free bond $Q_t(\mathbb{S})$ is

$$Q_t\left(\mathbb{S};s^t\right) = \beta \exp\left(-\gamma\left(\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a\right)\right) \int_{\mathbb{S}} \exp\left(-\gamma(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\omega_{t+1}\right) dF_{s,t+1} = Q_t\left(\mathbb{S}\right).$$

We now turn to the price function for insurance claims traded across islands. Because any contract that can be traded between islands can also be traded within an island, the inter-island price for a claim that pays δ^{-1} units of consumption iff $s_{t+1} \in Z$ must, by arbitrage, equal the corresponding within-island price, for any Z. This implies

$$Q_t^*\left(Z;s^t\right) = \Pr\left((\eta_{t+1},\theta_{t+1}) \in Z\right) \times Q_t(\mathbb{S}) = Q_t^*\left(Z\right).$$

Thus these prices are just probabilities times the price of a risk-free bond.[3]

Assuming log-normal distributions for ω_{t+1}, η_{t+1} and θ_{t+1} allows us to solve out the integral in the expression for the risk-free rate $Q_t(\mathbb{S})$. In this case,

$$\mathcal{C}_{t+1}^{a+1} - \mathcal{C}_t^a = \frac{(1-\tau)(1+\widehat{\sigma})\left(1-\tau\left(1+\widehat{\sigma}\right)\right)}{\left(\widehat{\sigma}+\gamma\right)\widehat{\sigma}}\left(\frac{v_{\eta,t+1}+v_{\theta,t+1}-v_{\theta,t}}{2}\right)$$

and thus

$$\begin{aligned} Q_t\left(\mathbb{S}\right) &= \beta \exp\left(-\gamma\frac{(1-\tau)(1+\widehat{\sigma})\left(1-\tau\left(1+\widehat{\sigma}\right)\right)}{\left(\widehat{\sigma}+\gamma\right)\widehat{\sigma}}\left(\frac{v_{\eta,t+1}+v_{\theta,t+1}-v_{\theta,t}}{2}\right)\right) \\ &\quad \times \exp\left(-\gamma(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\left(-\gamma(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}-1\right)\frac{v_{\omega,t+1}}{2}\right). \end{aligned} \tag{A5}$$

Expression (10) in the main text is a special case when $\tau = 0$.

Decentralization (asset purchases): We now derive expressions for insurance contract purchases, $B_t(s_{t+1};s^t)$ and $B_t^*(\eta_{t+1},\theta_{t+1};s^t)$ and verify that, given all conjectured prices and quantities, agents' budget constraints are satisfied.

Given that any available inter-island insurance contract can be purchased at the same price on the within-island market, $B_t^*(\eta_{t+1},\theta_{t+1};s^t) = 0$ for all $(\eta_{t+1},\theta_{t+1})$ is consistent with individual optimization (Proposition 1, part (iii)). Thus, agents optimize when purchasing all their insurance on the island on which they are located. At the same time, because $Q_t^*\left(Z;s^t\right) = Q_t^*\left(Z\right)$, no agent has an incentive to try to sell insurance to an agent located on another island. To understand this, note that the price at which one agent (say agent i_1) with

[3] If we allowed insurance contracts to be traded across islands contingent on ω_{t+1} then agents would pool ω_{t+1} risk and insurance prices would be $\Pr\left((\omega_{t+1},\eta_{t+1},\theta_{t+1}) \in S\right) \times \beta \neq Q_t\left(S\right)$.

history $s_{i_1}^t$ is willing to buy, on the margin, a set of claims that pay if and only if $(\eta_{t+1}, \theta_{t+1}) \in Z$ is the probability of that event times agent i_1's expected marginal rate of substitution, i.e. $\Pr((\eta_{t+1}, \theta_{t+1}) \in Z) \times Q_t\left(\mathbb{S}; s_{i_1}^t\right)$. The price at which a second agent on a different island (agent i_2 with history $s_{i_2}^t$) is willing to sell this insurance to agent i_1 is the same probability times agent i_2's expected marginal rate of substitution, $\Pr((\eta_{t+1}, \theta_{t+1}) \in Z) \times Q_t\left(\mathbb{S}; s_{i_2}^t\right)$. If agents i_1 and i_2 did not share the same marginal rate of substitution (i.e., if $Q_t\left(\mathbb{S}; s_{i_1}^t\right) \neq Q_t\left(\mathbb{S}; s_{i_2}^t\right)$), then there could be no equilibrium without inter-island trade, because any such equilibrium would feature unexploited gains from trade. Thus $Q_t(\mathbb{S}, s^t) = Q_t(\mathbb{S})$ is the crucial result supporting an absence of inter-island trade.

Finally, we now derive an expression for purchases of state-contingent claims, $B_t(s_{t+1}; s^t)$, and verify budget balance. Given $B_t^*(Z; s^t) = 0 \; \forall Z, \forall s^t$, realized wealth at s^t implicitly defines insurance purchases:

$$B_{t-1}(s_t; s^{t-1}) = \delta d_t(s^t).$$

We will now guess and verify the following solution for $d_t(s^t)$:

$$d_t(s^t) = \hat{d}_t\left(s^t\right) + T_t\left(s^t\right)$$

where

$$\begin{aligned} T_t\left(s^t\right) &= c_t\left(s^t\right) - \lambda\left(w_t\left(s^t\right) h_t\left(s^t\right)\right)^{1-\tau}, \\ \hat{d}_t\left(s^t\right) &= \mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty} \frac{(\beta\delta)^j c_{t+j}(s^{t+j})^{-\gamma}}{c_t(s^t)^{-\gamma}} T_{t+j}\left(s_{t+j}\right)\right]. \end{aligned}$$

The logic for this guess is that insurance payouts must deliver the appropriately discounted present value of lifetime differences between consumption and after-tax earnings.

We now need to check that the agent's budget constraint is satisfied. Given $B_t^*(Z; s^t) = 0$ $\forall Z, \forall s^t$ this amounts to checking that

$$c_t\left(s^t\right) + \int\int\int Q_t\left(\omega, \eta, \theta\right) B_t\left(\left(\omega, \eta, \theta\right); s^t\right) d\omega\, d\eta\, d\theta = \lambda\left(w_t\left(s^t\right) h_t\left(s^t\right)\right)^{1-\tau} + \hat{d}_t\left(s^t\right) + T_t\left(s^t\right).$$

Given the conjecture for $T\left(s^t\right)$ this simplifies to

$$\int\int\int Q_t\left(\omega, \eta, \theta\right) B_t\left(\left(\omega, \eta, \theta\right); s^t\right) d\omega\, d\eta\, d\theta = \hat{d}_t\left(s^t\right). \tag{A6}$$

To verify that this equation is in fact satisfied, we will write the functions $Q_t\left(\omega, \eta, \theta\right)$, $B_t\left(\left(\omega, \eta, \theta\right); s^t\right)$ and $\hat{d}_t\left(s^t\right)$ all in terms of the decision rule for consumption $c_t\left(s^t\right)$. The ratio of after-tax earnings to consumption is

$$\frac{\lambda\left(w_t\left(s^t\right)h_t\left(s^t\right)\right)^{1-\tau}}{c_t\left(s^t\right)}=\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\varepsilon_t-\frac{\gamma+\widehat{\sigma}}{\widehat{\sigma}}\mathcal{M}_t^a\right),$$

so

$$\begin{aligned}T_t\left(s^t\right) &= \left(1-\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_t+\theta_t)-\frac{\gamma+\widehat{\sigma}}{\widehat{\sigma}}\mathcal{M}_t^a\right)\right)c_t\left(s^t\right)\\ &= c_t(s^t)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_t+\theta_t)\right)}{\int\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_t+\theta_t)\right)\,dF_{\kappa,t}^a dF_{\theta,t}}\right),\end{aligned}$$

where the second line uses

$$\mathcal{M}_t^a=\frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log\int\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\left(\kappa_t+\theta_t\right)\right)\,dF_{\kappa,t}^a dF_{\theta.t}.$$

Substituting the definition for $T_{t+j}\left(s_{t+j}\right)$ into the one for $\hat{d}_t\left(s^t\right)$, and multiplying and dividing by $c_t(s^t)$, gives

$$\begin{aligned}\hat{d}_t\left(s^t\right) &= c_t(s^t)\mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+j}(s^{t+j})^{-\gamma}}{c_t(s^t)^{-\gamma}}\frac{c_{t+j}(s^{t+j})}{c_t(s^t)}\times\right.\\ &\quad\left.\times\left(1-\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\left(\kappa_t+\sum_{i=1}^{j}\eta_{t+i}+\theta_{t+j}\right)-\frac{\gamma+\widehat{\sigma}}{\widehat{\sigma}}\mathcal{M}_{t+j}^{a+j}\right)\right)\right]\\ &= c_t(s^t)\mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+j}(s^{t+j})^{1-\gamma}}{c_t(s^t)^{1-\gamma}}\times\right.\\ &\quad\left.\times\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\left(\kappa_t+\sum_{i=1}^{j}\eta_{t+i}+\theta_{t+j}\right)\right)}{\int....\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\left(\kappa_t+\sum_{i=1}^{j}\eta_{t+i}+\theta_{t+j}\right)\right)\,dF_{\kappa,t}^a\,dF_{\eta,t+1}\,...\,dF_{\eta,t+j}\,dF_{\theta,t+j}}\right)\right]\\ &= c_t(s^t)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF_{\kappa,t}^a}\right)\mathbb{E}_{s^t}\left[\frac{\sum_{j=1}^{\infty}(\beta\delta)^j\,c_{t+j}(s^{t+j})^{1-\gamma}}{c_t(s^t)^{1-\gamma}}\right],\end{aligned}\tag{A7}$$

where the second equation uses

$$\mathcal{M}_{t+j}^{a+j}=\frac{\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log\int....\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\left(\kappa_t+\sum_{i=1}^{j}\eta_{t+i}+\theta_{t+j}\right)\right)\,dF_{\kappa,t}^a\,dF_{\eta,t+1}\,...\,dF_{\eta,t+j}\,dF_{\theta,t+j}.$$

Thus

$$\begin{aligned}
& B_{t-1}((\omega_t,\eta_t,\theta_t)\,;s^{t-1}) && \text{(A8)}\\
=\ & \delta\left(\hat{d}_t\left(s^t\right)+T_t\left(s^t\right)\right)\\
=\ & \delta c_t(s^t)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+j}(s^{t+j})^{1-\gamma}}{c_t(s^t)^{1-\gamma}}\right]\\
& +\delta c_t(s^t)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_t+\theta_t)\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_t+\theta_t)\right)\,dF^a_{\kappa,t}dF_{\theta,t}}\right).
\end{aligned}$$

Substituting eq. (A7) and eq. (A3) into eq. (A6) gives

$$\begin{aligned}
& \beta c_t\left(s^t\right)^{\gamma}\mathbb{E}_{s^t}\left[c_{t+1}(s^t,s_{t+1})^{-\gamma}B_t\left(s_{t+1};s^t\right)\right] && \text{(A9)}\\
=\ & c_t\left(s^t\right)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\mathbb{E}_{s^t}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+j}(s^{t+j})^{1-\gamma}}{c_t(s^t)^{1-\gamma}}\right].
\end{aligned}$$

Let $LHS(s^t)$ denote the left-hand side of eq. (A9), substitute in eq. (A8) and simplify

$$\begin{aligned}
LHS(s^t)\ =\ & \beta c_t\left(s^t\right)^{\gamma}\mathbb{E}_{s^t}\left[c_{t+1}(s^t,s_{t+1})^{-\gamma}\,\delta c_{t+1}(s^{t+1})\times\right.\\
& \times\left\{\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_{t+1}\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_{t+1}\right)\,dF^{a+1}_{\kappa,t+1}}\right)\mathbb{E}_{s^{t+1}}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+1+j}(s^{t+1+j})^{1-\gamma}}{c_{t+1}(s^{t+1})^{1-\gamma}}\right]+\right.\\
& \left.\left.+\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_{t+1}+\theta_{t+1})\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}(\kappa_{t+1}+\theta_{t+1})\right)\,dF^{a+1}_{\kappa,t+1}dF_{\theta,t+1}}\right)\right\}\right]
\end{aligned}$$

$$\begin{aligned}
=\ & \beta\delta c_t\left(s^t\right)^{\gamma}\mathbb{E}_{s^t}\left[c_{t+1}(s^t,s_{t+1})^{1-\gamma}\times\right.\\
& \times\left\{\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\mathbb{E}_{s^{t+1}}\left[\sum_{j=1}^{\infty}\frac{(\beta\delta)^j\,c_{t+1+j}(s^{t+1+j})^{1-\gamma}}{c_{t+1}(s^{t+1})^{1-\gamma}}\right]+\right.\\
& \left.\left.+\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\right\}\right]\\
=\ & \beta\delta c_t\left(s^t\right)^{\gamma}\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\mathbb{E}_{s^t}\left[\mathbb{E}_{s^{t+1}}\left[\sum_{k=1}^{\infty}(\beta\delta)^{k-1}\,c_{t+k}(s^{t+k})^{1-\gamma}\right]\right]\\
=\ & c_t\left(s^t\right)\left(1-\frac{\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)}{\int\exp\left((1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}}\kappa_t\right)\,dF^a_{\kappa,t}}\right)\mathbb{E}_{s^t}\left[\sum_{k=1}^{\infty}(\beta\delta)^k\,\frac{c_{t+k}(s^{t+k})^{1-\gamma}}{c_t\left(s^t\right)^{1-\gamma}}\right],
\end{aligned}$$

which is the same as the right-hand side of eq. (A9). We conclude that the budget constraint is satisfied when state-contingent bond purchases are given by eq. (A8).

B Household Models

We begin with the household model of Section 3 where household composition is insurable (an abbreviated version is contained in Appendix A.2). Next, we present the alternative model, also briefly discussed in Section 3, where demographics are uninsurable.

B.1 The household model of Section 3

Suppose that utility for individual i in a household of g adult workers ("g" for "grownups") and k children ("k" for kids) is given by

$$u(c, h_i, g, k) = \frac{1}{1-\gamma}\left(\frac{c}{e(g,k)}\right)^{1-\gamma} - \frac{\exp(\varphi)}{1+\sigma} h_i^{1+\sigma},$$

where c is household consumption and h_i is i's hours worked. The equivalence scale is given by e and satisfies $e_g \in (0,1]$, $e_{gg} < 0$, $e_k \in (0,1]$, and $e(1,0) = 1$ for all $g \geq 1$ and $k \geq 0$. Assume that the household utility function attaches equal weights to all adults (and no weight to the children), so total utility is given by

$$U = \sum_{i=1}^{g} u(c, h_i, g) = \frac{g}{1-\gamma}\left(\frac{c}{e(g,k)}\right)^{1-\gamma} - \sum_{i=1}^{g}\left(\frac{\exp(\varphi)}{1+\sigma} h_i^{1+\sigma}\right) \tag{B1}$$

As in Section A, let $x_t = (a, \varphi, \alpha_t)$ denote the island-level components of the individual state. Each island-level planner can insure realizations of ε_t, g, and k. The planner's problem is to choose functions $c_t(x_t, g, k)$, $h_{it}(x_t, \varepsilon_t, g, k)$ for $i = 1, ..., g$ to solve

$$\max_{\{c(x_t,\cdot), h_{it}(x_t,\cdot)\}} \int \left[\frac{g}{1-\gamma}\left(\frac{c_t(x_t,g,k)}{e(g,k)}\right)^{1-\gamma} - \sum_{i=1}^{g}\int \frac{\exp((\gamma+\sigma)\varphi)}{1+\sigma} h_{it}(x_t,\varepsilon_t,g,k)^{1+\sigma}\, dF_{\varepsilon t}^{a}\right] dF_t(g,k)$$

subject to the after-tax resource constraint

$$\int \left[\sum_{i=1}^{g}\int \lambda\left[\exp(\alpha_t + \varepsilon_{it})\, h_{it}(x_t,\varepsilon_t,g,k)\right]^{1-\tau} dF_{\varepsilon t}^{a} - c_t(x_t,g,k)\right] dF_t(g,k) = 0,$$

where the objective function and the constraint recognize that there is a non-degenerate within-island distribution $F_t(g,k)$ of household workers and children. Moreover, in light of the result of Section A, we have imposed that consumption is independent of ε_t.

The first-order conditions with respect to $c_t(x_t, g, k)$ and $h_{it}(x_t, \varepsilon_t, g, k)$ are, respectively,

$$g e(g,k)^{\gamma-1} c_t(x_t,g,k)^{-\gamma} = \chi_t \tag{B2}$$

$$\exp(\varphi)\, h_{it}(x_t,\varepsilon_t,g,k)^{\sigma+\tau} = \chi_t \lambda (1-\tau)\exp((1-\tau)\alpha_t)\exp((1-\tau)\varepsilon_t), \tag{B3}$$

where χ_t is the multiplier on the date t island-level resource constraint.

Let $c_t(x_t,1,0)$ denote household consumption for a one-person household. Then equation (B2) implies

$$c_t\left(x_t,g,k\right)=c_t(x_t,1,0)\left(\frac{g}{e\left(g,k\right)^{1-\gamma}}\right)^{\frac{1}{\gamma}}.$$

Combining the two first-order conditions (B2)-(B3) gives

$$h_{it}(x_t,\varepsilon_t,g,k)=\left(\frac{gc_t(x_t,g,k)^{-\gamma}}{e\left(g,k\right)^{1-\gamma}}\right)^{\frac{1}{\sigma+\tau}}\left(\lambda\left(1-\tau\right)\right)^{\frac{1}{\sigma+\tau}}\exp\left(-\frac{\varphi}{\sigma+\tau}+\left(\frac{1-\tau}{\sigma+\tau}\right)\alpha_t+\left(\frac{1-\tau}{\sigma+\tau}\right)\varepsilon_t\right).$$

Substitute in the consumption expression for the one-person households:

$$h_{it}=c_t(x_t,1,0)^{-\frac{\gamma}{\sigma+\tau}}\left(\lambda\left(1-\tau\right)\right)^{\frac{1}{\sigma+\tau}}\exp\left(-\frac{\varphi}{\sigma+\tau}+\left(\frac{1-\tau}{\sigma+\tau}\right)\alpha_t+\left(\frac{1-\tau}{\sigma+\tau}\right)\varepsilon_t\right),$$

so individual hours are insensitive to household size.

Finally we can solve for $c_t(x_t,1,0)$ from the island resource constraint:

$$\begin{aligned}
0 &= \int\left\{\sum_{i=1}^{g}\int\lambda\left[\exp\left(\alpha_t+\varepsilon_t\right)h_{it}\left(x_t,\varepsilon_t,g,k\right)\right]^{1-\tau}dF^a_{\varepsilon t}\ -c_t(x_t,g,k)\right\}dF_t\left(g,k\right)\\
&= \int\left\{\sum_{i=1}^{g}\int\int\lambda\Bigg[\exp\left(\alpha_t+\kappa_{it}+\theta_{it}\right)c_t(x_t,1,0)^{-\frac{\gamma}{\sigma+\tau}}\left(\lambda\left(1-\tau\right)\right)^{\frac{1}{\sigma+\tau}}\right.\\
&\quad\times\exp\left(-\frac{\varphi}{\sigma+\tau}+\frac{1-\tau}{\sigma+\tau}\alpha_t+\left(\frac{1-\tau}{\sigma+\tau}\right)\varepsilon_t\right)\Bigg]^{1-\tau}dF^a_{\varepsilon t}\\
&\quad\left.-c_t(x_t,1,0)\left(\frac{g}{e\left(g,k\right)^{1-\gamma}}\right)^{\frac{1}{\gamma}}\right\}dF_t\left(g,k\right).
\end{aligned}$$

Collecting terms:

$$\begin{aligned}
&c_t(x_t,1,0)^{1+\left(\frac{1-\tau}{\sigma+\tau}\right)\gamma}\\
=\ &\bar{g}\exp\left(\frac{\left(1-\tau\right)\left(1+\sigma\right)}{\sigma+\tau}\alpha_t-\frac{1-\tau}{\sigma+\tau}\varphi\right)\times\int\exp\left(\frac{\left(1-\tau\right)\left(1+\sigma\right)}{\sigma+\tau}\varepsilon_t\right)dF^a_{\varepsilon t}\\
&\times\left(1-\tau\right)^{\frac{1-\tau}{\sigma+\tau}}\left(\lambda\right)^{\frac{1+\sigma}{\sigma+\tau}}\left(\int\left(\frac{g}{e\left(g,k\right)^{1-\gamma}}\right)^{\frac{1}{\gamma}}dF_t\left(g,k\right)\right)^{-1},
\end{aligned}$$

where $\bar{g}=\int g\,dF_t\left(g,k\right)$. Substitute out for $\widehat{\sigma}=\left(\sigma+\tau\right)/\left(1-\tau\right)$ and simplify the expression as:

$$\log c_t(x_t,1,0)=\left(1-\tau\right)\frac{1+\widehat{\sigma}}{\gamma+\widehat{\sigma}}\alpha_t-\frac{\varphi}{\gamma+\widehat{\sigma}}+\mathcal{C}^a_t,$$

where the constant $\mathcal{C}_t^a$ is defined as

$$\begin{aligned}\exp \mathcal{C}_t^a &= (\bar{g})^{\frac{\widehat{\sigma}}{\gamma+\widehat{\sigma}}} \left(\int \left(\frac{g}{e(g,k)^{1-\gamma}} \right)^{\frac{1}{\gamma}} dF_t(g,k) \right)^{-\frac{\widehat{\sigma}}{\gamma+\widehat{\sigma}}} (1-\tau)^{\frac{1}{\gamma+\widehat{\sigma}}} (\lambda)^{\frac{1+\widehat{\sigma}}{\gamma+\widehat{\sigma}}} \\ &\quad \times \left[\int \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau} \varepsilon_t \right) dF_{\varepsilon t}^a \right]^{\frac{\widehat{\sigma}}{\gamma+\widehat{\sigma}}}.\end{aligned}$$

The implied allocations for household consumption and individual labor supply are then

$$\begin{aligned}\log c_t^a(x_t, g, k) &= D(g,k) - (1-\tau)\widehat{\varphi} + (1-\tau)\left(\frac{1+\widehat{\sigma}}{\hat{\sigma}+\gamma} \right) \alpha_t + \mathcal{C}_t^a \\ \log h_{it}(x_t, \varepsilon_t, g, k) &= -\widehat{\varphi} + \frac{(1-\gamma)}{(\widehat{\sigma}+\gamma)}\alpha_t + \frac{1}{\widehat{\sigma}}\varepsilon_t + \mathcal{H}_t^a,\end{aligned}$$

where $\widehat{\varphi} = \varphi / ((1-\tau)(\widehat{\sigma}+\gamma))$ is the rescaled preference weight,

$$D(g,k) \equiv \log g + \frac{(1-\gamma)}{\gamma} \log \left(\frac{g}{e(g,k)} \right)$$

and

$$\mathcal{H}_t^a = -\frac{\gamma}{\sigma+\tau}\mathcal{C}_t^a + \frac{1}{\sigma+\tau}\log(\lambda(1-\tau)).$$

Note that the individual hours allocation is independent of (g,k), and the household consumption allocation is independent of ε_t.

B.2 Alternative household model with uninsurable demographics

Utility of a household with g adults and k children is still given by equation (B1). The island-level components of the individual state are now $x_t = (a, \varphi, \alpha_t, g, k)$. The planner can insure only against realizations of ε_t. The planner chooses functions $c_t(x_t)$ and $h_{it}(x_t, \varepsilon_t)$ for $i = 1, ..., g$ to solve

$$\max_{c_t(x_t), \{h_{it}(x_{t,}, \cdot)\}} \left\{ \frac{g}{1-\gamma} \left(\frac{c_t(x_t)}{e(g,k)} \right)^{1-\gamma} - \sum_{i=1}^{g} \int \frac{\exp(\varphi)}{1+\sigma} h_{it}(x_t, \varepsilon_t)^{1+\sigma} dF_{\varepsilon t}^a \right\}$$

subject to the island-level after-tax resource constraint

$$c_t(x_t) - \sum_{i=1}^{g} \int \lambda \left[\exp(\alpha_t + \varepsilon_t) h_{it}(x_t, \varepsilon_t) \right]^{1-\tau} dF_{\varepsilon t}^a = 0,$$

where, in light of the result of Section A, we have imposed that consumption is independent of ε_t.

The first-order conditions with respect to $c_t(x_t)$ and $h_{it}(x_t, \varepsilon_t)$ are, respectively,

$$\begin{aligned} \frac{g}{e(g,k)^{1-\gamma}} c_t(x_t)^{-\gamma} &= \chi_t \\ \exp(\varphi) h_{it}(x_t, \varepsilon_t)^{\sigma+\tau} &= \chi_t \lambda (1-\tau) \exp((1-\tau)\alpha_t) \exp((1-\tau)\varepsilon_t). \end{aligned}$$

Combining the two conditions and simplifying terms yields the expression for individual hours

$$\begin{aligned} h_{it}(x_t, \varepsilon_t) &= \left(\frac{g\, c_t(x_t)^{-\gamma}}{e(g,k)^{1-\gamma}}\right)^{\frac{1}{\sigma+\tau}} (\lambda(1-\tau))^{\frac{1}{\sigma+\tau}} \exp\left(-\frac{\varphi}{\sigma+\tau} + \frac{1-\tau}{\sigma+\tau}\alpha_t + \frac{1-\tau}{\sigma+\tau}\varepsilon_t\right) \qquad \text{(B4)} \\ &= g^{\frac{1}{\widehat{\sigma}(1-\tau)}} e(g,k)^{\frac{\gamma-1}{\widehat{\sigma}(1-\tau)}} (\lambda(1-\tau))^{\frac{1}{\widehat{\sigma}(1-\tau)}} \exp\left(\frac{1}{\widehat{\sigma}}\left(\alpha_t + \varepsilon_t - \frac{\varphi}{1-\tau}\right)\right) c_t(x_t)^{\frac{-\gamma}{\widehat{\sigma}(1-\tau)}} \end{aligned}$$

and the expression for household consumption

$$\begin{aligned} c_t(x_t) &= \sum_{i\in 1}^{g} \int \lambda \left[\exp(\alpha_t + \varepsilon_t) h_{it}(x_t, \varepsilon_t)\right]^{1-\tau} dF^a_{\varepsilon t} \\ &= g \exp((1-\tau)\alpha_t) \int \lambda \exp((1-\tau)\varepsilon_t) h_{it}(x_t, \varepsilon_t)^{1-\tau} dF^a_{\varepsilon t} \\ &= (g)^{1+\frac{1-\tau}{\sigma+\tau}} e(g,k)^{(\gamma-1)\frac{1-\tau}{\sigma+\tau}} c_t(x_t)^{-\frac{\gamma(1-\tau)}{\sigma+\tau}} \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\alpha_t - \frac{1-\tau}{\sigma+\tau}\varphi\right) \\ &\quad \times (1-\tau)^{\frac{1-\tau}{\sigma+\tau}} \lambda^{\frac{1+\sigma}{\sigma+\tau}} \int \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\varepsilon_t\right) dF^a_{\varepsilon t} \end{aligned}$$

Taking logs:

$$\begin{aligned} &\left(1 + \gamma\frac{1-\tau}{\sigma+\tau}\right) \log c_t(x_t) \\ = &\left(1 + \frac{1-\tau}{\sigma+\tau}\right) \log g + (\gamma-1)\frac{1-\tau}{\sigma+\tau} \log e(g,k) + \frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\alpha_t - \frac{1-\tau}{\sigma+\tau}\varphi \\ &+ \frac{1-\tau}{\sigma+\tau}\log(1-\tau) + \log\frac{1+\sigma}{\sigma+\tau}\lambda + \log\int \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\varepsilon_t\right) dF^a_{\varepsilon t} \end{aligned}$$

and rearranging

$$\begin{aligned} \log c_t(x_t) &= \frac{\frac{\sigma+\tau}{1-\tau}}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log g + \frac{1}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log g - \frac{1-\gamma}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log e(g,k) + \frac{1+\sigma}{\frac{\sigma+\tau}{1-\tau}+\gamma}\alpha_t \\ &\quad - \frac{\varphi}{\frac{\sigma+\tau}{1-\tau}+\gamma} + \frac{1}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log(1-\tau) + \frac{\frac{1+\sigma}{1-\tau}}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log\lambda \\ &\quad + \frac{\frac{\sigma+\tau}{1-\tau}}{\frac{\sigma+\tau}{1-\tau}+\gamma}\log\int \exp\left(\frac{(1-\tau)(1+\sigma)}{\sigma+\tau}\varepsilon_t\right) dF^a_{\varepsilon t}. \end{aligned}$$

Using the $\widehat{\sigma} = (\sigma+\tau)/(1-\tau)$ notation:

$$\log c_t(x_t) = \frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log g - \frac{1-\gamma}{\widehat{\sigma}+\gamma}\log e(g,k) + \frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}+\gamma}\alpha_t - \frac{\varphi}{\widehat{\sigma}+\gamma} + \mathcal{C}^a_t, \qquad \text{(B5)}$$

where $\mathcal{C}_t^a$ is defined as

$$\mathcal{C}_t^a \equiv \frac{1}{\widehat{\sigma}+\gamma}\left[\log(1-\tau)+(1+\widehat{\sigma})\log\lambda+\widehat{\sigma}\log\int\exp\left(\frac{(1-\tau)(1+\widehat{\sigma})}{\widehat{\sigma}}\varepsilon_t\right)\,dF_{\varepsilon t}^a\right].$$

Now substitute the expression for $c_t(x_t)$ of eq. (B5) into eq. (B4) to derive an expression for $h_{it}(x_t,\varepsilon_t)$,

$$\begin{aligned}\log h_{it}(x_t,\varepsilon_t) &= \frac{1}{1-\tau}\frac{1}{\widehat{\sigma}}\log g+\frac{\gamma-1}{1-\tau}\frac{1}{\widehat{\sigma}}\log e(g,k)-\frac{1}{\widehat{\sigma}}\frac{\varphi}{1-\tau}+\frac{1}{\widehat{\sigma}}\alpha_t\\ &\quad+\frac{1}{\widehat{\sigma}}\varepsilon_t+\frac{1}{\widehat{\sigma}}\frac{1}{1-\tau}\log(\lambda(1-\tau))-\frac{1}{\widehat{\sigma}}\frac{\gamma}{1-\tau}\log c_t(x_t)\\ &= \frac{1}{1-\tau}\frac{1-\gamma}{\widehat{\sigma}+\gamma}(\log g-\log e(g,k))+\frac{1}{\widehat{\sigma}}\varepsilon_t+\frac{1-\gamma}{\widehat{\sigma}+\gamma}\alpha_t-\frac{1}{\widehat{\sigma}+\gamma}\frac{\varphi}{1-\tau}+\mathcal{H}_t^a,\end{aligned}$$

where

$$\mathcal{H}_t^a \equiv -\frac{1}{\widehat{\sigma}}\frac{1}{1-\tau}\log(\lambda(1-\tau))-\frac{1}{\widehat{\sigma}}\frac{\gamma}{1-\tau}\mathcal{C}_t^a.$$

We conclude that household consumption and individual hours are given by

$$\begin{aligned}\log c_t(x_t) &= D^c(g,k)-(1-\tau)\widehat{\varphi}+(1-\tau)\frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\alpha_t+\mathcal{C}_t^a\\ \log h_{it}(x_t,\varepsilon_t) &= D^h(g,k)-\widehat{\varphi}+\frac{1-\gamma}{\widehat{\sigma}+\gamma}\alpha_t+\frac{\varepsilon_t}{\widehat{\sigma}}+\mathcal{H}_t^a,\ i=1,...,g\end{aligned}$$

where $\widehat{\varphi}=\varphi/((1-\tau)(\widehat{\sigma}+\gamma))$ is the rescaled preference weight, and the equivalization dummies $D^c(g,k)$ and $D^h(g,k)$ are given by

$$\begin{aligned}D^c(g,k) &= \frac{1+\widehat{\sigma}}{\widehat{\sigma}+\gamma}\log g-\frac{1-\gamma}{\widehat{\sigma}+\gamma}\log e(g,k)\\ D^h(g,k) &= \frac{1}{1-\tau}\frac{1-\gamma}{\widehat{\sigma}+\gamma}[\log g-\log e(g,k)].\end{aligned}$$

To sum up, in this case both household consumption and individual hours depend in general on the vector of household type (g,k). In the special case $\gamma=1$ hours are again independent of (g,k) and consumption is proportional to the number of adults g.

C Identification

This appendix contains proofs of Propositions 3 and a new Corollary 3.1 that extends Proposition 3 to allow for biannual panel data. Finally, this appendix also contains the additional identification assumptions made in the estimation of the model (see Section 4.3).

C.1 Proof of Proposition 3

PROPOSITION 3 [IDENTIFICATION WITH NO CONSUMPTION DATA] *With an unbalanced panel on wages and hours from $t = 1, ..., T$, and an external estimate of measurement error in earnings $v_{\mu y}$, the same parameters as in Proposition 2 are identified.*

PROOF The proof is organized in three sequential steps.

1. Given foreknowledge of $v_{\mu y}$, we identify $\widehat{\sigma}$, $v_{\mu h}$, and the sequence $\{v_{\theta,t}\}_{t=1}^{T-1}$ off moments involving (co-)variance of changes minus changes in (co-)variances:

   (a) The Frisch elasticity $1/\widehat{\sigma}$ is identified by $1/\widehat{\sigma}$ equal to

   $$\frac{cov_t^a(\Delta \log \hat{w}, \Delta \log \hat{h}) + var_t^a(\Delta \log \hat{h}) - \Delta cov_t^a(\log \hat{w}, \log \hat{h}) - \Delta var_t^a(\log \hat{h})}{cov_t^a\left(\Delta \log \hat{w}, \Delta \log \hat{h}\right) + var_t^a\left(\Delta \log \hat{w}\right) - \Delta cov_t^a\left(\log \hat{w}, \log \hat{h}\right) - \Delta var_t^a\left(\log \hat{w}\right) - 2v_{\mu y}}.$$

   This expression can equivalently be formulated as

   $$\frac{1}{\widehat{\sigma}} = \frac{cov_t^a(\Delta \log \hat{y}, \Delta \log \hat{h}) - \Delta cov_t^a(\log \hat{y}, \log \hat{h})}{cov_t^a(\Delta \log \hat{y}, \Delta \log \hat{w}) - \Delta cov_t^a(\log \hat{y}, \log \hat{w}) - 2v_{\mu y}}.$$

   (b) The sequence $\{v_{\theta,t}\}_{t=1}^{T-1}$ is then identified by panel data available from $t = 2, ..., T$:

   $$\begin{aligned} & cov_t^a(\Delta \log \hat{w}, \Delta \log \hat{h}) + var_t^a(\Delta \log \hat{h}) - \Delta cov_t^a(\log \hat{w}, \log \hat{h}) - \Delta var_t^a(\log \hat{h}) \\ = \; & 2\frac{(1+\widehat{\sigma})}{\widehat{\sigma}^2} v_{\theta,t-1}. \end{aligned}$$

   (c) Measurement error in hours is then identified from e.g.

   $$var_t^a\left(\Delta \log \hat{h}\right) - \Delta var_t^a\left(\log \hat{h}\right) = \frac{2}{\widehat{\sigma}^2} v_{\theta,t-1} + 2v_{\mu h}.$$

2. The parameter γ and the two sets of parameters $\{v_{\eta t}\}_{t=2}^{T}$ and $\{v_{\omega t}\}_{t=2}^{T}$ are then identified from within-cohort changes in the macro moments, $\Delta var_t^a\left(\log \hat{w}\right)$, $\Delta var_t^a\left(\log \hat{h}\right)$, and $\Delta cov_t^a\left(\log \hat{w}, \log \hat{h}\right)$, all available from $t = 2, ..., T$. These parameters can be identified recursively as follows:

3. Combine (24)-(25) to get

$$\frac{\left(\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)-\frac{1}{\widehat{\sigma}}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)^2}{\left(\Delta var_t^a\left(\log\hat{h}\right)-\frac{1}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)}=\frac{\left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2\left(v_{\omega t}\right)^2}{\left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 v_{\omega t}}=v_{\omega t}$$

Combine this with (23) to get an equation in $(v_{\eta t}+\Delta v_{\theta t})$,

$$\frac{\left(\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)-\frac{1}{\widehat{\sigma}}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)^2}{\left(\Delta var_t^a\left(\log\hat{h}\right)-\frac{1}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)}=\Delta var_t^a\left(\log\hat{w}\right)-\left(v_{\eta t}+\Delta v_{\theta t}\right).$$

Therefore, each element of the sequence $\{v_{\eta t}+\Delta v_{\theta t}\}_{t=2}^{T}$ is identified by[4]

$$\left(v_{\eta t}+\Delta v_{\theta t}\right)=\frac{\left(\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)\right)^2-\Delta var_t^a\left(\log\hat{w}\right)\cdot\Delta var_t^a\left(\log\hat{h}\right)}{\frac{\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)}{\widehat{\sigma}}-\frac{\Delta var_t^a\left(\log\hat{w}\right)}{\widehat{\sigma}^2}-\Delta var_t^a\left(\log\hat{h}\right)},$$

which, given $\{v_{\theta,t}\}_{t=1}^{T-1}$, pins down $\{v_{\eta,t}\}_{t=2}^{T-1}$.

(a) Given $v_{\eta t}+\Delta v_{\theta t}$, each element of the sequence $\{v_{\omega t}\}_{t=2}^{T}$ is identified by (23),

$$v_{\omega t}=\Delta var_t^a\left(\log\hat{w}\right)-\left(v_{\eta t}+\Delta v_{\theta t}\right).$$

(b) Given $v_{\eta t}+\Delta v_{\theta t}$ and $v_{\omega t}$, the risk aversion parameter γ is determined by (25) as the solution to the following equation:

$$\frac{1-\gamma}{\widehat{\sigma}+\gamma}=\frac{\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)}{v_{\omega t}}-\frac{1}{\widehat{\sigma}}\frac{\left(v_{\eta t}+\Delta v_{\theta t}\right)}{v_{\omega t}}.$$

[4]To see this, note that

$$\begin{aligned}
&\left(\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)-\frac{1}{\widehat{\sigma}}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)^2\\
=\;&\left(\Delta var_t^a\left(\log\hat{w}\right)-\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)\left(\Delta var_t^a\left(\log\hat{h}\right)-\frac{1}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)\right)\\
&\left(\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)\right)^2-\frac{\Delta cov_t^a\left(\log\hat{w},\log\hat{h}\right)}{\widehat{\sigma}}\left(v_{\eta t}+\Delta v_{\theta t}\right)+\frac{1}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)^2\\
=\;&\Delta var_t^a\left(\log\hat{w}\right)\cdot\Delta var_t^a\left(\log\hat{h}\right)-\frac{\Delta var_t^a\left(\log\hat{w}\right)}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)\\
&-\left(v_{\eta t}+\Delta v_{\theta t}\right)\cdot\Delta var_t^a\left(\log\hat{h}\right)+\frac{1}{\widehat{\sigma}^2}\left(v_{\eta t}+\Delta v_{\theta t}\right)^2
\end{aligned}$$

4. Given values for γ, $\widehat{\sigma}$, $\{v_{\theta t}\}_{t=1}^{T-1}$ and $\{v_{\mu h}, v_{\mu y}\}$, the following macro moments, available for all $t = 1, ..., T$ and evaluated for the youngest age group, identify the sequence of cohort effects $\{v_{\widehat{\varphi} t}, v_{\alpha^0 t}\}_{t=1}^{T}$, $\{v_{\kappa^0 t}\}_{t=2}^{T}$, and $(v_{\kappa^0 T} + v_{\theta T})$. We do it in two steps

   (a) First, the following two linearly independent macro moments, available for all $t = 1, ..., T$ and evaluated for the youngest age group, identify the sequence of cohort effects in insurable- and uninsurable initial wages, $\{v_{\alpha^0 t}\}_{t=1}^{T}$, $\{v_{\kappa^0 t}\}_{t=2}^{T}$, and $(v_{\kappa^0 T} + v_{\theta T})$,

$$\begin{aligned} var_t^0 (\log \hat{w}) &= v_{\alpha^0 t} + (v_{\kappa^0 t} + v_{\theta t}) + v_{\mu y} + v_{\mu h} \\ cov_t^0 \left(\log \hat{w}, \log \hat{h}\right) &= \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right) \cdot v_{\alpha^0 t} + \frac{1}{\widehat{\sigma}} (v_{\kappa^0 t} + v_{\theta t}) - v_{\mu h}. \end{aligned}$$

   (b) Finally, the following macro moments, available for all $t = 1, ..., T$ and evaluated for the youngest age group, identify the sequence of cohort effects in preference heterogeneity, $\{v_{\widehat{\varphi} t}\}_{t=1}^{T}$,

$$cov_t^0 \left(\log \hat{w}, \log \hat{h}\right) + var_t^0 \left(\log \hat{h}\right) = v_{\widehat{\varphi} t} + \frac{(1-\gamma)(1+\widehat{\sigma})}{(\widehat{\sigma}+\gamma)^2} v_{\alpha^0 t} + \frac{1+\widehat{\sigma}}{\widehat{\sigma}^2} (v_{\kappa^0 t} + v_{\theta t}),$$

   since every other parameter in those three moments is already known. This concludes the proof.

C.2 Extending Proposition 3 to biannual data

It is possible to extend Proposition 3 to allow for biannual panel data towards the end of the sample, so the proposition can be applied directly to the PSID. This amounts to combining Proposition 3 with Corollary 2.2. We state this formally as the following corollary.

Corollary 3.1 [extending Proposition 3 to biannual panel data] *Suppose one has access to an unbalanced panel on wages and hours, but no data on consumption. The panel data are available annually until year $\hat{t}$ and biannually thereafter, i.e. available for the years $t = 1, 2, ..., \hat{t}$ and $t = \hat{t}+2, \hat{t}+4, ..., T-2, T$. Suppose further that one has an exogenous estimate of measurement error in earnings, $v_{\mu y}$. Then, one can identify $\{\widehat{\sigma}, \gamma, v_{\mu h}, v_{\mu c}\}$, the sequences $\{v_{\widehat{\varphi} t}\}_{t=1}^{T}$, $\{v_{\omega t}, v_{\eta t}\}_{t=2}^{\hat{t}}$, $\{v_{\theta t}, v_{\kappa^0 t}, v_{\alpha^0 t}\}_{t=1}^{\hat{t}}$, and $\{v_{\theta t}, v_{\kappa^0 t}, v_{\alpha^0 t}, v_{\omega, t-1} + v_{\omega t}, v_{\eta, t-1} + v_{\eta t}\}$ for $t = \hat{t}+2, \hat{t}+4, ..., T-2$, as well as the sums $\{v_{\eta, T-1} + v_{\eta, T} + v_{\theta, T}\}$ and $\{v_{\kappa^0, T} + v_{\theta, T}\}$.*

Proof Start by following the proof of Proposition 3 for the years $t = 1, 2, ..., \hat{t}$. Consider then the time-varying parameters for the biannual sample, i.e., for $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2, T$. These parameters are identified in five sequential steps.

1. Identify $\{v_{\theta,t}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$, as well as $v_{\eta\hat{t}}$ and $v_{\kappa^0\hat{t}}$, by combining the following moments,

$$var_t^a\left(\Delta^2 \log \hat{w}\right) - \Delta^2 var_t^a\left(\log \hat{w}\right) = 2v_{\theta,t-2} + 2\left(v_{\mu y} + v_{\mu h}\right).$$

which is available for $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2, T$. Note that, since $v_{\theta,\hat{t}}$ is identified, so are $v_{\eta,\hat{t}}$ and $v_{\kappa^0\hat{t}}$.

2. Identify $\{v_{\eta t} + v_{\eta,t-1}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$, and the sum $(v_{\eta,T} + v_{\eta,T-1} + v_{\theta T})$. Start by combining the biannual versions of (24), (25), and (23) to get an equation where $(v_{\eta t} + v_{\eta,t-1} + \Delta^2 v_{\theta t})$ is the only unknown:

$$\begin{aligned}
&\frac{\left(\Delta^2 cov_t^a\left(\log \hat{w}, \log \hat{h}\right) - \frac{1}{\hat{\sigma}}\left(v_{\eta t} + v_{\eta,t-1} + \Delta^2 v_{\theta t}\right)\right)^2}{\Delta^2 var_t^a\left(\log \hat{h}\right) - \frac{1}{\hat{\sigma}^2}\left(v_{\eta t} + v_{\eta,t-1} + \Delta^2 v_{\theta t}\right)} \\
= &\frac{\left(\left(\frac{1-\gamma}{\hat{\sigma}+\gamma}\right)\left(v_{\omega t} + v_{\omega,t-1}\right)\right)^2}{\left(\frac{1-\gamma}{\hat{\sigma}+\gamma}\right)^2\left(v_{\omega t} + v_{\omega,t-1}\right)} = \left(v_{\omega t} + v_{\omega,t-1}\right) \\
= &\Delta^2 var_t^a\left(\log \hat{w}\right) - \left(v_{\eta t} + v_{\eta,t-1} + \Delta^2 v_{\theta t}\right).
\end{aligned}$$

This gives a linear equation in $(v_{\eta t} + v_{\eta,t-1} + \Delta^2 v_{\theta t})$,

$$\begin{aligned}
&\left(v_{\eta t} + v_{\eta,t-1} + v_{\theta t} - v_{\theta,t-2}\right) \\
= &\frac{\left(\Delta^2 cov_t^a\left(\log \hat{w}, \log \hat{h}\right)\right)^2 - \Delta^2 var_t^a\left(\log \hat{w}\right) \cdot \Delta^2 var_t^a\left(\log \hat{h}\right)}{\frac{1}{\hat{\sigma}} \cdot \Delta^2 cov_t^a\left(\log \hat{w}, \log \hat{h}\right) - \frac{1}{\hat{\sigma}^2} \cdot \Delta^2 var_t^a\left(\log \hat{w}\right) - \Delta^2 var_t^a\left(\log \hat{h}\right)}.
\end{aligned}$$

Since $\{v_{\theta t}\}$ is known for the years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$, this equation identifies $\{v_{\eta,t} + v_{\eta,t-1}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$, as well as the sum $(v_{\eta,T} + v_{\eta,T-1} + v_{\theta,T})$.

3. Given $\{v_{\eta,t} + v_{\eta,t-1}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$ and $(v_{\eta,T} + v_{\eta,T-1} + v_{\theta,T})$, the sequence of variances of uninsurable shocks $\{v_{\omega,t} + v_{\omega,t-1}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, \hat{t}+4, ..., T-2$ is identified from the growth in wage inequality:

$$\Delta^2 var_t^a\left(\log \hat{w}\right) = \left(v_{\omega t} + v_{\omega,t-1}\right) + \left(v_{\eta t} + v_{\eta,t-1} + v_{\theta t}\right) - v_{\theta,t-2}.$$

4. Consider now the cohort effects $\{v_{\alpha^0 t}, v_{\kappa^0 t}\}$ for the biannual years $t = \hat{t}, \hat{t}+2, ..., T$. The uninsurable component $v_{\alpha^0 t}$ is identified as

$$\begin{aligned} & \left(cov_t^0 \left(\log \hat{w}, \log \hat{h}\right) + v_{\mu h}\right) - \frac{1}{\widehat{\sigma}} \left(var_t^0 \left(\log \hat{w}\right) - \left(v_{\mu y} + v_{\mu h}\right)\right) \\ = & \frac{1-\gamma}{\widehat{\sigma}+\gamma} v_{\alpha^0 t} + \frac{1}{\widehat{\sigma}} \left(v_{\kappa^0 t} + v_{\theta t}\right) - \frac{1}{\widehat{\sigma}} \left(v_{\alpha^0 t} + \left(v_{\kappa^0 t} + v_{\theta t}\right)\right) \\ = & \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma} - \frac{1}{\widehat{\sigma}}\right) v_{\alpha^0 t}, \end{aligned}$$

which is available for $t = \hat{t}, \hat{t}+2, ..., T$. The wage inequality for new cohorts then identify the variance of the insurable cohort effect $\{v_{\kappa^0,t}\}$:

$$var_t^0 \left(\log \hat{w}\right) = v_{\alpha^0 t} + \left(v_{\kappa^0 t} + v_{\theta t}\right) + v_{\mu y} + v_{\mu h},$$

which is available for $t = \hat{t}, \hat{t}+2, ..., T-2$ since the other components on the right-hand side are known those years. For the final year $t = T$ we can only identify the sum $(v_{\kappa^0,T} + v_{\theta T})$.

5. Finally, the cohort effects $\{v_{\widehat{\varphi},t}\}_{t=\hat{t}}^{T}$ are identified by

$$var_t^0 \left(\log \hat{h}\right) = v_{\widehat{\varphi} t} + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 v_{\alpha^0 t} + \frac{1}{\widehat{\sigma}^2} \left(v_{\kappa^0 t} + v_{\theta t}\right) + v_{\mu h}$$

for the biannual years $t = \hat{t}, \hat{t}+2, ..., T$, and by

$$var_t^1 \left(\log \hat{h}\right) = v_{\widehat{\varphi},t-1} + \left(\frac{1-\gamma}{\widehat{\sigma}+\gamma}\right)^2 \left(v_{\alpha^0,t-1} + v_{\omega t}\right) + \frac{1}{\widehat{\sigma}^2} \left(v_{\kappa^0,t-1} + v_{\eta t} + v_{\theta t}\right) + v_{\mu h},$$

available at $t = \hat{t}, \hat{t}+2, ..., T-2, T$, to identify $\{v_{\widehat{\varphi},t}\}_{t=\hat{t}}^{T}$ in the in-between years.

D Test of model specification

Recall the notation of Section 4.3 of the paper. Λ is an $N \times 1$ vector of *parameters* of the model, $\mathbf{m}(\Lambda)$ is a $J \times 1$ vector of *theoretical moments* and $\hat{\mathbf{m}}$ is a $J \times 1$ vector of *empirical moments* from the data. The null hypothesis we want to test is that the model is "correctly" specified. For this purpose, we construct the Sargan test-statistic

$$\mathcal{T} = (\hat{\mathbf{m}} - \mathbf{m}(\hat{\Lambda}))' \hat{\mathbf{V}}^{-1} (\hat{\mathbf{m}} - \mathbf{m}(\hat{\Lambda}))$$

which, under the null hypothesis, is distributed as a χ^2_{J-N}.

$\hat{\mathbf{V}}$ is a consistent estimate of the fourth moments matrix. To make things concrete, we give an example of an element of $\hat{\mathbf{V}}$. The second moments used in estimation are of form $cov(x^i_{a,t}, y^i_{a,t})$. Consider two such moments from the PSID, for example $cov(w^i_{22,1990}, h^i_{22,1990})$ and $var(\Delta w^i_{26,1994})$. The corresponding entry in $\hat{\mathbf{V}}$ is computed as follows, where K is number of individuals with non-missing observations in the age/year cells needed to compute this entry:

$$\frac{1}{K}\sum_{i=1}^{K}\left[w^i_{22,1990}h^i_{22,1990} - cov(w^i_{22,1990}, h^i_{22,1990})\right] \times \left[\left(\Delta w^i_{26,1994}\right)^2 - var(\Delta w^i_{26,1994})\right].$$

In the baseline model, $N = 164$ and $J = 11,532$. The test statistic implies a $p-$value of 0.9991.

In the context of GMM estimation with dynamic panel data, it is known that overidentifying restrictions (OID) tests can have low power when the number of moment conditions is large relative to the number of observations used to calculate the empirical moments. Bowsher (2002) offers compelling evidence based on Monte Carlo experiments. The issue associated with the use of too many moments is attributed to the need to estimate $J(J+1)/2$ separate entries of the fourth moment matrix when J moment conditions are used. Intuitively, if this dimensionality is large relative to the sample size, the estimates of the $\mathbf{V}$ matrix may be poor. In the baseline case, $J-N$ is equal to $11,368$ and is therefore an order of magnitude larger than the average number of observations per moment. One may worry that the power of the $\mathcal{T}$ statistic is artificially low and the test result not too informative.

Bowsher (2002) and Roodman (2009) propose a method to restore power to the Sargan test. This method consists of testing the validity of a reduced number of moment conditions; the full set of moments is still used in estimation but a subset of moments is used in the construction of the test-statistic. For our purposes, one way of *collapsing* the number of moment conditions used in the test is to consider covariances by age and by year. For example, instead of using the $A \times T$ covariances $\{cov(w^i_{a,t}, h^i_{a,t})\}_{a=1,\ldots A, t=1,\ldots T}$, we restrict ourselves to just $A+T$ covariances: $\{cov(w^i_a, h^i_a)\}_{a=1,\ldots A}$ and $\{cov(w^i_t, h^i_t)\}_{t=1,\ldots T}$. And similarly, for all other moments. Doing so reduces the number of degrees of freedom from $11,368$ to 580 and increases the mean number of observations per moment from 285 to $8,291$ for PSID and from 179 to $5,318$. Therefore, the source of low power of the OID test should be much weaker under this "collapsed moments" version of the test. The test continues to yield a $p-$value above 0.99, failing to reject the null.

E Model Estimation: Additional Figures and Tables

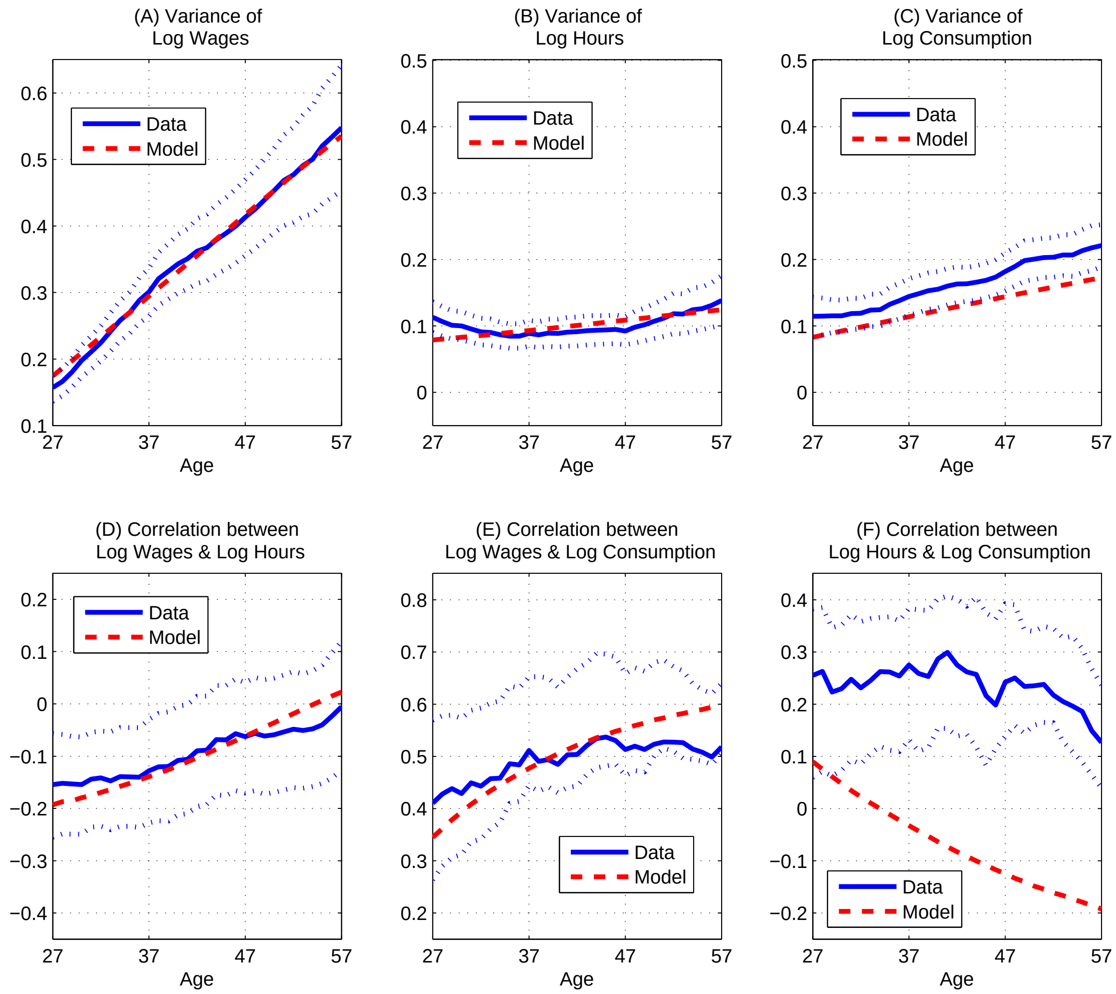

Figure E1: Estimation without CEX data. Data and model fit for moments in levels along the age dimension. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

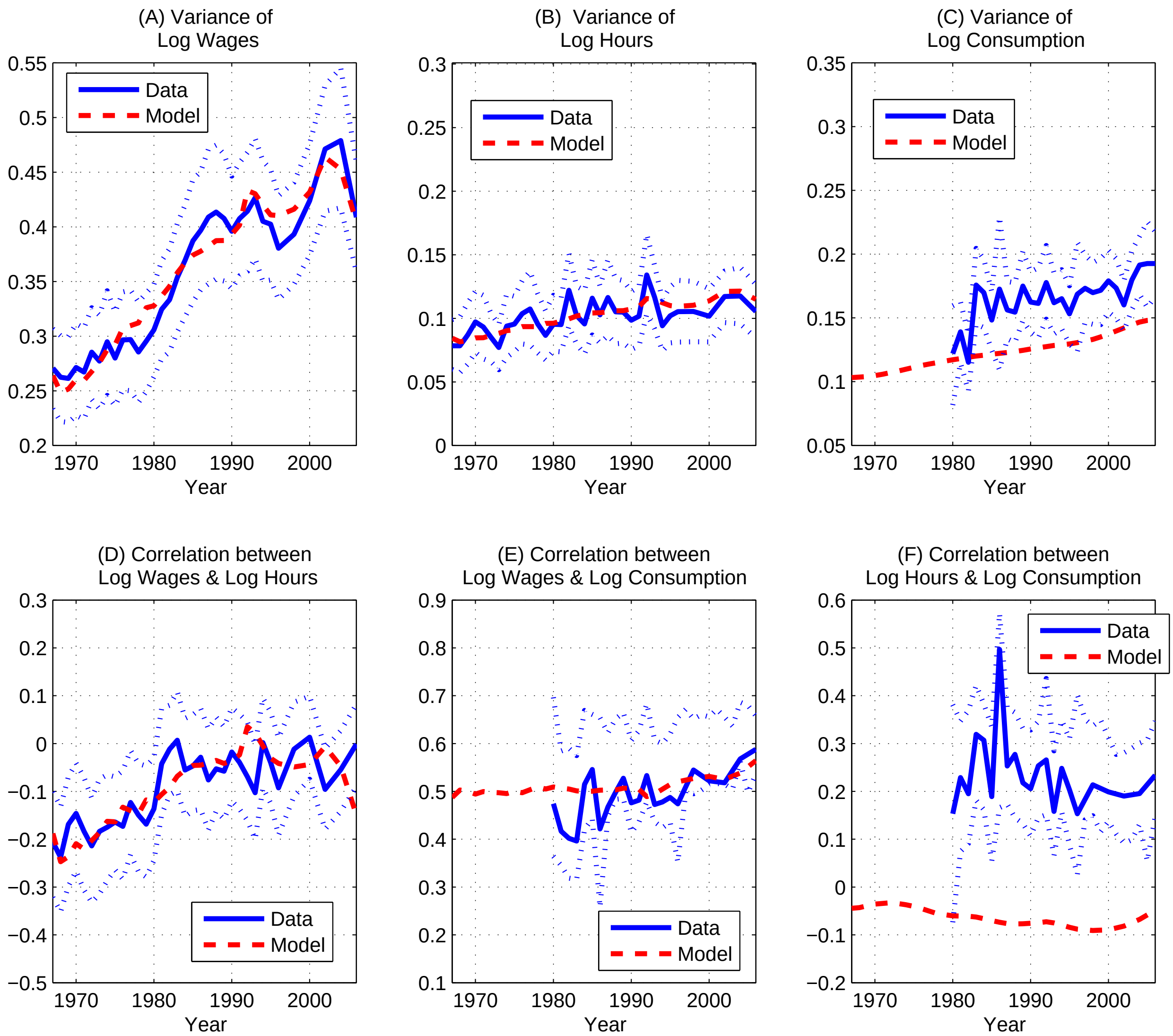

Figure E2: Estimation without CEX data. Data and model fit for moments in levels along the time dimension. Dotted lines denote 90–10 bootstrapped confidence intervals for the empirical moments.

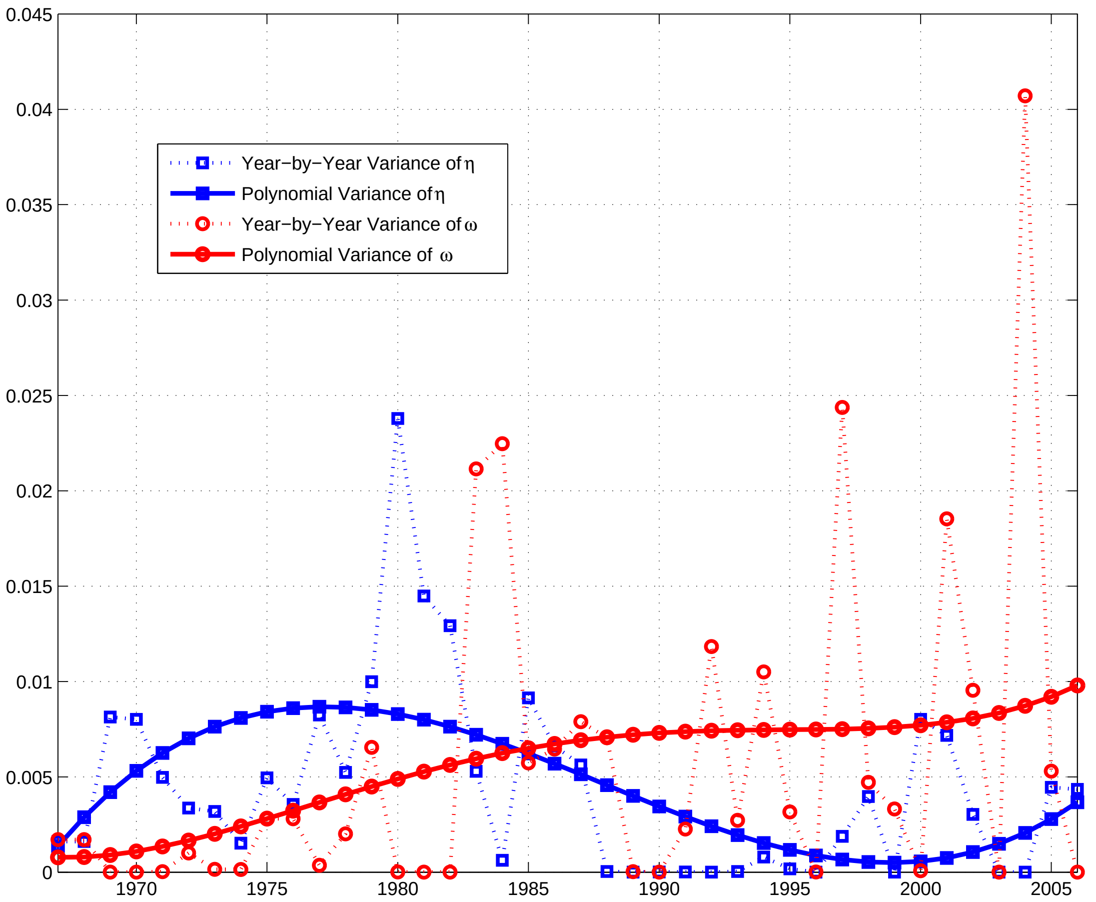

Figure E3: Comparison between (i) the baseline estimates for the variances of the insurable ($v_{\eta t}$) and uninsurable ($v_{\omega t}$) innovations modeled as time-polynomials and (ii) estimates for the same variances modeled as unrestricted sequences.

Table E1: Parameter Estimates: Baseline Model

	$v_{\alpha^0 t}$	$v_{\kappa^0 t}$	$v_{\widehat{\varphi} t}$	$v_{\omega t}$	$v_{\eta t}$	$v_{\theta t}$
1967	0.136	0.029	0.045	0.0008	0.0014	0.027
	(0.020)	(0.020)	(0.009)	(0.0005)	(0.0010)	(0.005)
1968	0.001	0.000	0.089	0.0008	0.0029	0.010
	(0.019)	(0.003)	(0.025)	(0.0004)	(0.0007)	(0.003)
1969	0.037	0.001	0.054	0.0009	0.0042	0.011
	(0.026)	(0.008)	(0.013)	(0.0005)	(0.0008)	(0.003)
1970	0.140	0.000	0.049	0.0011	0.0053	0.018
	(0.037)	(0.003)	(0.015)	(0.0006)	(0.0009)	(0.003)
1971	0.156	0.000	0.049	0.0013	0.0063	0.013
	(0.047)	(0.010)	(0.014)	(0.0007)	(0.0010)	(0.003)
1972	0.133	0.003	0.051	0.0017	0.0070	0.017
	(0.034)	(0.015)	(0.013)	(0.0008)	(0.0010)	(0.003)
1973	0.001	0.001	0.055	0.0020	0.0076	0.021
	(0.010)	(0.008)	(0.019)	(0.0009)	(0.0010)	(0.003)
1974	0.047	0.138	0.059	0.0024	0.0081	0.025
	(0.034)	(0.047)	(0.013)	(0.0009)	(0.0010)	(0.004)
1975	0.150	0.014	0.058	0.0028	0.0084	0.022
	(0.034)	(0.015)	(0.014)	(0.0009)	(0.0010)	(0.004)
1976	0.152	0.052	0.036	0.0032	0.0086	0.029
	(0.032)	(0.031)	(0.013)	(0.0008)	(0.0010)	(0.004)
1977	0.089	0.000	0.039	0.0037	0.0087	0.023
	(0.030)	(0.012)	(0.011)	(0.0008)	(0.0010)	(0.004)
1978	0.008	0.044	0.055	0.0041	0.0086	0.018
	(0.014)	(0.025)	(0.019)	(0.0008)	(0.0010)	(0.003)
1979	0.038	0.080	0.070	0.0045	0.0085	0.025
	(0.030)	(0.029)	(0.016)	(0.0007)	(0.0010)	(0.004)
1980	0.068	0.016	0.052	0.0049	0.0083	0.019
	(0.030)	(0.014)	(0.013)	(0.0007)	(0.0010)	(0.004)
1981	0.143	0.030	0.046	0.0053	0.0080	0.021
	(0.029)	(0.020)	(0.014)	(0.0007)	(0.0011)	(0.003)
1982	0.152	0.040	0.042	0.0056	0.0076	0.023
	(0.029)	(0.024)	(0.011)	(0.0007)	(0.0011)	(0.004)
1983	0.115	0.054	0.047	0.0059	0.0072	0.031
	(0.019)	(0.031)	(0.018)	(0.0007)	(0.0011)	(0.005)
1984	0.037	0.102	0.051	0.0062	0.0067	0.035
	(0.029)	(0.036)	(0.012)	(0.0007)	(0.0011)	(0.005)
1985	0.068	0.070	0.058	0.0065	0.0062	0.036
	(0.030)	(0.030)	(0.013)	(0.0006)	(0.0011)	(0.004)
1986	0.136	0.097	0.050	0.0067	0.0057	0.034
	(0.028)	(0.032)	(0.014)	(0.0006)	(0.0011)	(0.005)
1987	0.149	0.078	0.051	0.0069	0.0051	0.034
	(0.020)	(0.028)	(0.013)	(0.0006)	(0.0010)	(0.004)
1988	0.105	0.044	0.056	0.0071	0.0046	0.036
	(0.020)	(0.027)	(0.020)	(0.0005)	(0.0010)	(0.005)
1989	0.053	0.098	0.062	0.0072	0.0040	0.033
	(0.029)	(0.035)	(0.013)	(0.0005)	(0.0009)	(0.004)

Note: Bootstrapped standard errors based on 500 replications in parenthesis.

Table E1: (Continued) Parameter Estimates: Baseline Model

	$v_{\alpha^0 t}$	$v_{\kappa^0 t}$	$v_{\widehat{\varphi} t}$	$v_{\omega t}$	$v_{\eta t}$	$v_{\theta t}$
1990	0.057	0.078	0.039	0.0073	0.0034	0.036
	(0.028)	(0.030)	(0.011)	(0.0005)	(0.0009)	(0.005)
1991	0.138	0.077	0.055	0.0074	0.0029	0.042
	(0.028)	(0.033)	(0.013)	(0.0005)	(0.0008)	(0.005)
1992	0.152	0.000	0.048	0.0074	0.0024	0.075
	(0.032)	(0.030)	(0.013)	(0.0005)	(0.0008)	(0.007)
1993	0.058	0.071	0.038	0.0074	0.0019	0.072
	(0.022)	(0.029)	(0.017)	(0.0005)	(0.0008)	(0.006)
1994	0.071	0.027	0.040	0.0075	0.0015	0.061
	(0.031)	(0.022)	(0.012)	(0.0006)	(0.0008)	(0.007)
1995	0.182	0.059	0.057	0.0075	0.0012	0.051
	(0.033)	(0.029)	(0.017)	(0.0007)	(0.0009)	(0.006)
1996	0.160	0.063	0.051	0.0075	0.0009	0.052
	(0.031)	(0.032)	(0.016)	(0.0007)	(0.0009)	(0.006)
1997	0.180	0.154	0.062	0.0075	0.0007	0.055
	(0.035)	(0.019)	(0.016)	(0.0008)	(0.0009)	(0.004)
1998	0.024	0.044	0.060	0.0075	0.0005	0.059
	(0.025)	(0.022)	(0.022)	(0.0008)	(0.0009)	(0.007)
1999	0.134	0.036	0.057	0.0076	0.0005	0.066
	(0.037)	(0.016)	(0.015)	(0.0008)	(0.0010)	(0.005)
2000	0.154	0.029	0.063	0.0077	0.0006	0.074
	(0.032)	(0.023)	(0.018)	(0.0008)	(0.0011)	(0.008)
2001	0.129	0.031	0.067	0.0079	0.0007	0.091
	(0.033)	(0.018)	(0.018)	(0.0008)	(0.0013)	(0.008)
2002	0.188	0.033	0.068	0.0081	0.0011	0.107
	(0.035)	(0.022)	(0.017)	(0.0009)	(0.0016)	(0.011)
2003	0.047	0.057	0.061	0.0084	0.0015	0.101
	(0.029)	(0.018)	(0.022)	(0.0010)	(0.0020)	(0.008)
2004	0.098	0.060	0.053	0.0087	0.0021	0.094
	(0.039)	(0.028)	(0.024)	(0.0013)	(0.0025)	(0.009)
2005	0.030	0.084	0.085	0.0092	0.0028	0.066
	(0.040)	(0.027)	(0.027)	(0.0018)	(0.0034)	(0.008)
2006	0.123	0.109	0.048	0.0098	0.0037	0.037
	(0.039)	(0.028)	(0.014)	(0.0023)	(0.0045)	(0.013)

Note: Bootstrapped standard errors based on 500 replications in parenthesis.