LISER

LUXEMBOURG INSTITUTE OF SOCIO-ECONOMIC RESEARCH

n° 2021-04

May 2021

WORKING PAPERS

Consumption Inequality across Heterogeneous Families

Alexandros **THELOUDIS**[1,2]

[1] Luxembourg Institute of Socio-Economic Research (LISER), Luxembourg
[2] Tilburg University, Netherlands

www.liser.lu

Consumption Inequality across Heterogeneous Families[*]

Alexandros Theloudis[**]

May 6, 2021

Abstract

What does preference heterogeneity imply for consumption inequality? This paper studies the link from wage to consumption inequality within a lifecycle model of consumption and family labor supply. Its distinctive feature is that households have general heterogeneous preferences over consumption and labor supply. The paper shows identification of the joint distribution of unobserved household preferences separately from the observed distributions of incomes and outcomes. Estimation on data from the Panel Study of Income Dynamics in the US reveals substantial unexplained heterogeneity in consumption preferences but little unexplained heterogeneity in labor supply preferences. Preference heterogeneity accounts for about a third of consumption inequality in recent years and implies, on average, lower partial insurance of wage shocks compared to recent studies in the literature.

Keywords: unobserved preference heterogeneity, family labor supply, lifecycle model, partial insurance, PSID

JEL classification: D12, D30, D91, E21

[*]This working paper supersedes the previous version of the article, LISER Working Paper 2017-18.

[**]Alexandros Theloudis: Luxembourg Institute for Socio-Economic Research, and Department of Econometrics and Operations Research, Tilburg University (email: a.theloudis@gmail.com). This paper is based on the third chapter of my dissertation (Theloudis, 2017b). I am grateful to the editor Robert Sauer, an associate editor, and three referees for their detailed comments. I thank Richard Blundell, Pierre-André Chiappori, Sam Cosaert, Jan Eeckhout, Eric French, Jonathan Heathcote, Christos Koulovatianos, Magne Mogstad, Luigi Pista-ferri, Ian Preston, Itay Saporta-Eksten and participants in various seminars and conferences for many helpful comments. Thanks also to Dirk Krueger and Chunzan Wu who provided me with the programs that simulate a quantitative consumption model with endogenous family labor supply. I acknowledge financial support from a Royal Economic Society Junior Fellowship and R.Blundell's ERC advanced grant 'MicroConLab'.

1 Introduction

This paper studies the link from wage to consumption inequality using a lifecycle model of family labor supply, consumption and wealth. Its distinctive feature is that households have heterogeneous preferences over consumption and labor supply. The paper shows identification of the joint distribution of unobserved household preferences, namely consumption and labor supply Frisch elasticities, separately from the observed distributions of incomes and outcomes. It then studies how preference heterogeneity matters for the transmission of wage into consumption inequality. The model is implemented empirically on recent data from the Panel Study of Income Dynamics (PSID) revealing large amounts of consumption preference heterogeneity across households with substantial implications for inequality and partial insurance.

There is extensive empirical, experimental and survey evidence of preference heterogeneity.[1] Such heterogeneity has potentially important implications for consumption inequality. Consider two households who have similar wages, wealth and demographics but differ in their respective consumption preferences. This preference heterogeneity may reflect differences in the composition of their consumption baskets, in the complementarity between consumption goods and leisure, or a myriad other aspects. The households will likely adjust their consumption *differently* in response to a similar wage change and subsequent consumption inequality in this simple cross-section will reflect both the wage change *and* preference heterogeneity. This has implications for how much of inequality is policy relevant, for how inequality responds to redistributive policies, and for the extent and distribution of consumption insurance to wage fluctuations in the economy.

Family labor supply matters for the transmission of wage into consumption inequality. A model with endogenous labor supply implies the right degree of comovement between consumption and income (Wu and Krueger, 2021), while a setting with exogenous labor supply typically does not (Kaplan and Violante, 2010). Wu and Krueger (2021) show that the bulk of consumption insurance against wage changes is provided through the *added worker* in the household, namely the income and labor supply of the secondary earner. It then follows that in a world with heterogeneous preferences across households, heterogeneity in the secondary earner's labor supply gives rise to heterogeneity in the household response to wage changes. Furthermore, the way such heterogeneity relates to other aspects of household preferences (for example, to the primary earner's labor supply) may further matter for the extent and distribution of consumption insurance to wage fluctuations, therefore also for the transmission of wage into consumption inequality.

[1] For example, Alan and Browning (2010) find heterogeneity in the discount factor and the elasticity of intertemporal substitution across education groups in the PSID. Andersen et al. (2008) and other experimental studies find substantial dispersion in risk and time preferences while Guiso and Paiella (2008) observe directly from survey data large amounts of unexplained heterogeneity in risky preferences. Abowd and Card (1989) find large unexplained dispersion in working hours at fixed wages. See Heckman (2001) for a theoretical discussion.

This paper incorporates unobserved preference heterogeneity into a lifecycle model for consumption and labor supply of two-earner households. The specific workings of the household are as follows. Two spouses make unitary lifecycle choices over consumption and labor supply, while they also save and possibly borrow at the market interest rate. For each hour of work, they earn a wage that is subject to wage/productivity shocks of various persistence. The treatment of unobserved heterogeneity is general: (i) heterogeneity is non-separable from within-period consumption and labor supply preferences; (ii) preferences are nonparametric; (iii) heterogeneity is not restricted to a single dimension (to a single parameter in the analog of parametric preferences); instead it is multi-dimensional meaning that any preference parameter in the parametric analog might be independently or jointly heterogeneous; (iv) the multivariate distribution of preferences is itself nonparametric.

Following Blundell and Preston (1998) and several papers thereafter, I derive analytical expressions for consumption and labor supply by approximating the lifetime budget constraint and the problem's optimality conditions. The analytical expressions relate the growth rates of consumption and hours to wage shocks, preferences (household-specific Frisch elasticities of consumption and labor supply), and financial and human wealth in the household. This approximation has been shown to perform well in a variety of contexts (Blundell et al., 2013). Thanks to these expressions, the second and higher moments of the empirical joint distribution of consumption, earnings and wages have straightforward theoretical counterparts. This link between data and theory provides restrictions that can be used to identify moments of the cross-sectional distribution of household preferences. Estimation of preferences is then by means of estimating a demand-like system based on these analytical expressions.

I show that *any* moment of the distribution of wage elasticities of consumption and labor supply is identified separately from the distribution of consumption, hours, wages, and wealth. Such elasticities describe preferences in an ordinal way and are not specific to a particular parametrization of the household utility function. Identification rests on the idea that cross-sectional heterogeneity in consumption and hours growth that occurs at fixed prices, observables, and initial levels of consumption and hours, masks heterogeneity in consumption and labor supply preferences. Identification requires panel data on consumption, hours and wages.

To illustrate these points empirically I fit second and third moments of the joint distribution of consumption, earnings and wages in the PSID in years 1999-2011. This permits the estimation of up to third moments of wage shocks and up to second moments of preferences (wage elasticities of consumption and labor supply). The model fits the data reasonably well. There are four main new findings from this exercise.

First, there is substantial unexplained heterogeneity in consumption elasticities unrelated to the levels of consumption the elasticities naturally depend on. By contrast, unexplained heterogeneity in labor supply elasticities is at least an order of magnitude smaller. Consumption in the average household (one with average preferences) is separable from leisure. However, there is substantial cross-household heterogeneity in the magnitude and sign of the consumption-leisure complementarity; for example, one standard deviation of the wage elasticities of consumption about the mean is in the range $(-0.54, 0.43)$. Average male and female labor supply elasticities are lower than average values reported by Keane (2011). Their magnitudes drop as the model attempts to match third moments of consumption and earnings.

Second, preference heterogeneity magnifies the transmission of wage into consumption inequality. I decompose consumption inequality into terms that pertain to wage inequality, wealth inequality, and preference heterogeneity; this allows me to sign how preference heterogeneity affects consumption inequality. The decomposition suggests that preference heterogeneity accounts for approximately 31% of unexplained consumption inequality across US households, wage inequality for 47% and wealth inequality for 22%. Preferences and wealth amount together to a combined 53%, similar to the contribution of preferences and household-specific initial conditions to the variance of consumption in Heathcote et al. (2014).

Third, on average consumption tracks permanent wage shocks more closely than the preference *homogeneity* benchmarks of Blundell et al. (2016) and Wu and Krueger (2021). Consumption partial insurance, namely the extent to which consumption is insured against shocks, is lower here because of a limited insurance role of family labor supply. As the model matches third moments of the data not targeted in the homogeneous case, labor supply elasticities drop and family labor supply becomes less effective in insuring against shocks. While on average 45% (27%) of a male (female) permanent shock passes through to consumption compared to 34% (20%) in Blundell et al. (2016), a non-negligible fraction of households lacks partial insurance as in Hryshko and Manovskii (2017).

Fourth, the distributions of most shocks to male and female wages have a long left tail. This negative skewness implies that negative shocks are more unsettling than positive ones as they are on average further away from the zero mean. This is similar to Guvenen et al. (2015) who find negative skewness of *earnings* shocks in US Social Security Administration data.

Possible misspecification in the model may confound preference heterogeneity with heterogeneity in omitted variables. I argue that this is not the case in regard to wage heterogeneity, time aggregation in wages, taxes or home production, all of which the baseline model abstracts from. I also explore the implications of unobserved liquidity constraints and adjustment costs of work. Finally, I use the parametric model of Wu and Krueger (2021) to show that the approximation method I use here correctly recovers many aspects of the underlying preferences even in the presence of pervasive preference heterogeneity.

Contribution and literature. The paper relates to the literature that studies the mapping from income to consumption inequality and the consumption response to income changes. The literature *either* allows for heterogeneity in the transmission of income into consumption but abstracts from family labor supply *or* allows for family labor supply but abstracts from heterogeneity. This paper offers four contributions: (i) it introduces general preference heterogeneity and family labor supply in a consumption lifecycle model; given the importance of heterogeneity for how much of inequality is policy relevant and the importance of family labor supply for the household response to wage shocks, it is crucial to allow for both features; (ii) it establishes identification of any moment of the distribution of wage Frisch elasticities of consumption and labor supply; (iii) it estimates the location and spread of these elasticities in the US; (iv) it quantifies their implications for consumption inequality and partial insurance.

The relationship between income and consumption inequality is the focus of a large literature. Deaton and Paxson (1994), Blundell and Preston (1998), Krueger and Perri (2006) and Primiceri and van Rens (2009) investigate the link from income into consumption inequality, while Blundell et al. (2008), Kaplan and Violante (2010), and Guvenen and Smith (2014) measure the degree of consumption insurance. These papers abstract from labor supply so they are agnostic about the nature of idiosyncratic earnings fluctuations or the insurance possibilities family labor supply offers. Hyslop (2001) and Attanasio et al. (2002) introduce family labor supply and show that it matters for the transmission of wage into earnings and consumption inequality respectively. Attanasio et al. (2005) quantify the substantial insurance female labor supply offers when male earnings uncertainty rises. Huggett et al. (2011) develop an one-earner lifecycle model with labor supply to study the relative contributions of luck (wage shocks) and initial conditions (wealth, human capital) to inequality.[2]

Blundell, Pistaferri, and Saporta-Eksten (2016), BPS hereafter, study the transmission of wage shocks into consumption in a model of family labor supply, savings, and external insurance. BPS estimate preferences homogeneously using a nonparametric demand-like system of analytical expressions for consumption and labor supply. They find that once family labor supply, wealth and transfers are accounted for, there is little room for additional insurance. Wu and Krueger (2021), WK hereafter, confirm their findings in a parametric setting. The model in this paper is conceptually similar to BPS and WK so several results subsequently are compared to theirs. The present paper, however, extends these studies to allow households to differ in their consumption and labor supply preferences, therefore also in their willingness to smooth consumption. It employs similar analytical expressions to BPS which, unlike a fully specified structural model, do not require to restrict preferences or wages to a particular distribution. Using the same data (but augmented by one wave and using additional moments), heterogeneity turns out to be important for both consumption inequality and partial insurance.

With the exception of the following three studies, a consistent feature in this literature is that households are ex ante identical: conditional on observables and idiosyncratic incomes, any two households behave the same when confronted with a given income change. This is clearly counterfactual given the extensive evidence of preference heterogeneity. Alan et al. (2018) allow for parametric heterogeneity in income and preferences while Arellano et al. (2017) develop a nonlinear framework for the consumption response to income allowing for flexible heterogeneity.

[2] Blundell and Etheridge (2010) and Heathcote et al. (2010) provide additional empirical evidence on the role of labor supply in the transmission of inequality. Jappelli and Pistaferri (2010) and Meghir and Pistaferri (2011) provide an overview of the extensive literature in income and consumption inequality.

Both studies find large amounts of heterogeneity in income risk and in the response of consumption to it. However, they both abstract from family labor supply. So what they estimate as income heterogeneity may actually reflect heterogeneity in family labor supply which may also project into consumption (e.g. through leisure-consumption complementarities). Heathcote et al. (2014) admit that part of the cross-sectional dispersion in consumption and hours is unrelated to income or price variation; they allow for individual labor supply and unobserved heterogeneity which is, however, additively separable and specific to the parametrization of preferences they employ. By contrast, I allow for joint heterogeneity in family labor supply and consumption while leaving preferences and their distribution nonparametric.

Finally, the paper shares a common goal with the extensive literature in consumer demand, namely the identification of preferences from observed behavior.[3] This paper identifies first and higher moments of consumption and hours elasticities and estimates a subset of them. These novel parameters can inform studies of welfare programs (French, 2005), inequality (De Nardi et al., 2019) or optimal taxation (Gayle and Shephard, 2019), where heterogeneity in the behavioral response of consumption or labor supply may crucially affect the policy conclusions.

The paper is organized as follows. Section 2 presents the model and the analytical expressions for consumption and hours. Section 3 establishes identification, section 4 presents the empirical implementation, and section 5 shows the main results. Section 6 discusses the implications of preference heterogeneity for consumption inequality and partial insurance; it also investigates alternative explanations and the robustness of the results. Section 7 concludes. An online appendix includes technical details and additional results.

2 Lifecycle Model with Two Earners

A household consists of two earners, each one subscripted by j. To fix ideas suppose the two earners are a male ($j = 1$) and a female ($j = 2$) spouse. In lifecycle period t the spouses make choices over household consumption C_t, future assets A_{t+1}, and hours of market work H_{1t} and H_{2t} respectively (intensive margin labor supply only). I model the household problem as unitary, that is, as the problem of a single economic agent. This facilitates the discussion of cross-household preference heterogeneity without confounding it with issues pertaining to intra-household heterogeneity and commitment. The length of the lifecycle is a known T.

Household i in the cross-section chooses $\{C_{it}, A_{it+1}, H_{1it}, H_{2it}\}$ over its lifecycle to maximize its expected discounted lifetime utility

$$\max_{\{C_{it},A_{it+1},H_{1it},H_{2it}\}_{t=0}^{T}} \mathbb{E}_0 \sum_{t=0}^{T} U_{it}(C_{it}, H_{1it}, H_{2it}; \mathbf{Z}_{it}) \tag{1}$$

[3] Several recent studies, e.g. Matzkin (2003), Blundell et al. (2017), Lewbel and Pendakur (2017), Cosaert and Demuynck (2018), present identification results and empirical applications when preferences exhibit non-separable heterogeneity as in this paper. Lewbel (2001) argues that the usual practice to restrict heterogeneity to additive errors is similar to enforcing a representative agent assumption.

subject to a lifetime budget constraint, the sequential version of which is

$$P_t(1+r_t)A_{it} + \sum_{j=1}^{2} W_{jit}H_{jit} = P_tC_{it} + P_tA_{it+1}. \tag{2}$$

In the budget constraint, P_t is the deterministic market price of consumption (the numeraire in the empirical application), r_t is the deterministic market interest rate, A_{it} is beginning-of-period assets, and W_{jit} is spouse j's stochastic hourly wage in the labor market.

In the objective function, U_{it} is household utility from consumption and labor supply. The dependence on time reflects discounting. Vector $\mathbf{Z}_{it}$ includes observable taste shifters, such as spouses' education or age, and captures *observed* preference heterogeneity. In addition, preferences U_{it} are subscripted by i to reflect *unobserved* preference heterogeneity across households, namely household-specific preferences not captured by the conditioning observables. This general way to model unobserved heterogeneity is consistent with various different sources heterogeneity may stem from, such as cross-household differences in the composition of the consumption basket (Aguiar and Hurst, 2013), in labor market attachment (Attanasio et al., 2018), consumption-leisure complementarity, risk or time preferences. I do not parameterize U_{it} but I assume geometric discounting and continuous first and second order derivatives.

I model wages as a permanent-transitory process.[4] Specifically, log wage $\ln W_{jit}$ is the sum of a deterministic component, a permanent stochastic component that follows a unit root, and a transitory shock; namely

$$\begin{aligned} \ln W_{jit} &= \mathbf{X}'_{jit}\boldsymbol{\alpha}_{W_j} + \ln W^p_{jit} + u_{jit} \\ \ln W^p_{jit} &= \ln W^p_{jit-1} + v_{jit}. \end{aligned}$$

Here $\mathbf{X}_{jit}$ is a vector of observables (age, education, etc.) with coefficient $\boldsymbol{\alpha}_{W_j}$. $\ln W^p_{jit}$ is the permanent component, u_{jit} is the transitory shock, and v_{jit} is the permanent shock for spouse $j = \{1, 2\}$ in household i at time $t = \{1, \dots, T\}$. This process can be written compactly as

$$\Delta w_{jit} = v_{jit} + \Delta u_{jit} \tag{3}$$

where $\Delta w_{jit} = \Delta \ln W_{jit} - \Delta \mathbf{X}'_{jit}\boldsymbol{\alpha}_{W_j}$ and $\Delta(\cdot)$ denotes first difference. The permanent shock reflects a permanent change in the returns to one's skills in the labor market such as a skill-specific technical change; the transitory shock indicates short-lived mean reverting fluctuations in productivity. Wage shocks are the only source of uncertainty the household encounters. In section 6.3 I explore the implications of a general wage process with heterogeneous persistence and the implications of neglected time aggregation in wages in the PSID.

[4]The permanent-transitory process has been used extensively in the income dynamics literature and beyond, for example in MaCurdy (1982), Abowd and Card (1989), Attanasio et al. (2002), Meghir and Pistaferri (2004), Attanasio et al. (2008), Blundell et al. (2008) and BPS. In this process agents have the same ex-ante income growth conditional on observables but face idiosyncratic shocks. An alternative family of income processes assumes ex-ante idiosyncratic income growth (see, for example, Guvenen, 2007; Browning et al., 2010).

Properties of shocks. Wage shocks are idiosyncratic (with zero cross-sectional mean) and possibly non-normal as in De Nardi et al. (2019). Their n^{th} moments ($n > 1$) are given by

$$\mathbb{E}(v_{jit}^{\nu} v_{j'it+s}^{n-\nu}) = \begin{cases} m_{v_j^{\nu} v_{j'}^{n-\nu}(t)} & \text{for } s = 0 \text{ and } \nu = \{1, \ldots, n\} \\ 0 & \text{for } s \neq 0 \end{cases}$$

$$\mathbb{E}(u_{jit}^{\nu} u_{j'it+s}^{n-\nu}) = \begin{cases} m_{u_j^{\nu} u_{j'}^{n-\nu}(t)} & \text{for } s = 0 \text{ and } \nu = \{1, \ldots, n\} \\ 0 & \text{for } s \neq 0 \end{cases}$$

and v_{jit} independent of $u_{j'it+s}$ for any combination of $j, j' = \{1, 2\}$ and $s = \{-t+1, \ldots, T-t\}$.

As an illustration for $s = 0$ and $j \neq j'$, the second moments ($n = 2$) of shocks are given by

$$\mathbb{E}(v_{jit}^{\nu} v_{j'it}^{2-\nu}) = \begin{cases} \sigma_{v_j(t)}^2 & \text{if } \nu = 2 \\ \sigma_{v_1 v_2(t)} & \text{if } \nu = 1 \end{cases}$$

$$\mathbb{E}(u_{jit}^{\nu} u_{j'it}^{2-\nu}) = \begin{cases} \sigma_{u_j(t)}^2 & \text{if } \nu = 2 \\ \sigma_{u_1 u_2(t)} & \text{if } \nu = 1 \end{cases}$$

and the third moments ($n = 3$) by

$$\mathbb{E}(v_{jit}^{\nu} v_{j'it}^{3-\nu}) = \begin{cases} \gamma_{v_j(t)} & \text{if } \nu = 3 \\ \gamma_{v_1 v_2^2(t)} & \text{if } (\nu, j) = (1, 1) \text{ or } (2, 2) \\ \gamma_{v_1^2 v_2(t)} & \text{if } (\nu, j) = (2, 1) \text{ or } (1, 2) \end{cases}$$

$$\mathbb{E}(u_{jit}^{\nu} u_{j'it}^{3-\nu}) = \begin{cases} \gamma_{u_j(t)} & \text{if } \nu = 3 \\ \gamma_{u_1 u_2^2(t)} & \text{if } (\nu, j) = (1, 1) \text{ or } (2, 2) \\ \gamma_{u_1^2 u_2(t)} & \text{if } (\nu, j) = (2, 1) \text{ or } (1, 2). \end{cases}$$

$\mathbb{E}(\cdot)$ denotes the cross-sectional mean over i. I assume the spouses hold no advance information about future shocks; this is testable and typically not rejected (Meghir and Pistaferri, 2011). I allow for non-zero *cross*-moments to reflect possible assortative matching between spouses. The indexing of moments by t indicates that moments may vary with lifecycle time.

Dynamics of consumption and hours. I derive analytical expressions for the growth rates of consumption and labor supply in terms of (the growth in) spousal wages and the marginal utility of wealth. A Taylor approximation to the intra-temporal first-order conditions of household problem (1) *s.t.* (2) yields

$$\begin{aligned} \Delta c_{it} &\approx \eta_{c,w_1(i,t-1)} \Delta w_{1it} + \eta_{c,w_2(i,t-1)} \Delta w_{2it} \\ &\quad + \left(\eta_{c,p(i,t-1)} + \eta_{c,w_1(i,t-1)} + \eta_{c,w_2(i,t-1)} \right) \Delta \ln \lambda_{it} \\ \Delta h_{jit} &\approx \eta_{h_j,w_1(i,t-1)} \Delta w_{1it} + \eta_{h_j,w_2(i,t-1)} \Delta w_{2it} \\ &\quad + \left(\eta_{h_j,p(i,t-1)} + \eta_{h_j,w_1(i,t-1)} + \eta_{h_j,w_2(i,t-1)} \right) \Delta \ln \lambda_{it}, \qquad j = \{1, 2\}, \end{aligned} \tag{4}$$

Table 1 – Frisch Elasticities in Household i at Time t

Consumption	$\eta_{c,w_1(i,t)}$	:	with respect to male wage W_1
	$\eta_{c,w_2(i,t)}$	:	with respect to female wage W_2
	$\eta_{c,p(i,t)}$	:	with respect to the price of consumption P
Male hours	$\eta_{h_1,w_1(i,t)}$	:	with respect to male wage W_1
	$\eta_{h_1,w_2(i,t)}$	:	with respect to female wage W_2
	$\eta_{h_1,p(i,t)}$	:	with respect to the price of consumption P
Female hours	$\eta_{h_2,w_1(i,t)}$	:	with respect to male wage W_1
	$\eta_{h_2,w_2(i,t)}$	:	with respect to female wage W_2
	$\eta_{h_2,p(i,t)}$	:	with respect to the price of consumption P

Notes: The table presents the Frisch elasticities in household i at time t. These elasticities constitute an ordinal representation of within-period preferences. There are 3 own-price and 6 cross-price elasticities. The rule governing the notation is: $\eta_{o,x}$ is the elasticity of outcome variable $o = \{c, h_1, h_2\}$ with respect to price $x = \{w_1, w_2, p\}$. Each elasticity is subscripted by i to denote it is household-specific, and by t to reflect dependence on time-t levels of consumption and hours (see appendix B).

with details reported in appendix A. The notation is as follows: $\Delta c_{it} = \Delta \ln C_{it}$ and $\Delta h_{jit} = \Delta \ln H_{jit}$, both net of the effect of taste observables $\mathbf{Z}_{it}$ and wage observables $\mathbf{X}_{jit}$. λ_{it} is the marginal utility of wealth, namely the Lagrange multiplier on the sequential budget constraint. Parameters η are Frisch (λ-constant) elasticities defined at the household level. As an illustration, $\eta_{c,w_1(i,t)} = \frac{\partial C}{\partial W_1}\frac{W_1}{C}\Big|^{i,t}_{\lambda\text{-const.}}$ is the elasticity of consumption with respect to male wage W_1, $\eta_{c,p(i,t)}$ is the elasticity of consumption with respect to its price P, and $\eta_{h_j,w_2(i,t)}$ is the labor supply elasticity of spouse j with respect to female wage W_2. The full list of elasticities appears in table 1 and is defined formally in appendix B.

The Frisch elasticities, 9 in total per household, provide an ordinal representation of *within*-period preferences over consumption and labor supply. They are i-specific because the utility function is household-specific. The elasticities typically depend on the *levels* of consumption and hours (e.g. Attanasio et al., 2018, and appendix B), so they are also t-specific to reflect dependence on time t levels. Note, however, that the nature of the approximation is such that the Frisch elasticities in (4) depend on consumption and hours at $t-1$, i.e. on *past* levels.[5]

There is a multivariate distribution of elasticities across households, denoted by $F_{\boldsymbol{\eta}_t}$. This reflects the distribution of within-period preferences *and* the distribution of consumption and hours levels on which the elasticities depend. The distribution *conditional* on those levels is the epicenter of unobserved preference heterogeneity in this paper.

Expression (4) is empirically unattractive because the marginal utility of wealth is unobserved. Following Blundell and Preston (1998) and BPS, I overcome this in two steps. First, I

[5]The approximation involves the expansion of the problem's optimality conditions at t around consumption, hours and wages at $t-1$. This recasts the optimality conditions in terms of consumption and hours *growth* from $t-1$ to t, expressed as functions of Frisch elasticities that depend on consumption and hours at $t-1$.

approximate the Euler equation to decompose $\Delta \ln \lambda_{it}$ into two terms: the anticipated gradient of outcome growth (a function of the interest and household-specific discount rates) and an innovation that captures idiosyncratic revisions to λ due to wage shocks. Second, I approximate the lifetime budget constraint to map the innovation to λ into wage shocks. The details of both steps appear in appendix A (see also Campbell, 1993; Blundell et al., 2013).

These approximations combined lead to analytical expressions for the growth rates of outcomes as functions of wage shocks, Frisch elasticities, and two parameters pertaining to financial and human wealth (defined below). The analytical expressions are given by

$$\begin{aligned}\Delta c_{it} \approx\ & \eta_{c,w_1(i,t-1)}\Delta u_{1it} + \eta_{c,w_2(i,t-1)}\Delta u_{2it} \\ & + \left(\eta_{c,w_1(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{1it} \\ & + \left(\eta_{c,w_2(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{2it}\end{aligned} \tag{5}$$

$$\begin{aligned}\Delta h_{1it} \approx\ & \eta_{h_1,w_1(i,t-1)}\Delta u_{1it} + \eta_{h_1,w_2(i,t-1)}\Delta u_{2it} \\ & + \left(\eta_{h_1,w_1(i,t-1)} + \bar{\eta}_{h_1(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{1it} \\ & + \left(\eta_{h_1,w_2(i,t-1)} + \bar{\eta}_{h_1(i,t-1)}\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{2it}\end{aligned} \tag{6}$$

$$\begin{aligned}\Delta h_{2it} \approx\ & \eta_{h_2,w_1(i,t-1)}\Delta u_{1it} + \eta_{h_2,w_2(i,t-1)}\Delta u_{2it} \\ & + \left(\eta_{h_2,w_1(i,t-1)} + \bar{\eta}_{h_2(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{1it} \\ & + \left(\eta_{h_2,w_2(i,t-1)} + \bar{\eta}_{h_2(i,t-1)}\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{2it},\end{aligned} \tag{7}$$

where $\bar{\eta}_c = \eta_{c,p} + \eta_{c,w_1} + \eta_{c,w_2}$, $\bar{\eta}_{h_j} = \eta_{h_j,p} + \eta_{h_j,w_1} + \eta_{h_j,w_2}$, and $\boldsymbol{\eta}$ is the vector of all elasticities. Given the focus on the transmission of wage shocks, consumption and hours growth in (5)-(7) are net of taste and wage observables, and net of household-specific intercepts that determine the gradient of consumption and hours growth in the absence of shocks.[6]

Unlike permanent shocks, transitory shocks have a negligible effect on the lifetime budget constraint as long as the time horizon of the household is sufficiently long and period earnings are a small fraction of lifetime earnings.[7] Consumption responds to such shocks because of the non-separability with labor supply in the utility function. The response reflects the intertemporal substitution between consumption and leisure and measures the consumption-wage

[6]As per appendix A and for $j, j' = \{1, 2\}$, the left hand side of equations (5)-(7) is given by

$$\Delta c_{it} = \Delta \ln C_{it} - \eta_{c,p(i,t-1)}\Delta(\mathbf{Z}'_{it}\boldsymbol{\alpha}_C) - \textstyle\sum_j \eta_{c,w_j(i,t-1)}\left(\Delta(\mathbf{Z}'_{it}\boldsymbol{\alpha}_{H_j}) + \Delta(\mathbf{X}'_{jit}\boldsymbol{\alpha}_{W_j})\right) - \bar{\eta}_{c(i,t-1)}\omega_{it}$$

$$\Delta h_{jit} = \Delta \ln H_{jit} - \eta_{h_j,p(i,t-1)}\Delta(\mathbf{Z}'_{it}\boldsymbol{\alpha}_C) - \textstyle\sum_{j'} \eta_{h_j,w_{j'}(i,t-1)}\left(\Delta(\mathbf{Z}'_{it}\boldsymbol{\alpha}_{H_{j'}}) + \Delta(\mathbf{X}'_{j'it}\boldsymbol{\alpha}_{W_{j'}})\right) - \bar{\eta}_{h_j(i,t-1)}\omega_{it}.$$

The $\boldsymbol{\alpha}$'s reflect the effects of taste observables $\mathbf{Z}_{it}$ on outcomes and wage observables $\mathbf{X}_{jit}$ on wages. $\bar{\eta}_{c(i,t-1)}\omega_{it}$ and $\bar{\eta}_{h_j(i,t-1)}\omega_{it}$ are the intercepts reflecting the anticipated heterogeneous gradients of consumption and hours growth in the absence of shocks. The gradients are functions of the heterogeneous discount factor (appendix A) and the past consumption and hours levels that $\bar{\eta}_{c(i,t-1)}$ and $\bar{\eta}_{h_j(i,t-1)}$ depend on. One can obtain Δc_{it} or Δh_{jit} as the residual from regressions of $\Delta \ln C_{it}$ and $\Delta \ln H_{jit}$ on observables and past consumption and hours.

[7]This is a result rather than an assumption; appendix A illustrates this point analytically. The statement is true if 1.) the time horizon of the household is sufficiently long; and 2.) period earnings are a small fraction of remaining lifetime earnings. The second condition is true if a given transitory shock is not very large.

elasticity η_{c,w_j}. Heterogeneity in such elasticity gives rise to heterogeneity in the consumption response to transitory shocks. Similarly, hours respond to own transitory shocks reflecting the intertemporal substitution between work and leisure induced by a temporary wage shift. By definition, such response measures the own-wage Frisch labor supply elasticity η_{h_j,w_j}. Hours also respond to the partner's transitory shock; this measures the cross-wage elasticity $\eta_{h_j,w_{-j}}$ ($-j$ denotes the partner) and reflects the complementarity between spouses' leisure.[8]

Permanent shocks induce similar *substitution* effects; in addition, they induce *income* effects as they shift the lifetime budget constraint. While the response of consumption and hours to transitory shocks measures *Frisch* elasticities, the response to permanent shocks measures *lifecycle Marshallian* elasticities (Attanasio et al., 2018). ε_j in (5)-(7) reflects the impact of j's permanent shock on the budget constraint, presented analytically in appendix A. It is a function of household-specific preferences $\boldsymbol{\eta}_{it-1}$ and two financial and human wealth parameters, or wealth shares, π_{it} and $\mathbf{s}_{it} = (s_{1it}, s_{2it})'$. $\pi_{it} \approx \text{Assets}_{it}/(\text{Assets}_{it} + \text{Human Wealth}_{it})$ reflects financial wealth; a larger π_{it} implies a smaller impact of permanent shocks on consumption as the household holds more assets to absorb such shocks. $s_{jit} \approx \text{Human Wealth}_{jit}/\text{Human Wealth}_{it}$ is the share of j's human wealth (expected lifetime earnings) in total human wealth. A high s_{jit} implies that j's permanent shocks are relatively more important because this spouse contributes a larger share into family earnings. BPS and WK offer a comprehensive discussion of the transmission of shocks in the *representative* household and I refer to them for details.

Expressions (5)-(7) offer a theoretical interpretation to the local dynamics of consumption and hours and form the basis of identification of preferences. These equations are household-specific not only due to idiosyncratic shocks but also because of household-specific preferences. Accounting for preference heterogeneity can have important implications for how wage inequality translates into consumption inequality, and for the overall transmission of wage shocks into consumption and earnings. For example, consumption may *on average* be fully insured against some shock while at the same time, due to preference heterogeneity, households may respond very heterogeneously to the same shock.

3 Identification

There are two items of identification interest: the moments of wage shocks and the moments of preferences. Identification of the moments of wage shocks is straightforward; it follows Meghir and Pistaferri (2004) and earlier studies so I relegate these details (including a discussion of measurement error in wages) to appendix C.1. The remaining of this section thus focuses on identification of preferences. Section 3.1 outlines the main proposition; section 3.2 provides details on identification of most Frisch elasticities; section 3.3 explains why moments of a particular elasticity, that of consumption with respect to its price, are not identified.

[8] Given the permanent-transitory process for wages, it is the *change* in transitory shock at t that matters for outcomes at t. The loading factor of Δu changes with time as is clear from iterating equations (5)-(7).

3.1 Outline

The preference parameters of interest are the moments of the 9-dimensional joint distribution of Frisch elasticities *conditional* on the levels of consumption and hours; I denote the conditional distribution by $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ where $\mathbf{O}_t = (C_t, H_{1t}, H_{2t})'$ is the conditioning variable. In other words, $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ is the distribution of Frisch elasticities holding $\mathbf{O}_t$ fixed across households. Its moments characterize the unobserved heterogeneity in Frisch elasticities, i.e. heterogeneity not driven by differences in consumption and hours across households. For a *given* value of $\mathbf{O}_t$, there are 9 parameters for the conditional first moment: $\mathbb{E}(\eta_{c,w_j(i,t)}|\mathbf{O}_t)$, $\mathbb{E}(\eta_{c,p(i,t)}|\mathbf{O}_t)$, $\mathbb{E}(\eta_{h_{j'},w_j(i,t)}|\mathbf{O}_t)$, and $\mathbb{E}(\eta_{h_{j'},p(i,t)}|\mathbf{O}_t)$ for $j, j' = \{1, 2\}$. There are 45 parameters for the conditional second moment: the cross-sectional variance of each elasticity of table 1 (9 parameters) and all possible covariances between them (36 parameters). Appendix table C.1 lists these parameters. In general, there are $\prod_{i=1}^{8}(n+i)/8!$ parameters for the $n^{\text{th}} = \{1, 2, 3, \dots\}$ moment conditional on a value for $\mathbf{O}_t$, assuming such moment exists and is finite.

The focus of the paper is on moments of the Frisch elasticities rather than also of the geometric discount factor. The latter determines consumption and hours growth in the absence of shocks (the intercepts in (5)-(7) absorbed in the first-stage regression on observables) and the present framework, revolving around the short-term local response of outcomes to shocks, is not suited to identify moments of the discount factor (see appendix A for further discussion).

Assumption 1a. *Exogeneity of wage shocks.* *Conditional on observables, wage shocks do not depend on household consumption and labor supply preferences.*

Assumption 1a enables using wages as the source of exogenous variation that will help identify preferences. This is a standard assumption in BPS, WK, and other related studies.[9] By contrast, preferences (the Frisch elasticities) *do* depend on wage shocks. This is because the elasticities generally -but not always- depend on the levels of consumption and hours which, given the model, depend on wage shocks. For example, the consumption-wage elasticity $\eta_{c,w_j(i,t)}$ directly depends on C_{it}, H_{1it} and H_{2it} through the partial derivatives of the utility function (appendix B). Expressions (5)-(7) show that C_{it}, H_{1it} and H_{2it} depend on shocks at t and the sequence of all past shocks through their recursive structure. This complicates identification because the distribution $F_{\boldsymbol{\eta}_t}$ depends on all shocks up to t through its dependence on consumption and hours at t. However, if shocks enter the elasticities through the levels of consumption and hours *only*, then the dependence of $F_{\boldsymbol{\eta}_t}$ on shocks ceases by keeping such levels fixed. This is another reason why I focus on moments of $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ rather than of the unconditional $F_{\boldsymbol{\eta}_t}$.

[9] Exogeneity of prices is a standard assumption in the related consumer demand literature (e.g. Lewbel, 2001) and in incomplete-markets models where productivity shocks do not depend on preferences (e.g. Heathcote et al., 2014). A caveat is that in a dynamic context past choices that reflect preferences may affect current wages. A typical example is past labor supply affecting current wages. Here I assume that any possible effect of past choices on wages materializes only through observables, which is straightforward to account for empirically.

Assumption 1b. ***Conditional independence of elasticities and wage shocks.*** *Preferences $\boldsymbol{\eta}_{it}$ are independent of time-t wage shocks, namely $\boldsymbol{\eta}_{it} \perp v_{jit}, u_{jit}$ for all j, i, t, conditional on observables and the levels of consumption C_{it} and hours H_{1it}, H_{2it} that $\boldsymbol{\eta}_{it}$ depends on.*[10]

Assumption 1b restricts wage shocks to enter the Frisch elasticities only through the levels of consumption and hours that the elasticities depend on. This is generally true unless shocks enter U_{it} directly. This occurs, for example, when the true household structure is collective and shocks shift spousal bargaining power; in such case identification requires additional assumptions (assuming transitory shocks do *not* shift bargaining power would suffice).

To see how assumptions 1a and 1b help with identification, consider expression (5) for consumption growth and focus on the part driven by the male transitory shock only, namely $\Delta c_{it} \approx \eta_{c,w_1(i,t-1)} \Delta u_{1it} + \dots$. While u_{1it} enters the expression only through Δu_{1it}, u_{1it-1} enters in two places: through Δu_{1it} and the time $t-1$ levels of consumption and hours that $\eta_{c,w_1(i,t-1)}$ depends on. Assumption 1b guarantees that, after conditioning on consumption and hours at $t-1$, the elasticity is independent of u_{1it-1}. Assumption 1a guarantees that $\eta_{c,w_1(i,t-1)}$ and u_{1it} are also independent because the elasticity precedes the shock and shocks (serially uncorrelated by assumption) are exogenous to preferences. One can then use standard covariance restrictions to separate moments of the elasticity from moments of the shock.

More formally, I group the moments of $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ into three categories. The first involves moments of the 6 *wage elasticities* of consumption and labor supply ($\eta_{c,w_j(i,t)}$, $\eta_{h_1,w_j(i,t)}$, $\eta_{h_2,w_j(i,t)}$). The second involves moments of the 2 labor supply elasticities *with respect to the price of consumption* ($\eta_{h_1,p(i,t)}$, $\eta_{h_2,p(i,t)}$), including cross-moments with wage elasticities. The third includes all other parameters, namely any moment that involves the consumption elasticity *with respect to its price* $\eta_{c,p(i,t)}$. The main identification proposition reads as follows.

Proposition 1. ***Identification of elasticities.*** *Let assumptions 1a and 1b hold. If (1.) measurement error in wages, hours, and consumption is classical and Gaussian and (2.) the second moments of wage and hours measurement error are known from validation studies, then:*

- *All moments of the conditional joint distribution of wage elasticities of consumption and labor supply are identified from panel data on wages, hours, consumption and wealth;*
- *All moments of the conditional joint distribution of labor supply elasticities with respect to the price of consumption are mechanically also identified;*
- *No moment of the elasticity of consumption with respect to its price is identified without collapsing its distribution to degenerate;*
- *The variance of consumption measurement error is identified only under additional restrictions on the second moments of wage elasticities of consumption.*

[10] Wealth shares π_{it} and s_{jit} also are independent of wage shocks at t. This is a result rather than an assumption as both π_{it} and s_{jit} pertain to $t-1$ expectations, therefore are non-random at t (see appendix A).

Identification rests on the idea that cross-sectional variation in consumption or hours growth that occurs at *fixed* prices, covariates, and consumption/hours levels reflects heterogeneity in preferences. This is motivated by Abowd and Card (1989) who find that "most of the covariation of earnings and working hours occurs at fixed wage rates". It suggests that the variation in outcomes that remains after removing variation in wages/prices and observables masks heterogeneity in household preferences.

The following sections provide details on this proposition while appendix C.2 shows all identification results formally – the reader interested in the empirical application may thus jump to section 4. In what follows, and for the sake of simplicity, I assume that conditional elasticities in immediately adjacent ages are approximately equal, i.e. $\eta_{c,w_j(i,t-1)}|\mathbf{O}_{t-1}, \mathbf{O}_t \approx \eta_{c,w_j(i,t)}|\mathbf{O}_{t-1}, \mathbf{O}_t$. This holds exactly if within-period utility depends on age through observable taste shifter $\mathbf{Z}_{it}$ only. This restriction is *not* needed for identification but simplifies all subsequent statements considerably – see e.g. the derivation of the consumption autocovariance in appendix C.2.

3.2 Identified Elasticities and Consumption Measurement Error

First moments of wage elasticities. Identification follows BPS and relies on the transmission of transitory shocks into consumption and hours. This transmission is captured by moments $\mathbb{E}(\Delta w_{jit} \Delta c_{it+1}|\mathbf{O}_t)$ and $\mathbb{E}(\Delta w_{j'it} \Delta h_{jit+1}|\mathbf{O}_t)$ respectively, $j, j' = \{1, 2\}$. I condition on the levels of consumption & hours at t because Δc_{it+1} and Δh_{jit+1} are, by the nature of the approximation in (5)-(7), functions of $\boldsymbol{\eta}_{it}$. The relevant elasticities thus depend on outcome levels at t so $\mathbf{O}_t$ is the appropriate variable to condition these moments on.

These moments reflect the variance of the mean-reverting transitory shock weighed by the *average* loading factor of such shock onto consumption and hours, that is, by the average consumption-wage and labor supply elasticity respectively. The rationale is as follows: $\mathbb{E}(\eta_{h_j,w_j(i,t)}|\mathbf{O}_t)$ reflects the average sensitivity of hours h_j to a lifetime-income-constant (i.e. transitory) wage change at a given level of outcomes $\mathbf{O}_t$; this is precisely what the latter moment captures. As transitory shocks temporarily shift labor supply, the average response of consumption to such shocks (the former moment) reflects the average complementarity of consumption and hours $\mathbb{E}(\eta_{c,w_j(i,t)}|\mathbf{O}_t)$. Similarly, the average response of hours to the partner's transitory shock uncovers the average complementarity between spouses' hours $\mathbb{E}(\eta_{h_j,w_{-j}(i,t)}|\mathbf{O}_t)$.

Second moments of wage elasticities. Consider expressions (5) and (6) for consumption and male hours; assume for now that female transitory shocks are zero ($u_{2it} = 0$) and there is no measurement error. Variation in consumption growth across consecutive periods, given by $\mathbb{E}(\Delta c_{it} \Delta c_{it+1}|\mathbf{O}_{t-1}, \mathbf{O}_t) = -\mathbb{E}(\eta^2_{c,w_1(i,t)}|\mathbf{O}_{t-1}, \mathbf{O}_t)\sigma^2_{u_1(t)}$, is due to the variance of the mean-reverting transitory shock *and* heterogeneity in the consumption response to such shock, that is, in the consumption elasticity η_{c,w_1}.[11] Similarly, intertemporal variation in hours growth, given

[11] I condition here on both $\mathbf{O}_{t-1}$ and $\mathbf{O}_t$ because, by the nature of the approximation, Δc_{it} is a function of

by $\mathbb{E}(\Delta h_{1it}\Delta h_{1it+1}|\mathbf{O}_{t-1},\mathbf{O}_t) = -\mathbb{E}(\eta^2_{h_1,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_1(t)}$, is due to the variance of the shock *and* heterogeneity in the male labor supply elasticity. Covariation in consumption and hours growth across periods, given by $\mathbb{E}(\Delta c_{it}\Delta h_{1it+1}|\mathbf{O}_{t-1},\mathbf{O}_t) = -\mathbb{E}(\eta_{c,w_1(i,t)}\eta_{h_1,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_1(t)}$, is due to the variance of the shock *and* the *joint* variation in the consumption and hours elasticities. Scaling these moments by the inverse variance of the shock removes variation in wages and identifies the conditional second moments of the elasticities.

The previous lines convey the intuition behind identification in the simplest terms. Reinstating the female shock renders identification slightly more demanding but the basic logic remains unchanged. One must account for the covariance between elasticities and the joint variation in spouses' shocks. Appendix C.2 shows that the first-order autocovariance becomes

$$\begin{aligned}\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t) = &-\mathbb{E}(\eta^2_{c,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_1(t)} - \mathbb{E}(\eta^2_{c,w_2(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_2(t)} \\ &- 2\mathbb{E}(\eta_{c,w_1(i,t)}\eta_{c,w_2(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma_{u_1u_2(t)}.\end{aligned}$$

The autocovariance is due to variation in spouses' transitory shocks as well as marginal and joint heterogeneity in all consumption-wage elasticities. It no longer identifies $\mathbb{E}(\eta^2_{c,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)$ so we are in need of additional identifying equations. Consumption preference heterogeneity affects a number of higher moments of consumption and wages such as the intertemporal covariance between wage and squared consumption growth $\mathbb{E}(\Delta w_{jit}(\Delta c_{it+1})^2|\mathbf{O}_t)$, a form of impulse response function. The extent to which squared consumption growth varies with past wage growth reflects *skewness* in the distribution of shocks (the γ parameters) *and* dispersion in consumption preferences, thus provides additional information to (over-)identify such dispersion. Similar arguments apply to the identification of all second moments of wage elasticities. Appendix C.2 provides step-by-step statements for the identification of all second moments and an extension to higher moments.

Consumption measurement error. If consumption is contaminated with error, the variance of the error $\sigma^2_{e^c}$ enters $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t)$ additively and it is not separately identified from heterogeneity. An additional restriction, e.g. fixing $\mathrm{Cov}(\eta_{c,w_1(i,t)},\eta_{c,w_2(i,t)})$, enables identification of both preference dispersion and $\sigma^2_{e^c}$. Alternatively, one may fix $\sigma^2_{e^c}$ at a constant proportion of the variance of consumption growth.

Labor supply elasticities with respect to $\boldsymbol{P_t}$. Symmetry of the matrix of Frisch substitution effects, a natural theoretical restriction discussed in appendix B, identifies the moments of the hours elasticities with respect to the price of consumption $\eta_{h_j,p}$. These are simply linear transformations of the moments of the wage elasticities of consumption η_{c,w_j} (Phlips, 1974).

$\boldsymbol{\eta}_{it-1}$ while Δc_{it+1} is a function of $\boldsymbol{\eta}_{it}$. Then $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t)$ identifies $\mathbb{E}(\eta_{c,w_1(i,t-1)}\eta_{c,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)$ which, given that elasticities in adjacent ages are approximately equal, is equivalent to $\mathbb{E}(\eta^2_{c,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)$.

3.3 Elasticity of Consumption with respect to its Price

The consumption substitution elasticity $\eta_{c,p}$ matters for the sensitivity of the lifetime budget constraint to permanent shocks. In the absence of observed cross-sectional variation in the price of consumption, identification of moments of $\eta_{c,p}$ must come from the transmission of permanent shocks into consumption and hours. In practice, however, such transmission can identify at most a homogeneous $\eta_{c,p}$, and this is not without strong additional restrictions.

Assume for simplicity that utility is separable, $\pi_{it} = 0$ (no assets) and $s_{1it} = s_{2it} = 1/2$ (the spouses have equal expected lifetime earnings). The transmission of the male permanent shock into consumption is given by the lifecycle Marshallian elasticity

$$\kappa_{c,v_1(i,t)} \equiv \kappa_{c,v_1}(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}) = \frac{\eta_{c,p(i,t-1)}(1 + \eta_{h_1,w_1(i,t-1)})}{2\eta_{c,p(i,t-1)} - \eta_{h_1,w_1(i,t-1)} - \eta_{h_2,w_2(i,t-1)}}.$$[12]

Concurrent consumption and wage data easily identify moments of κ_{c,v_1}. Abstracting from preference heterogeneity, BPS recover a homogeneous $\eta_{c,p}$: conditional on the labor supply elasticities, $\mathbb{E}(\kappa_{c,v_1(i,t)})$ reflects the willingness of households to trade consumption intertemporally, thus pins down $\eta_{c,p}$. In the presence of preference heterogeneity, however, $\mathbb{E}(\kappa_{c,v_1(i,t)})$ depends nonlinearly and implicitly on the mean *and* variance of $\eta_{c,p}$, its covariance with the labor supply elasticities, higher moments, and moments pertaining exclusively to η_{h_j,w_j}. Identification even of the average $\eta_{c,p}$ fails as one cannot separate the mean of $\eta_{c,p}$ from its variance or any of its covariances. No other Marshallian elasticity (e.g. of hours) or other moments can help overcome this as the number of involved parameters is large, especially if preferences are non-separable. When the marginal distribution of $\eta_{c,p}$ is degenerate, however, $\mathbb{E}(\kappa_{c,v_1(i,t)})$ identifies a homogeneous $\eta_{c,p}$. $\mathbb{E}(\kappa_{c,v_1(i,t)})$ still depends implicitly on the moments of wage elasticities, thus necessitating to parametrize their joint distribution or approximate $\kappa_{c,v_1(i,t)}$ by a Taylor series whose order would practically need to be low. Appendix C.2 expands this point further and illustrates why $\mathbb{E}(\kappa_{c,v_1(i,t)})$ depends on various moments of $\eta_{c,p}$.

4 Empirical Application

4.1 Data

I use data from seven waves of the PSID (1999-2011; biennial data).[13] The PSID started in 1968 tracking a -then- nationally representative sample of households ('SRC' sample) as well as a second sample of lower-income households ('SEO' sample). Repeated annually until 1997 the survey collected detailed information on incomes, employment, food expenditure and demographics of the adult household members and their linear descendants should they split

[12] One can obtain this from (5) for $\eta_{c,w_j(i,t-1)} = \eta_{h_j,p(i,t-1)} = \eta_{h_j,w_{j'}(i,t-1)} = 0$, $\pi_{it} = 0$, and $s_{jit} = 0.5$.

[13] I also use data from 1997 for the estimation of wage shocks. More information on the PSID, as well as access to the data, is available online at psidonline.isr.umich.edu.

off and establish their own households. The survey became biennial after 1997; thereafter it started collecting detailed information on household consumption and wealth.

Sample selection and variable definitions are close to BPS. I select married opposite-sex couples between 30 and 60 years old from the 'SRC' sample with consistent information on demographics and employment. I focus on stable couples; if a spouse remarries, I drop the household in the year when remarriage occurs and reinstate it subsequently as a new household. The estimating equations are in first differences so I drop households that appear only once.

Identification requires wages for both spouses; therefore I select spouses who participate in the labor market and earn positive amounts.[14] I discuss potential selectivity issues below. I construct hourly wages as earnings over total hours of work on a yearly basis and I drop observations for which wages are below half the applicable state minimum wage.

I construct consumption (expenditure) as the Hicksian aggregate of a number of elementary consumption items.[15,16] I treat missing values in elementary items as zeros and I drop a few observations for which (i) total consumption is zero, or (ii) an elementary item is censored, or (iii) the reported time period of a given expenditure is missing.

I remove observations for which wages, earnings or consumption exhibit an extreme drop (jump) in a given period followed by an extreme jump (drop) in the next one as such movements may reflect measurement error.[17] I also remove households with wealth higher than \$20M.[18]

Table 2 presents measures of volatility and skewness in the baseline sample and in four subsamples of relatively wealthy households (I postpone the discussion of the wealthy to section

[14]Earnings are defined as labor income from all jobs (including overtime, tips, bonuses, commissions etc.) plus the labor part of business income from unincorporated businesses. I exclude the labor part of farm income because this is not measured consistently over time. Participation requires strictly positive earnings *and* hours of work. The latter is defined as actual realized hours on all jobs including overtime.

[15]Elementary items are: food (at home and away from home; all for recipients and non-recipients of food stamps), vehicle expenses (car insurance, repairs, parking, gasoline), transportation costs (bus and train tickets, taxicabs, other costs), child care, school expenses for children, medical costs (out of pocket health insurance, nursing homes and hospital bills, doctor, surgery, and dentist costs, prescriptions), utilities (gas, electricity, water and sewage costs, other utilities such as telecommunications), home insurance, rent for renters and rent equivalent for homeowners or people in other housing arrangements. I impute the expenditure value of housing for homeowners as 6% of the self-reported value of their house.

[16]Several expenditure categories change materially after 2011: out of pocket health insurance jumps massively from 2013 onwards; the report period for nursing homes and hospital bills changes in 2013 making the equivalent annual series drop unexpectedly; the timing of expenses on doctors and prescriptions changes in 2015 in an unclear way; vehicle repair costs have a coding error in 2015 making the series inconsistent over time. Consequently, use of a consistent consumption measure over the entire 1999-2017 period requires removing all medical expenditures and a large part of vehicle costs from such measure. I thus use data up to 2011 for consistency between the measure of consumption here and BPS (who use data up to 2009).

[17]Given the biennial nature of the data, I construct the distribution of $(\chi_{it} - \chi_{it-2})(\chi_{it+2} - \chi_{it})$ by available year and I drop observations with values in the bottom 0.25% ($\chi_{it} = \{\ln W_{jit}, \ln Y_{jit}, \ln C_{it}\}$, $j = \{1, 2\}$).

[18]Household wealth comprises the present value of the primary dwelling (net of outstanding mortgages), other real estate, savings, IRA and annuities, the value of vehicles, farms and businesses, investments in stocks and shares, and other assets net of credit card debt, student loans, medical or legal bills and loans from relatives.

Table 2 – Volatility and Skewness

	baseline sample	$A > \overline{C}$	$A > 2\overline{C}$	$A > \overline{C}$ no debt	$A > \overline{C}$ liquid
	(1)	(W1)	(W2)	(W3)	(W4)
Volatility [standard deviation of corresponding variable]					
$\Delta \ln W_{1it}$	0.494	0.516	0.531	0.487	0.517
$\Delta \ln W_{2it}$	0.449	0.446	0.454	0.452	0.460
$\Delta \ln Y_{1it}$	0.537	0.539	0.549	0.560	0.588
$\Delta \ln Y_{2it}$	0.608	0.586	0.595	0.609	0.622
$\Delta \ln C_{it}$	0.289	0.283	0.281	0.268	0.271
Skewness [third moment of corresponding variable]					
$\Delta \ln W_{1it}$	0.161	0.188	0.141	-0.676	-0.688
$\Delta \ln W_{2it}$	-0.492	-0.472	-0.439	-0.564	-0.551
$\Delta \ln Y_{1it}$	-1.017	-1.129	-1.133	-1.520	-1.477
$\Delta \ln Y_{2it}$	-0.380	-0.223	-0.286	1.247	1.219
$\Delta \ln C_{it}$	-0.062	-0.085	-0.016	0.050	0.041
Obs. [households $\times \Delta_t$]	6071	4614	3789	1806	1374

Notes: The table reports volatility and skewness in the baseline (column 1) and in four samples of wealthy households (W1-W4). Sample sizes are smaller than in summary statistics appendix table D.1 because the variables here are in first differences.

6.3). Volatility in wage and earnings growth in the baseline sample is comparable between men and women; skewness however is not. Women's wage growth is substantially *negatively* skewed, which reflects women facing considerable risk of experiencing a large wage drop, while men's wage growth exhibits positive skewness from a small number of top observations. This does not translate to earnings: men's earnings growth is substantially negatively skewed, and more so than women's. Consumption dispersion and skewness is lower than that of wages and earnings suggesting the availability of consumption insurance in the household. Appendix table D.1 presents general descriptive statistics for the sample; the summary demographic, employment, income and consumption statistics conform to expectations.

4.2 Implementation

In the empirical application I fit second and third moments of the cross-sectional joint distribution of consumption, earnings, and wage growth in the PSID.[19] This enables me to estimate up to third moments of wage shocks and up to second moments of wage elasticities. While in principle I could also estimate higher than second moments of elasticities, the data requirements for this go beyond the relatively small sample here so I leave this for future research.

[19] I follow BPS and WK and use earnings Y_j in lieu of hours H_j as the outcome variable in the empirical exercise. Given the tautology $Y_j = W_j H_j$, this choice does not make a difference in the results.

First stage. To construct empirical variables that align with their theoretical counterparts in section 2, I carry out a number of first-stage regressions for wages, earnings, and consumption. I regress $\Delta \ln W_{jit}$ on a set of observable characteristics including year, age, education, race, and state of residence dummies, as well as year-education and year-race interactions.[20] I carry out the regression separately by spouse. The residual is $\Delta w^*_{jit} = \Delta w_{jit} + \Delta e^w_{jit}$ where w_{jit} is the unexplained part of wages and e^w_{jit} is measurement error. The statistical/theoretical counterpart of Δw_{jit} is given by (3).

I regress earnings growth $\Delta \ln Y_{jit}$, separately by spouse, on (i.) the above set of observable characteristics for oneself *and* their spouse; (ii.) on indicators for the number of children, the number of household members, employment status of each spouse, and whether the household supports or receives income from members outside the household (also including changes in these variables over time and interactions with year dummies); (iii.) the past levels of consumption C_{it-1} and hours H_{1it-1}, H_{2it-1}. The residual is $\Delta y^*_{jit} = \Delta y_{jit} + \Delta e^y_{jit}$ where y_{jit} is the unexplained part and e^y_{jit} is measurement error. The theoretical counterpart of Δy_{jit} is $\Delta w_{jit} + \Delta h_{jit}$ with Δh_{jit} given by (6)-(7). Similarly, I regress consumption growth $\Delta \ln C_{it}$ on the same variables as above. The residual is $\Delta c^*_{it} = \Delta c_{it} + \Delta e^c_{it}$ where c_{it} is the unexplained part and e^c_{it} is measurement error. The theoretical counterpart of Δc_{it} is given by (5).

I control for past consumption and hours to obtain empirical growth rates that parallel their model counterparts in (5)-(7). Theoretically, growth in consumption and earnings in the absence of shocks is driven by household-specific terms that, under preference heterogeneity, depend on the past levels of consumption and hours. I call these terms the 'intercepts' of system (5)-(7) because they do not directly depend on wage shocks. These terms are not useful for identification of the Frisch elasticities so I simply absorb them in the first stage regressions on observables.[21] However, as these terms depend on past outcome levels, the first stage regressions must control for such levels in additional to other observables. Appendix table D.2 shows the results from the first stage regressions on past consumption and hours.[22]

Overall, the first stage serves to capture *observed* heterogeneity in wages and outcomes, be it through $\mathbf{X}_{jit}$ and $\mathbf{Z}_{it}$ in model notation or through past levels C_{it-1} and H_{jit-1}. Following the model and abstracting from measurement error, any remaining variation in wages is due to wage shocks and any remaining variation in earnings or consumption is due to wage shocks, heterogeneity in Frisch elasticities, and heterogeneity in wealth shares π_{it} and $\mathbf{s}_{it}$.

Conditional moments. Estimation of Frisch elasticities under preference heterogeneity requires, as per section 3, moments of Δc_{it} and Δy_{jit} conditional on consumption and hours.

[20]Given biennial data, the theoretical notation $\Delta \ln W_{jit} = \ln W_{jit} - \ln W_{jit-1}$ points to the empirical $\Delta^2 \ln W_{jit} = \ln W_{jit} - \ln W_{jit-2}$. I maintain the first notation throughout the text and for all variables.

[21]Details on this point are in footnote 6 and in paragraph 'Estimable equations' in appendix A.

[22]BPS do not control for past levels as the gradient of consumption and earnings growth under preference homogeneity does not depend on such levels. This departure from BPS seems to matter little in practice. Tables 4-5 replicate the BPS results fairly closely even after controlling for past consumption and hours levels.

Even if the first-stage regressions remove the variation in consumption and earnings growth that is due to past outcomes, this does not address the dependence of identifying moments $\mathbb{E}(\Delta w_{jit}\Delta c_{it+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta w_{j'it}\Delta y_{jit+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1}, \mathbf{O}_t)$, etc. on $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$.

There are two main ways to obtain the empirical conditional moments. The first way involves the estimation of $\mathbb{E}(\Delta w_{jit}\Delta c_{it+1})$, $\mathbb{E}(\Delta w_{j'it}\Delta y_{jit+1})$, $\mathbb{E}(\Delta c_{it}\Delta c_{it+1})$, etc. directly in the data along the entire support of conditioning variables $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$. This method thus estimates the empirical moments nonparametrically in each realization of $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$, hoping that there are enough observations per realization to permit this estimation. The obvious downside of this method is the excessive amount of data it requires.[23]

The second way is to explicitly model the dependence of moments on outcomes $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$. This involves imposing a functional form for each of the empirical conditional moments that appear in the structural estimation, i.e. a functional form of the dependence of each moment on the conditioning variables. Consider for example

$$\begin{aligned} \Delta w_{jit}\Delta c_{it+1} &= \gamma_0 + \boldsymbol{\gamma}_1'\mathbf{O}_{it} + \epsilon_{it}^{w_j c} \\ \Delta c_{it}\Delta c_{it+1} &= \delta_0 + \boldsymbol{\delta}_1'\mathbf{O}_{it} + \boldsymbol{\delta}_2'\mathbf{O}_{it-1} + \epsilon_{it}^{cc}. \end{aligned} \tag{8}$$

For given values $\mathbf{O}_t = (C_t, H_{1t}, H_{2t})'$ and $\mathbf{O}_{t-1} = (C_{t-1}, H_{1t-1}, H_{2t-1})'$, one may obtain the conditional moments as $\mathbb{E}(\Delta w_{jit}\Delta c_{it+1}|\mathbf{O}_t) = \widehat{\gamma}_0 + \widehat{\boldsymbol{\gamma}}_1'\mathbf{O}_t$ and $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1}, \mathbf{O}_t) = \widehat{\delta}_0 + \widehat{\boldsymbol{\delta}}_1'\mathbf{O}_t + \widehat{\boldsymbol{\delta}}_2'\mathbf{O}_{t-1}$, assuming that the conditional mean of the error term is zero in both cases. This assumption is internally consistent with the theoretical model which postulates that, net of observables, these moments depend only on wage shocks and outcome levels. One may estimate (8) using OLS pooling all data together. Of course, the parametric form for (8) may matter so one should try a flexible specification that also involves higher order polynomial terms in consumption and hours. The advantage of this method is its feasibility in relatively small datasets; it also permits to address the possible dependence of moments on other variables, for example on age or calendar time (e.g. if different times may be associated with different levels of earnings dispersion, skewness etc).

Both methods deliver a mapping from conditioning variables into conditional moments, namely $\mathbf{O}_t, \mathbf{O}_{t-1} \to \mathbb{E}(\Delta w_{jit}\Delta c_{it+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta w_{j'it}\Delta y_{jit+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1}, \mathbf{O}_t)$, etc. I implement the second method in practice using a linear specification in the conditioning variables (high order terms seem to matter little). In my baseline results I report estimates of Frisch elasticities at the sample average of consumption and hours, i.e. for $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$ and $\mathbf{O}_{t-1} = \mathbb{E}(\mathbf{O}_{it-1})$. From the law of iterated expectations, the conditional moments in this case are equal to their unconditional counterparts. I also show how the Frisch elasticities change along the distribution of outcomes as well as in different ages and years.

[23] One may consider grouping observations over *similar* values for $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$. Suppose the grouping is over deciles of C_t, H_{1t}, H_{2t}, and their $t-1$ counterparts. This simple grouping results in 10^6 outcome cells; most datasets could never afford enough observations within each cell to make this work in practice.

Structural estimation. Once I obtain moments $\mathbb{E}(\Delta w_{jit}\Delta c_{it+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta w_{j'it}\Delta y_{jit+1}|\mathbf{O}_t)$, $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t)$ etc. given a value for $\mathbf{O}_t$ and $\mathbf{O}_{t-1}$, I carry out the estimation of Frisch elasticities in two steps. First, I fit second and third moments of the joint distribution of wages to estimate second and third moments of wage shocks. I then fit second and, depending on the model specification, third moments of the joint distribution of wages, earnings and consumption to estimate up to second moments of wage elasticities. I deal with multiple moments and over-identifying restrictions using GMM and, in the baseline, an identity weighting matrix. I show that the choice of weighting matrix does not matter for the results.

Wealth shares π_{it} and $\mathbf{s}_{it}$ can be inferred directly from the data in a way that I describe in appendix D. These parameters enter the earnings and consumption moments through the transmission of permanent shocks. Such transmission, however, also depends on the consumption substitution elasticity that is not identified without additional strong restrictions (section 3.3). To avoid subjecting the estimation of wage elasticities to parametric form restrictions, I fit moments of earnings and consumption that pertain to the transmission of transitory shocks only. These are joint moments across *consecutive* rather concurrent time periods. The only exception is when I replicate BPS as, under preference homogeneity, the transmission of permanent shocks does not require parametric restrictions.

Measurement error. I remove a priori the variability in wages and earnings that is due to measurement error, accounting for the fact that wage and earnings errors are correlated because wages are constructed as earnings over hours. I obtain the error variance from Bound et al. (1994) who report findings from a validation study of the PSID.[24] While identification of the variance of consumption error is possible with certain restrictions (section 3.2), these restrictions apply only to certain model specifications. To avoid inconsistency across specifications, I fix consumption error at 15% of the variance of consumption growth, which is higher than the fractions that wage (13%) and earnings (4%) error are of the corresponding variances in Bound et al. (1994). I explore alternative values for consumption error in a robustness check.

Selection into labor market. Participation in the labor market is relatively high in the (prior to participation selection) sample: around 94% of men and 81% of women work. Given the focus on stable married couples, most of female non-participation is attributed to the presence of young children (included in the first-stage observables) rather than, for example, involuntary unemployment. BPS account for endogenous participation of women following two

[24]The validation study surveys workers from a single manufacturing company in 1983 and 1987. It obtains information on earnings and hours in a way that parallels the PSID questionnaire and coding practices and compares the responses to administrative data from the firm. A caveat is that the sample of workers comes from two decades prior to the data in this paper. Whether and how the nature of measurement error changed ever since is unknown. Another caveat comes from using the same estimates to correct male and female earnings or wages, although the validation study sampled male workers only.

empirical approaches: the first corrects for participation using conditional covariance restrictions as in Low et al. (2010); the second assumes the decision to work is driven by wages and demographics. Neither approach makes a difference to their results and average preferences are indistinguishable with or without the correction. In the light of the high participation rates and given the difficulty to find a convincing exclusion restriction or to model participation explicitly (this would render the approximations infeasible), I do not apply a participation correction. Importantly, volatility of $\Delta \ln C_{it}$ among all participating and non-participating individuals is 0.293 while skewness equals -0.067, both very similar to the baseline sample in table 2.

Inference. Given the multi-stage estimation, I adopt the block bootstrap as means to conduct inference (Horowitz, 2001). I draw 1,000 random samples from the baseline sample and I repeat all stages of the estimation for each such sample. A major challenge arises because some parameters are on the boundary of the parameter space and the bootstrap is inconsistent in such cases (Andrews, 2000). This applies to all heterogeneity parameters that are subject to non-negativity inequality constraints, namely the variances $\mathrm{Var}(\eta_{c,w_j(i,t)})$, $\mathrm{Var}(\eta_{h_1,w_j(i,t)})$, $\mathrm{Var}(\eta_{h_2,w_j(i,t)})$, $j = \{1, 2\}$. It does not apply to the variances of wage shocks because these are always away from the zero lower bound.

I adopt the following rule to overcome this challenge. I report standard errors $\hat{\sigma}$ for wage or preference parameters *not* subject to non-negativity constraints. I calculate this as $\hat{\sigma} = (\hat{F}^{-1}(0.75) - \hat{F}^{-1}(0.25))/(\Phi^{-1}(0.75) - \Phi^{-1}(0.25))$; the numerator is the interquartile range of the bootstrap distribution $\hat{F}$ of the relevant parameter and Φ is the standard normal cdf.[25] For parameters subject to non-negativity constraints I report the p-value associated with the null hypothesis that the respective parameter θ is zero ($\mathcal{H}_0 : \theta = 0$) against the one-sided alternative ($\mathcal{H}_a : \theta > 0$). I obtain the p-value as one minus the share of bootstrap replications for which $n^{1/2}\hat{\theta} > n^{1/2}(\hat{\theta}^* - \hat{\theta})$, where $\hat{\theta}^*$ is a bootstrap estimate of θ and $\hat{\theta}$ is the original parameter estimate. Andrews (2000) shows that the bootstrap test of $\mathcal{H}_0$ against $\mathcal{H}_a$ has the correct asymptotic rejection rate for $p \in (0, 1/2)$.

5 Empirical Results

5.1 Wages

Table 3 presents the estimates of second and third moments of wage shocks in the baseline sample, assuming stationarity over age and calendar time (note that age and time effects

[25]For normal distributions $iqr = \sigma(\Phi^{-1}(0.75) - \Phi^{-1}(0.25))$ where iqr is the interquartile range of the distribution (the difference between the 75^{th} and 25^{th} percentiles), σ is its standard deviation, and Φ is the standard normal cdf. Calculating the standard errors in this way is equivalent to applying a normal approximation to the distribution of bootstrap replications, thus shielding standard errors from extreme bootstrap draws. Before imposing normality I verify that the unrestricted distributions are approximately normal.

Table 3 – Estimates of Wage Parameters

	men		women		household			
			Panel A: *Second moments*					
permanent	$\sigma^2_{v_1}$	0.041 (0.006)	$\sigma^2_{v_2}$	0.037 (0.004)	$\sigma_{v_1v_2}$	0.005 (0.002)	$\rho_{v_1v_2}$	0.118 (0.054)
transitory	$\sigma^2_{u_1}$	0.024 (0.006)	$\sigma^2_{u_2}$	0.012 (0.005)	$\sigma_{u_1u_2}$	0.004 (0.002)	$\rho_{u_1u_2}$	0.223 (0.161)
			Panel B: *Third moments*					
permanent	γ_{v_1}	0.022 (0.030)	γ_{v_2}	-0.015 (0.005)	$\gamma_{v_1v_2^2}$	0.001 (0.002)	$\gamma_{v_1^2v_2}$	0.005 (0.002)
transitory	γ_{u_1}	-0.018 (0.008)	γ_{u_2}	-0.008 (0.006)	$\gamma_{u_1u_2^2}$	0.000 (0.001)	$\gamma_{u_1^2u_2}$	0.000 (0.004)

Notes: The table presents GMM estimates of the parameters of the wage process, pooling all years and ages together. Block bootstrap standard errors are in parentheses based on 1,000 bootstrap replications subject to a normal approximation to the interquartile range of bootstrap replications. Observations: 7034 (households $\times\Delta_t$) including data from the 1997 wave.

are partly captured in the first-stage observables).[26] Panel A presents the second moments. The variance of permanent shocks is similar between men and women. Permanent shocks covary within the couple ($\rho_{v_1v_2} = 0.118$) suggesting that spouses face similar wage risks due to similar skills, working in similar industries, or pursuing similar occupations (e.g. as a result of assortative matching). The variance of men's transitory shocks is double that of women's indicating higher wage instability possibly due to higher job mobility (Gottschalk and Moffitt, 2009). Transitory shocks covary positively ($\rho_{u_1u_2} = 0.223$) but statistically insignificantly.

Panel B presents the third moments, which are needed in the estimation of the second moments of preferences (see section 3.2 and appendix C.2). With the exception of male permanent shocks, all other wage shocks feature left skewness, that is, they have a long left tail. The corresponding third standardized moments are $\widetilde{\gamma}_{v_1} = 2.62$ (but statistically insignificant), $\widetilde{\gamma}_{v_2} = -2.07$, $\widetilde{\gamma}_{u_1} = -5.07$ and $\widetilde{\gamma}_{u_2} = -5.92$. Such left tail suggests that negative shocks are more devastating than positive ones as they are on average further away from the zero mean. Guvenen et al. (2015) obtain similar results for earnings shocks of million of workers using data from the Social Security Administration. In addition, all cross-moments except $\gamma_{v_1^2v_2}$ are effectively zero implying there is limited co-skewness between spouses' shocks.

5.2 Preferences

Tables 4 and 5 present the cross-sectional first and second moments of Frisch elasticities. Column 1 shows the main results from BPS (table 4 column 2 therein) while column 2 replicates BPS using my data. Then columns 3 through 8 present results introducing additional moments

[26]BPS allow the second moments of shocks to vary over pre-defined age brackets. The variance of permanent shocks follows an U-shape over the lifetime but this does not affect their estimates of average preferences.

and preference heterogeneity. I obtain these results at the sample average of consumption and hours, i.e. for $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$, pooling all ages and calendar years together, weighing all empirical moments equally, and fixing the variance of consumption measurement error at 15% of the variance of consumption growth. After I discuss the baseline results, I show parameter estimates along the entire distribution of $\mathbf{O}_t$, over different ages and years, using a different weighting matrix, and for different values of consumption error.

Replication of BPS. Column 2 replicates BPS using an additional wave of data (wave 2011). I target the same second moments that BPS target (i.e. including moments of the transmission of permanent shocks) but I deviate from them by controlling for past consumption and hours in the first-stage regressions. Attanasio et al. (2018) argue that aggregate labor supply elasticities depend on hours worked, for which BPS do not account. So accounting for hours may lead to meaningful differences from BPS. Nevertheless, the results are quite similar to theirs; the main difference is that the female labor supply elasticity drops by approximately 10% to 0.78. The difference from BPS is not statistically significant.[27]

New specifications. Column 3 estimates preferences without heterogeneity from second moments of wages, earnings and consumption that pertain exclusively to the transmission of transitory shocks. This specification uses less information than column 2 as it no longer targets moments involving the transmission of permanent shocks such as the variance of consumption growth $\text{Var}(\Delta c_{it})$. As explained earlier, such moments depend on the consumption substitution elasticity $\eta_{c,p}$ which is not identified in the presence of heterogeneity. Nevertheless, the performance of the model with respect to $\text{Var}(\Delta c_{it})$ will be subsequently assessed.

Two things are worth noting. First, both consumption-wage elasticities η_{c,w_j} are now statistically not different from zero. This implies separability between consumption and labor supply at the intensive margin. However, the response to transitory shocks alone may partly reflect liquidity constraints, a point to which I return in section 6.3. Second, the own-wage labor supply elasticities drop to $\eta_{h_1,w_1} = \mathbb{E}(\eta_{h_1,w_1(i)}) = 0.41$ for men and $\eta_{h_2,w_2} = \mathbb{E}(\eta_{h_2,w_2(i)}) = 0.45$ for women while the *cross*-elasticities are statistically zero (more on the cross-elasticities below).[28]

Column 4 estimates preferences without heterogeneity from second *and third* moments of wages, earnings and consumption. This specification uses the moments of column 3 and additional information from third moments of wages and outcomes that will later contribute to the

[27] In replicating BPS, I fix $\eta_{c,p} = -0.45$, i.e. at the midrange of estimates from BPS. This is for two reasons: first, I cannot identify this elasticity in subsequent model specifications (see section 3.3); second, it has been practically hard to pin down $\eta_{c,p}$ in my data as it subjects the problem to many similarly-valued local minima. Fixing $\eta_{c,p}$ rectifies this but slightly decreases the standard errors of all other estimates compared to BPS.

[28] In section 6.4 I study the information content of moments of transitory shocks and I show that they convey sufficient information for identification. I repeat the estimation of column 3 using data up to 2009 (as in BPS) and data up to the 2017 wave of the PSID. While the elasticities drop in the former case, the labor supply elasticities with respect to *own wages* are mostly similar across columns 2 & 3 in the latter.

Table 4 – Estimates of Preferences: Means and Variances at Sample Average of Consumption and Hours

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
specification/moments:	BPS	BPS-repl.	2^{nd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}
heterogeneity:	no	no	no	no	marginal	joint	joint-restr.	preferred
change from column on the left:		2011 data, past levels	only transitory	third moments	marg. heterogeneity	joint heterogeneity	$\eta_{h_j,w_{j'}} = 0$	mild homogeneity
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.148	-0.237	0.029	-0.020	-0.020	-0.041	-0.042	-0.033
	(0.060)	(0.058)	(0.062)	(0.059)	(0.053)	(0.062)	(0.062)	(0.032)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.030	0.047	-0.107	-0.054	-0.045	-0.032	-0.031	-0.053
	(0.059)	(0.049)	(0.117)	(0.084)	(0.070)	(0.075)	(0.074)	(0.052)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.594	0.589	0.406	0.283	0.282	0.276	0.275	0.276
	(0.155)	(0.174)	(0.151)	(0.125)	(0.126)	(0.125)	(0.121)	(0.120)
$\mathbb{E}(\eta_{h_1,w_2(i)})$	0.104	0.089	-0.049	-0.010	-0.010	-0.009		
	(0.053)	(0.054)	(0.062)	(0.048)	(0.047)	(0.047)		
$\mathbb{E}(\eta_{h_2,w_1(i)})$	0.212	0.186	-0.102	-0.020	-0.020	-0.019		
	(0.108)	(0.112)	(0.128)	(0.100)	(0.098)	(0.098)		
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.871	0.776	0.446	0.256	0.255	0.254	0.247	0.247
	(0.221)	(0.153)	(0.294)	(0.220)	(0.224)	(0.227)	(0.198)	(0.196)
$\mathrm{Var}(\eta_{c,w_1(i)})^{\#}$					0.081	0.097	0.090	0.091
					[0.144]	[0.106]	[0.119]	[0.014]
$\mathrm{Var}(\eta_{c,w_2(i)})^{\#}$					0.269	0.259	0.246	0.233
					[0.137]	[0.130]	[0.120]	[0.014]
$\mathrm{Var}(\eta_{h_1,w_1(i)})^{\#}$					0.002	0.020	0.021	0.018
					[0.132]	[0.214]	[0.215]	[0.209]
$\mathrm{Var}(\eta_{h_1,w_2(i)})^{\#}$					0.000	0.000		
					[0.178]	[0.232]		
$\mathrm{Var}(\eta_{h_2,w_1(i)})^{\#}$					0.001	0.003		
					[0.195]	[0.231]		
$\mathrm{Var}(\eta_{h_2,w_2(i)})^{\#}$					0.001	0.004	0.004	0.003
					[0.055]	[0.167]	[0.174]	[0.152]

Notes: See page 27. Observations across columns 2-8: 6071 (households $\times \Delta_t$). $^{\#}p$-value in square brackets for the one-sided test that the respective parameter equals zero.

Table 5 – Estimates of Preferences *(continues from previous table)*: Covariances at Sample Average of Consumption and Hours

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$						0.001	0.128	0.146
						(0.009)	(0.128)	(0.050)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$						0.040	0.040	0.036
						(0.035)	(0.036)	(0.024)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_2(i)})$						0.000		
						(0.005)		
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_1(i)})$						-0.001		
						(0.009)		
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.002	-0.009	-0.008
						(0.011)	(0.021)	(0.017)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$						0.016	0.055	0.058
						(0.025)	(0.047)	(0.038)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_2(i)})$						0.001		
						(0.004)		
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_1(i)})$						0.002		
						(0.009)		
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$						-0.020	-0.018	-0.013
						(0.029)	(0.034)	(0.027)
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_1,w_2(i)})$						0.000		
						(0.003)		
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_1(i)})$						0.000		
						(0.006)		
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.002	-0.004	-0.003
						(0.016)	(0.024)	(0.019)
$\mathrm{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_1(i)})$					0.000	0.000		
					(0.001)	(0.001)		
$\mathrm{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_2(i)})$						0.000		
						(0.001)		
$\mathrm{Cov}(\eta_{h_2,w_1(i)}, \eta_{h_2,w_2(i)})$						0.000		
						(0.002)		

Notes: See page 27. Observations across columns 2-8: 6071 (households $\times \Delta_t$).

Notes for tables 4 and 5: The tables present GMM estimates of first and second moments of wage elasticities. Except column 1, all moments are conditional on the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. Column 1 reports the original BPS estimates from table 4, column 2 therein. Column 2 replicates BPS using additional PSID data in 2011 and controlling for past levels of consumption and hours in the first stage. Column 3 estimates preferences without heterogeneity from second moments of wages, earnings and consumption that pertain to the transmission of transitory shocks only. With unobserved preference heterogeneity subsequently, I cannot target moments of the consumption and hours response to permanent shocks unless I impose strong parametric assumptions. Column 4 estimates preferences without heterogeneity from second *and* third moments of wages and outcomes. Third moments are needed to subsequently identify preference heterogeneity. Column 5 allows for marginal heterogeneity in preferences. Column 6 reports estimates for the unrestricted multivariate preference distribution. Column 7 reports estimates for the multivariate preference distribution, shutting down the labor supply cross-elasticities. Column 8 reports the preferred specification where I impose a mild homogeneity restriction across households, namely $\eta_{c,w_2(i)} = \alpha \times \eta_{c,w_1(i)}$, $\forall i$. Standard errors appear in parentheses and, whenever applicable, p-values in square brackets for the one-sided test that the respective parameter equals zero. In the estimation of the covariances I require that the Pearson correlation coefficients of any pair of elasticities are within $[-1; 1]$ and that the matrix of second moments is positive semi-definite. Frisch symmetry implies that $\text{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_1(i)})$ is a positive transformation of $\text{Var}(\eta_{h_1,w_2(i)})$. The standard error of this covariance is consistent because the parameter is on the space boundary defined by an *equality* constraint (Andrews, 2000).

estimation of preference heterogeneity. Two things are worth noting. First, the consumption-wage elasticities are attenuated as the model attempts to match third moments of consumption and wages, particularly $\mathbb{E}\big((\Delta w_{jit})^2 \times \Delta c_{it+1}\big)$; both remain statistically insignificant. Second, men's labor supply elasticity drops by 30% to 0.28 while women's drops by 42% to 0.26. The model attenuates these elasticities because, first, left skewness in female earnings is lower than in female wages and, second, male earnings exhibit opposite skewness (negative) from male wages (see table 2). The model attempts to reproduce these patterns by reducing the average labor supply elasticities in order to strike a wedge between wages and earnings.

Column 5 estimates preferences *with* heterogeneity from second and third moments of wages, earnings and consumption. Heterogeneity here is restricted as the elasticities can vary marginally, but not jointly, across households. While the first moments remain unchanged from column 4, two new findings are worth noting.

First, the variances of consumption elasticities at $\text{Var}(\eta_{c,w_1(i)}) = 0.08$ and $\text{Var}(\eta_{c,w_2(i)}) = 0.27$ are large pointing to substantial unobserved heterogeneity in consumption preferences, unrelated to the underlying levels of consumption and hours the elasticities depend on. Although these parameters are not statistically significant at conventional levels, at face value they imply that: (i) one standard deviation of η_{c,w_1} about its cross-sectional mean falls approximately in $(-0.30; 0.27)$; (ii) one standard deviation of η_{c,w_2} about its mean falls in $(-0.56; 0.47)$. These intervals suggest that for some households consumption and hours are Frisch substitutes while for other households they are Frisch complements.

Second, the variances of labor supply elasticities are economically zero (and one of them statistically significantly so) implying that there is not much heterogeneity in *intensive margin* labor supply elasticities once wage variation, observables, and the levels of consumption and hours are accounted for. While this result appears puzzling at first, especially given the large amount of consumption preference heterogeneity, all earnings moments that over-identify these variances are matched quite well. With measurement error taken into account, the variation in female and (especially) male *hours* in the sample of stable working families is a small fraction of the variation in their respective wages and earnings. The model can match the involved earnings moments scaling up the wage moments by the appropriate average labor supply elasticities. This is not the case for consumption where average elasticities are zero and the model requires consumption preference heterogeneity in order to match the consumption second moments.

Column 6 estimates preferences with heterogeneity from second and third moments of wages and outcomes. Heterogeneity here is the most general as the elasticities can vary jointly. Three things are worth noting. First, the variances of consumption elasticities remain unchanged at $\text{Var}(\eta_{c,w_1(i)}) = 0.10$ and $\text{Var}(\eta_{c,w_2(i)}) = 0.26$ reflecting substantial consumption preference heterogeneity across households. The first variance now marginally misses significance at the 10% level. Second, the variances of labor supply elasticities remain economically zero even though $\text{Var}(\eta_{h_1,w_1(i)})$ increases tenfold from column 5. Third, although η_{h_1,w_1} varies positively with η_{c,w_1} ($corr = 0.90$) and η_{c,w_2} ($corr = 0.23$), all other covariances are statistically and economically zero. The fully flexible model, albeit appealing from a theoretical perspective, is more flexible than required to fit the joint distribution of wages and outcomes across households. Restrictions are thus needed to reduce the number of parameters and improve efficiency.

The labor supply *cross*-elasticities reflect the added worker effect, namely the response of hours to the partner's permanent shock (Lundberg, 1985). Moments pertaining to transitory shocks alone do not seem to identify this effect and these elasticities are estimated at zero in all specifications up to this point. A natural first restriction is thus to shut down these elasticities. Column 7 presents the results: the main conclusions remain unchanged from column 6 although most covariances now imply economically large correlations between pairs of elasticities.

Interim summary and preferred specification. A summary of the results so far is: (1.) the average consumption-wage elasticities are economically zero implying that preferences for the *average* household are separable between consumption and hours; (2.) the average labor supply elasticities drop as the model attempts to match third moments of earnings and wages; (3.) the consumption elasticities exhibit substantial heterogeneity implying consumption and leisure are complements *or* substitutes away from the average household; (4.) once wages, observables, and the levels of consumption and hours are accounted for, there is little remaining heterogeneity in intensive margin labor supply elasticities; (5.) preference heterogeneity helps better fit the joint distribution of wages and outcomes but the fully unrestricted model, albeit theoretically appealing, neither provides efficient estimates nor fares better over a version with restricted heterogeneity (the value of the GMM metric improves little from column 5 to 6).

Despite parameters such as $\text{Var}(\eta_{c,w_j(i)})$ being economically large, statistical significance is at best rather weak even when the labor supply cross-elasticities are shut. The estimation is underpowered due to the large number of parameters and the fact that relatively few moments contribute to the estimation of heterogeneity. To improve power, I impose a mild *homogeneity* restriction across households. I posit $\eta_{c,w_2(i)} = \alpha_i \times \eta_{c,w_1(i)}$ (which is always true for α_i unrestricted) and estimate preferences holding $\alpha_i = \alpha$ fixed across households. There are several underlying utilities that imply proportional elasticities.[29] This delivers a number of linear restrictions across moments of consumption elasticities that improve efficiency drastically without compromising fit. Neither this nor any of several alternative restrictions that I explore in appendix table E.1 meaningfully alters the parameter point estimates from column 7.[30]

Estimates from this preferred specification are in column 8. Four observations emerge but the previous conclusions remain intact. First, the average consumption-wage elasticities are economically small and statistically zero implying that preferences in the *average* household are separable between consumption and leisure. Second, the average labor supply elasticities are small; their magnitude drops as the model attempts to match third moments of earnings and wages. Third, consumption preference heterogeneity remains substantial. The variance of consumption elasticities, now significant almost at 1%, implies that one st.d. of η_{c,w_1} about its mean falls approximately in the range $(-0.33; 0.27)$ while one st.d. of η_{c,w_2} in $(-0.54; 0.43)$. Fourth, the variance of labor supply elasticities is at least an order of magnitude lower than that of consumption elasticities; these parameters remain insignificant. This result should not be surprising. Heathcote et al. (2014) find lower preference heterogeneity when they use hours but not consumption data. It is their consumption data that warrant large amounts of preference heterogeneity suggesting that it is consumption preferences that exhibit such heterogeneity. Moreover, Attanasio et al. (2018) find that most heterogeneity in labor supply elasticities is due to demographics and hours worked, both of which I control for here. Appendix tables D.4-D.6 provide numerical evidence for the fit of the preferred model.

Robustness. All results are robust to a large number of alternative restrictions (table E.1), to trimming extreme observations in wages, earnings and consumption that may matter for the third moments (table E.2), and to additional first-stage observables (tables E.3-E.4). They are also robust to an alternative weighting scheme that takes into account the precision of the empirical moments in the data (diagonally weighted GMM; tables E.5-E.7).

The variance of consumption measurement error is identified in certain specifications only after imposing additional restrictions. To avoid inconsistencies across specifications, I fixed this at 15% of the variance of consumption growth. This number is higher than error in wages

[29] Abstracting from observables, let the household utility function be of constant relative risk aversion form and separable between spouses' hours; for example $U_i(C_{it}, H_{1it}, H_{2it}) = \frac{(C_{it}H_{1it})^{1-\gamma_{1i}}}{1-\gamma_{1i}} + \frac{(C_{it}H_{2it})^{1-\gamma_{2i}}}{1-\gamma_{2i}}$. The consumption-wage elasticities are determined by γ_{1i} and γ_{2i}, and the levels of consumption and hours. A restriction $\eta_{c,w_2(i)} \approx \alpha \times \eta_{c,w_1(i)}$ across households implies a restriction between γ_{1i} and γ_{2i}, given that the estimation holds the outcome levels fixed. In the implementation of this restriction at $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$, I first estimate α freely, I find $\widehat{\alpha} = 1.59$, then I fix $\alpha = 1.6$. Results treating α as free parameter are in appendix table E.1.

[30] Alternative restrictions: no heterogeneity in labor supply elasticities; no *joint* variation in labor supply elasticities; $\eta_{c,w_2(i)} = \alpha \times \eta_{c,w_1(i)}$ and treat α as a free parameter; equal labor supply elasticities between spouses, i.e. $\eta_{h_1,w_1(i)} = \eta_{h_2,w_2(i)}$; equal consumption-wage elasticities, i.e. $\eta_{c,w_1(i)} = \eta_{c,w_2(i)}$. All results are remarkably similar to column 7 and to the preferred specification in column 8 of tables 4-5.

(13%) or earnings (4%) in Bound et al. (1994) but it is ultimately arbitrary. So it is important to check how the results may change with a different amount of consumption error. Appendix figure E.1 plots the variances of consumption and labor supply elasticities against the variance of consumption error. Consumption preference heterogeneity remains substantial for amounts of consumption error up to about 40-45% of the variance of consumption growth; heterogeneity in labor supply elasticities remains always negligible.

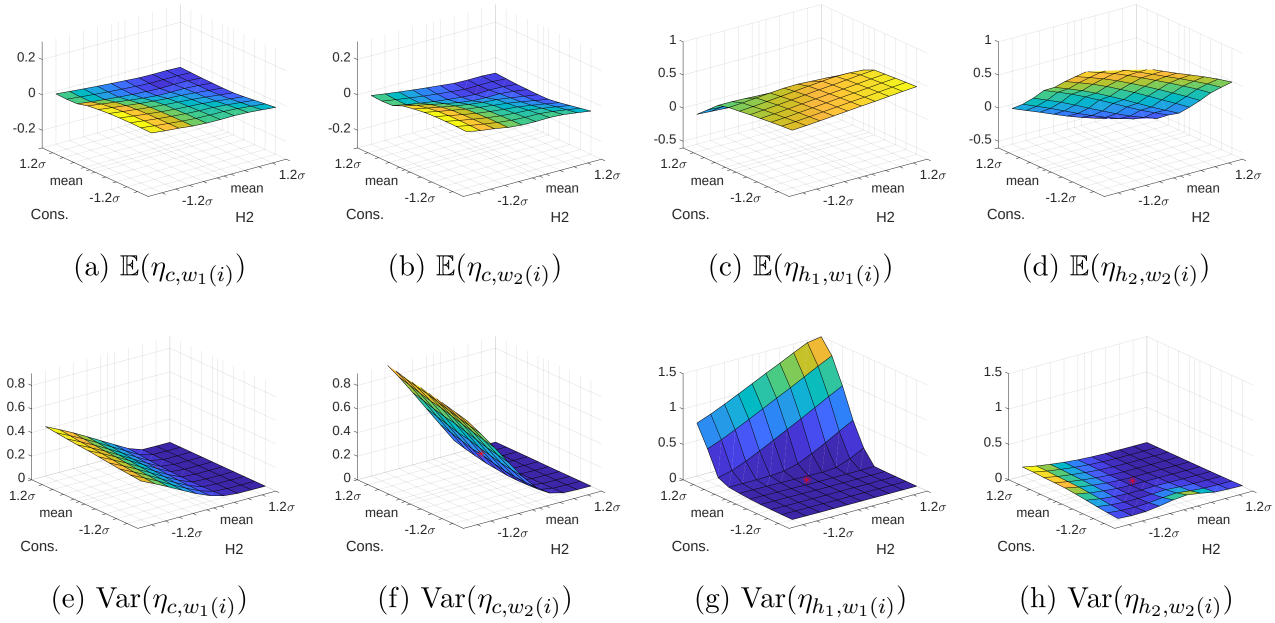
Naturally, a bigger amount of consumption error reduces the scope of preferences in the model especially as the amounts of earnings & wage error do not increase in parallel. A larger consumption error reduces the 'true' variance of consumption; the model then gradually attributes a larger share of consumption variation to wage variation, which is constant throughout the counterfactual of figure E.1. This is not true if the amounts of earnings & wage error increase proportionally to the amount of consumption error. In that case, the 'true' variance of both consumption and wages decreases and consumption preference heterogeneity retains its relative importance. For most values of consumption, wage and earnings error, heterogeneity in labor supply preferences remains negligible (see appendix E for further discussion).

Parameters over age, time, and the distribution of outcomes. The results so far are for the Frisch elasticities at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. Does consumption preference heterogeneity retain its relative importance away from the mean? Do average labor supply elasticities remain low? To answer these questions I estimate the preferred specification along the distribution of consumption and hours. I use (8) to obtain the target conditional empirical moments over a fine grid of values for consumption and hours within ± 1.5 standard deviations from their mean, and I repeat the structural estimation in each point of this grid.

Figure 1 plots the parameter estimates along the distribution of consumption and female hours, at the sample average of male hours. Average consumption elasticities are stable around zero or weakly negative while average labor supply elasticities remain low, centered around 0.25-0.35, and never exceed 0.5. They drop to zero or become negative at very high values of consumption, indicating strong income effects in hours. The variance of consumption elasticities remains substantial except at high values of female hours. Finally, heterogeneity in labor supply elasticities is negligible except at high values of consumption where $\mathrm{Var}(\eta_{h_1,w_1(i)})$ suddenly jumps because of lack of identification numerically in that sample. Results away from the mean of male hours are similar.

Subsequently, I estimate the preferred specification in different age brackets and calendar years. Behavior may vary across stages of the lifecycle, while different calendar years may exhibit different amounts of inequality or earnings/consumption skewness. I augment (8) to include age and, separately, time controls, and I use it to obtain target conditional empirical moments in various age/time brackets. I condition on average consumption and hours within each age/time bracket. The results appear in appendix figures E.3 and E.4. Average consump-

Figure 1 – Parameter Estimates along the Distribution of Consumption and Hours



(a) $\mathbb{E}(\eta_{c,w_1(i)})$ (b) $\mathbb{E}(\eta_{c,w_2(i)})$ (c) $\mathbb{E}(\eta_{h_1,w_1(i)})$ (d) $\mathbb{E}(\eta_{h_2,w_2(i)})$

(e) $\mathrm{Var}(\eta_{c,w_1(i)})$ (f) $\mathrm{Var}(\eta_{c,w_2(i)})$ (g) $\mathrm{Var}(\eta_{h_1,w_1(i)})$ (h) $\mathrm{Var}(\eta_{h_2,w_2(i)})$

Notes: This figure plots selected parameter estimates in the preferred specification, along the distribution of consumption and female hours within ± 1.5 standard deviations from their respective mean. I condition on the sample average of male hours.

tion elasticities are stable over *age* while average labor supply elasticities become negative at older ages when individuals have accumulated wealth and income effects may dominate labor supply. The pattern of heterogeneity within each age bracket is remarkably similar to the baseline: consumption preference heterogeneity is substantial, labor supply preference heterogeneity is not. First and second moments of consumption elasticities do not vary over *time*; however, $\mathbb{E}(\eta_{h_j,w_j(i)})$, $j = \{1, 2\}$, drops in 2009-2011 while $\mathrm{Var}(\eta_{h_1,w_1(i)})$ suddenly jumps. These results are fragile so they do not entertain an economic explanation. Overall, the results are mostly similar to the baseline so baseline heterogeneity is not picking up heterogeneity in age/time.

Consumption substitution elasticity. While estimation of the wage elasticities does not require parametric restrictions, estimation of $\eta_{c,p}$ is not possible without strong distributional assumptions. Such assumptions are not needed when preferences are homogeneous so BPS estimate $\eta_{c,p}$ together with the wage elasticities. To avoid subjecting *all* parameters to parametric restrictions, I estimate $\eta_{c,p}$ in a second stage after estimating the wage elasticities without restrictions. I use the variance of consumption growth at the sample average of consumption and hours in order to estimate a homogeneous $\eta_{c,p}$; this involves moments of the Marshallian elasticity $\kappa_{c,v_j(i,t)}$ that was shown in section 3.3 to depend on $\eta_{c,p}$. The Marshallian elasticity also depends on the *entire* distribution of wage elasticities thus necessitating fixing their joint distribution. A natural benchmark is to use the joint normal, parameterized at the first and second moments in the preferred specification (this does not contradict the previous use of third moments as consumption and hours can still exhibit skewness because wage shocks are skewed). I operationalize this using simulated minimum distance as $\mathrm{Var}(\Delta c_{it})$ is a nonlinear implicit function of $\eta_{c,p}$ and all other elasticities. The distance metric is minimized at $\eta_{c,p} = -0.598$ implying a coefficient of relative risk aversion close to 1.67. I report details in appendix D and the sensitivity to consumption measurement error in figure E.2.

6 Discussion

6.1 Implications for Consumption Inequality

In this section I examine how wage inequality translates into consumption inequality, focusing on the variance of consumption growth across households. This describes the evolution of consumption inequality in response to wage shocks. In other words, it characterizes how wage risk and uncertainty translate into heterogeneous consumption paths across households who start off their lives at similar consumption levels. Its main limitation is that, as in Blundell et al. (2008) and BPS, it is agnostic about inequality in consumption levels.

Using expression (5), the properties of shocks and assumptions 1a and 1b, the variance of consumption growth is given by

$$
\begin{aligned}
\operatorname{Var}(\Delta c_{it}) \approx & \sum_{j=1}^{2} \underbrace{\mathbb{E}(\eta^2_{c,w_j(i,t-1)})}_{\text{involves heterogeneity in } \eta_{c,w_j}} \times \left(\sigma^2_{u_j(t)} + \sigma^2_{u_j(t-1)}\right) \\
& + 2\underbrace{\mathbb{E}(\eta_{c,w_1(i,t-1)}\eta_{c,w_2(i,t-1)})}_{\text{involves joint heterogeneity in } \eta_{c,w_1} \text{ and } \eta_{c,w_2}} \times \left(\sigma_{u_1u_2(t)} + \sigma_{u_1u_2(t-1)}\right) \\
& + \sum_{j=1}^{2} \underbrace{\mathbb{E}\left(\left(\eta_{c,w_j(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right)^2\right)}_{\text{involves joint heterogeneity in preferences, and heterogeneity in financial \& human wealth}} \times \sigma^2_{v_j(t)} \\
& + 2\underbrace{\mathbb{E}\left(\prod_{j=1}^{2}\left(\eta_{c,w_j(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right)\right)}_{\text{involves joint heterogeneity in preferences, and heterogeneity in financial \& human wealth}} \times \sigma_{v_1v_2(t)}.
\end{aligned}
\tag{9}
$$

To obtain (9) I use results for the second moments of products of random variables from Goodman (1960) and Bohrnstedt and Goldberger (1969). Appendix F details this derivation.

Theoretical interpretations. Expression (9) shows that consumption inequality is the result of three distinct forces: (1.) wage inequality (second moments of shocks), (2.) preference heterogeneity, (3.) heterogeneity in financial and human wealth. Consumption inequality can be decomposed into two parts given the type of wage inequality behind it: *consumption instability* due to transitory shocks and *permanent inequality* due to permanent shocks.[31] Ceteris paribus, the higher the variances of wage shocks, the higher consumption inequality is.

[31] Gottschalk and Moffitt (2009) use the term 'income instability' to refer to short-term transitory fluctuations in income that contribute to overall income inequality.

Preferences contribute to consumption inequality through their first and second moments. Consider the loading factor of the variance of transitory shock $\sigma^2_{u_j(t)}$, given by $\mathbb{E}(\eta^2_{c,w_j(i,t-1)}) = \mathbb{E}(\eta_{c,w_j(i,t-1)})^2 + \mathrm{Var}(\eta_{c,w_j(i,t-1)})$. Consumption inequality increases with the absolute value of the mean consumption-wage elasticity as well as with its variance. A large average elasticity implies that the representative household is more responsive to transitory wage fluctuations while a large variance implies there are households for whom consumption responds even more substantially. Both magnify the transmission of wage into consumption inequality.

Now consider the loading factor of the variance of permanent shock $\sigma^2_{v_j(t)}$. Suppose for the sake of the illustration that $\varepsilon_j = 1$ (more on this to follow). The loading factor is

$$\begin{aligned}\mathbb{E}\left(\left(\eta_{c,w_j(i,t-1)} + \bar{\eta}_{c(i,t-1)}\right)^2\right) &= \mathbb{E}(\eta_{c,w_j(i,t-1)})^2 + \mathbb{E}(\bar{\eta}_{c(i,t-1)})^2 + 2\mathbb{E}(\eta_{c,w_j(i,t-1)})\mathbb{E}(\bar{\eta}_{c(i,t-1)}) \\ &\quad + \mathrm{Var}(\eta_{c,w_j(i,t-1)}) + \mathrm{Var}(\bar{\eta}_{c(i,t-1)}) + 2\mathrm{Cov}(\eta_{c,w_j(i,t-1)}, \bar{\eta}_{c(i,t-1)}),\end{aligned}$$

where $\mathrm{Var}(\bar{\eta}_{c(i,t-1)})$ involves marginal and joint heterogeneity in all three consumption elasticities of table 1. Like before, consumption inequality increases with the absolute value of the mean consumption elasticities (first line above) and with preference heterogeneity (last line).

These expressions show that, conditional on wage inequality and a value for ε_j, marginal preference heterogeneity *always* increases consumption inequality. Let ϕ_{it} be a generic transmission parameter of a shock into consumption (i.e. $\phi_{it} = \eta_{c,w_1(i,t-1)}$ for u_{1it} or $\phi_{it} = \eta_{c,w_1(i,t-1)} + \bar{\eta}_{c(i,t-1)}$ for v_{1it}); then $\partial\mathrm{Var}(\Delta c_{it})/\partial\mathrm{Var}(\phi_{it}) > 0$ and preference heterogeneity increases consumption inequality compared to the homogeneity benchmark ($\mathrm{Var}(\phi_{it}) = 0$). The logic is straightforward: households who differ in preferences respond differently to a given wage shock and their consumption grows apart. The greater preference heterogeneity is, the further apart their consumption responses are, and the higher consumption inequality is.

While this analytical result is *always* true for consumption instability, it is derived for permanent inequality under the ad hoc value $\varepsilon_j = 1$. ε_j is a complicated implicit function of preferences and wealth shares (section 3.3). However, simulations of consumption inequality (reported below) that fully account for the structure of ε_j confirm that preference heterogeneity indeed *always* increases permanent inequality. As I show below, the same applies to heterogeneity in financial and human wealth shares π_{it} and $\mathbf{s}_{it}$. For completeness, appendix F presents expressions for earnings and hours inequality and offers additional remarks.

Empirical results. The wage and preference parameters enable the decomposition of observed consumption inequality to *consumption instability* and *permanent inequality*. Table 6 reports the results. At the sample average of consumption and hours, and after controlling for observables, consumption instability (driven by transitory shocks) accounts for 21.5% of overall inequality while permanent inequality (driven by permanent shocks) accounts for the remaining 78.5%. The two components add up to overall consumption inequality by definition, given that the consumption substitution elasticity $\eta_{c,p}$ is chosen such that consumption inequality in the model matches its empirical counterpart.

Table 6 – Accounting Decomposition of Consumption Inequality

		% of cons. inequality	% of cons. instability	% of perm. inequality
$\mathrm{Var}(\Delta c_{it})$	0.055	100%		
standard error	(0.001)			
consumption instability	0.012	21.5%	100%	
without pref. heterogeneity	0.000		1.2%	
pref. heterogeneity induced	0.012		98.8%	
permanent inequality	0.043	78.5%		100%
without heterogeneity	0.026			59.6%
heterogen. in preferences only	0.031			72%
heterogen. in preferences, π_{it}, $\mathbf{s}_{it}$	0.043			100%

Notes: The table presents the decomposition of consumption inequality into consumption instability and permanent inequality. $\mathrm{Var}(\Delta c_{it})$ is estimated at the sample average of consumption and hours (i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$) and is net of measurement error; a block bootstrap standard error from 1,000 replications is in parentheses. Simulations of permanent inequality assume preferences are jointly normal and maintain $\eta_{c,p} = -0.598$. Wealth shares π_{it} and $\mathbf{s}_{it}$ are drawn from their empirical distributions.

Consumption preference heterogeneity is responsible for nearly all (98.8% of) consumption instability. This is because the average consumption-wage elasticities are almost zero and preferences in the representative household are separable in consumption and leisure. Consumption would hardly respond to transitory shocks if every household had the same average preferences and consumption instability would measure a tiny 1.2% of the baseline figure implied by the parameter estimates. In practice, a distribution of consumption-wage elasticities about the mean implies a distribution of consumption responses to transitory shocks inducing consumption instability. In such case, a 98.8% reduction in the second moments of transitory shocks is needed to compensate for instability induced by preference heterogeneity.

While the contribution of preference heterogeneity to consumption instability is deduced analytically, its contribution to permanent inequality is not. Term $\varepsilon_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})$ that captures the effect of permanent shocks on lifetime income is an implicit function of preferences, thus necessitating a numerical investigation of the quantitative role of heterogeneity. To address this I simulate permanent inequality across 10 million households with and without heterogeneity. For the simulations *with* heterogeneity, I assume that preferences are jointly normal and I maintain $\eta_{c,p} = -0.598$ from section 5.2.

In a first simulation, households exhibit no heterogeneity in preferences or financial and human wealth: they all have the same average preferences from the preferred specification and share the same average wealth shares. Permanent inequality in this case (0.026) amounts to 59.6% of the baseline figure. Even without heterogeneity, wage inequality alone implies a substantial amount of consumption inequality. In a new simulation, households exhibit preference heterogeneity but still share the same average wealth shares. Permanent inequality increases by 19% (to 0.031) amounting to 72% of the baseline figure. In a final simulation, I introduce independent heterogeneity in financial and human wealth, in addition to heterogeneity in preferences. Permanent inequality rises by a further 39% (to 0.043) to match the baseline figure (by construction, empirical and simulated inequality are exactly matched at $\eta_{c,p} = -0.598$).

The simulations suggest that preference heterogeneity increases permanent inequality. This result is not driven by a few extreme preference draws. Partitioning the 10 million households in a big number of random subgroups, and calculating inequality with and without heterogeneity in each subgroup, I observe that preference heterogeneity *always* increases inequality within each subgroup. The result is also robust to trimming the distribution of preferences at a number of alternative thresholds. The same applies to heterogeneity in financial and human wealth; such heterogeneity *always* increases consumption inequality and it has a larger impact on inequality when coupled with preference heterogeneity rather than without it.

A back-of-the-envelope calculation over the different components of the decomposition suggests that preference heterogeneity accounts for 31% of overall consumption inequality (98.8% of consumption instability and 12.4% of permanent inequality), while heterogeneity in financial and human wealth for 22% (28% of permanent inequality).[32] Together they amount to a combined 53%, similar to the contribution of preferences and initial conditions to the variance of consumption (52.6%) in Heathcote et al. (2014). Wage inequality accounts for the remaining portion of inequality, that is, for approximately 47% of the empirical figure. This is numerically similar to the contribution of wage shocks into inequality in Huggett et al. (2011). Unsurprisingly, heterogeneity in financial and human wealth plays a lesser role in the variance of consumption growth; one expects a greater influence of wealth on the variance of consumption and hours levels. Additional simulations not shown here indicate that the means $\mathbb{E}(\pi_{it})$ and $\mathbb{E}(\mathbf{s}_{it})$ have much greater effect on consumption inequality than dispersion around them.

6.2 Implications for Consumption Partial Insurance

The degree of consumption partial insurance is the fraction of a wage shock that does *not* pass through to consumption. The *pass-through rate* of transitory shocks is measured by the partial derivative $\partial\Delta c_{it}/\partial u_{jit}$ whereas that of permanent shocks by $\partial\Delta c_{it}/\partial v_{jit}$. Then $1 - \partial\Delta c_{it}/\partial u_{jit}$ and $1 - \partial\Delta c_{it}/\partial v_{jit}$ are the degrees of partial insurance against transitory and permanent shocks respectively. Preference heterogeneity implies a distribution of partial insurance across households; and the structural model allows me to infer this distribution. Here I refer to the distribution of partial insurance *implied* by the parameter estimates; given that shocks are unobserved, I cannot check whether this matches the *empirical* distribution of partial insurance.

[32] The contribution of preference heterogeneity to consumption inequality varies with the amount of consumption measurement error but it remains always substantial. Appendix table E.8 illustrates this.

Table 7 – Pass-Through Rates of Transitory Shocks into Consumption

	$\mathbb{E}(\frac{\partial \Delta c_{it}}{\partial u_{jit}})$	$\partial \mathbf{\Delta c_{it}} / \partial \mathbf{u_{jit}}$ for household with preferences at: mean	mean + 0.5 s.d.	mean + 1.5 s.d.	mean − 0.5 s.d.	mean − 1.5 s.d.
u_{1it}	-0.033	-0.033	0.118	0.420	-0.184	-0.486
u_{2it}	-0.053	-0.053	0.189	0.672	-0.294	-0.777

Notes: The table presents the pass-through rates of transitory wage shocks into consumption at the sample average of consumption and hours (i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$). The first column reports the average pass-through rates across households; these equal the pass-through rates for the representative household (one with average preferences) in the second column. The remaining columns report pass-through rates for households with preferences up to 1.5 standard deviations away from the mean.

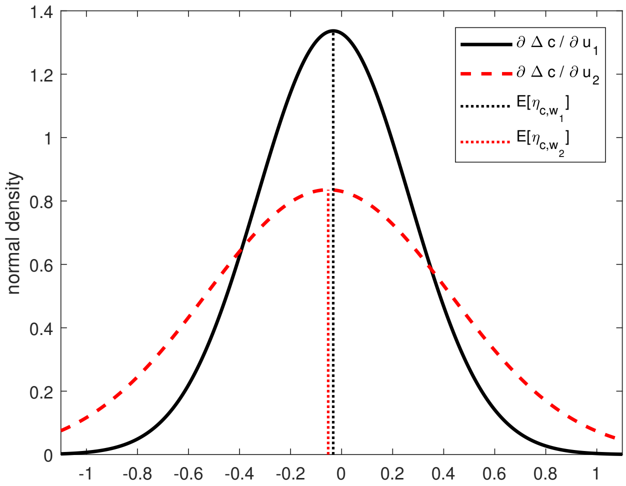
Insurance against transitory shocks. The average pass-through rate of transitory shocks is $\mathbb{E}(\partial \Delta c_{it} / \partial u_{jit}) = \mathbb{E}(\eta_{c,w_j(i,t-1)})$, estimated at -0.033 (*s.e.* 0.032) for male and -0.053 (*s.e.* 0.052) for female shocks in the preferred specification. By construction, these are also the pass-through rates in the representative household (the household with average preferences). Both are indistinguishable from the full insurance benchmark reflecting that consumption is *on average* fully insured against transitory shocks (e.g. Attanasio and Davis, 1996). Blundell et al. (2008) attribute this to self insurance over the lifecycle.

Away from the average, transitory shocks transmit into consumption. Panel (a) of figure 2 illustrates the implied distribution of pass-through rates when the consumption-wage elasticities are jointly normal parameterized at the estimated first and second moments (recall that these moments are estimated without the normality restriction). While consumption for many households is fully insured against transitory shocks, there are households for whom consumption moves considerably with or against transitory shocks. Note that the discussion so far is about the unrestricted response to *lifetime-income-constant* wage changes; this reflects the typical substitution between consumption and leisure net of income effects (unlike the response to permanent shocks below). Finally, figure 2 artificially puts mass over various pass-through rates, including values close to ± 1, due to the normality assumption. In reality, the model offers no information about *how many* households really exhibit these large responses.

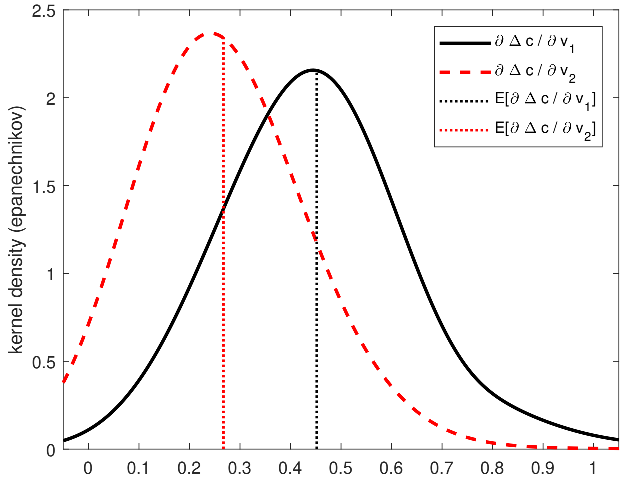
Table 7 reports the pass-through rates for households with preferences 0.5 and 1.5 standard deviations away from the mean; these numbers are *not* specific to a particular preference distribution. The consumption response to transitory shocks already becomes substantial (about ± 0.20) when preferences are 0.5 standard deviation from the mean; at 1.5 standard deviations consumption responds approximately 1-to-$\pm$1/2.

Heterogeneity in pass-through rates reflects heterogeneity in the complementarity between consumption and leisure but it may also reflect liquidity constraints. If the true relationship between consumption and hours is one of Frisch substitution ($\eta_{c,w_j} < 0$) as in BPS, then liquidity constraints mitigate this or flip its sign. Constrained households tend to move consumption

Figure 2 – Distributions of Pass-Through Rates of Shocks into Consumption



(a) transitory shocks



(b) permanent shocks

Notes: The figures visualize the distributions of pass-through rates of shocks across 10 million households whose preferences are drawn from the joint normal parameterized at the estimates from the preferred specification. For the pass-through rates of permanent shocks, π_{it} and $\mathbf{s}_{it}$ are drawn from their empirical distributions, and $\eta_{c,p} = -0.598$. The mass placed over specific pass-through rates is arbitrary: it follows the normality assumption that is neither used in nor inferred by the estimation.

in the same direction with wages and a varying degree of tightness of such constraints induces heterogeneity in the consumption response. I return to this issue in section 6.3.[33]

Insurance against permanent shocks. The average pass-through rate of permanent shocks is $\mathbb{E}(\partial\Delta c_{it})/\partial v_{jit}) = \mathbb{E}(\kappa_{c,v_j(i,t)})$. As per section 3.3, this cannot be expressed in closed form in terms of the preference parameters but it can be constructed numerically. I simulate 10 million households assuming preferences are jointly normal, maintaining $\eta_{c,p} = -0.598$, and drawing wealth shares from their empirical distributions.

The first column of table 8 reports $\mathbb{E}(\partial\Delta c_{it}/\partial v_{jit})$ at 0.45 for male and 0.27 for female permanent shocks. These rates suggest that *on average* 55% of male and 73% of female permanent shocks are insured. There is more insurance against female shocks simply because female earnings are a smaller share in total household earnings. The pass-through rates in the representative household (reported in the second column) are similar.

The pass-through rates are substantially higher (thus less insurance) than BPS who estimate them at 0.34 and 0.20 respectively. There are two main reasons for this. First, matching third moments reduces the female labor supply elasticity and limits the ability of family labor supply to provide insurance to wage shocks. Ghosh (2016) reaches a similar conclusion albeit in the extreme: once she targets consumption and earnings third moments she finds no insurance against persistent shocks but full insurance against transitory ones. However, she abstracts from labor supply so her estimates likely overestimate the pass-through rates. Second, a consumption substitution elasticity $\eta_{c,p} = -0.598$ close to the largest among the BPS estimates (in

[33]Appendix figure E.5 illustrates the implied pass-through rates of shocks when the preference parameter estimates account for different amounts of consumption measurement error.

Table 8 – Pass-Through Rates of Permanent Shocks into Consumption

			no labor supply responses by:				
			men		**women**		**both**
	$\mathbb{E}(\frac{\partial \Delta c_{it}}{\partial v_{jit}})$	average hh	$\mathbb{E}(\frac{\partial \Delta c_{it}}{\partial v_{jit}})$	average hh	$\mathbb{E}(\frac{\partial \Delta c_{it}}{\partial v_{jit}})$	average hh	$\mathbb{E}(\frac{\partial \Delta c_{it}}{\partial v_{jit}})$
v_{1it}	0.452	0.455	0.411	0.430	0.498	0.506	0.497
v_{2it}	0.267	0.271	0.326	0.331	0.250	0.253	0.312

Notes: The table presents the average pass-through rates of permanent wage shocks into consumption and the pass-through rates for the representative household (the household with average preferences).

absolute terms) renders consumption intertemporally more variable and responsive to shocks.[34]

Away from the average, figure 2 illustrates that the distribution of pass-through rates includes both the full insurance (complete markets; $\partial \Delta c_{it}/\partial v_{jit}$ close to 0) and no insurance (autarky; $\partial \Delta c_{it}/\partial v_{jit}$ close to 1) benchmarks as there is some mass near both extrema. This is consistent with Hryshko and Manovskii (2017) who find that partial insurance in the PSID features two polar modes. While they provide an interpretation on the basis of heterogeneity in the wage process, my interpretation is on the basis of preferences. I discuss wage heterogeneity in section 6.3 and I show that it is distinguishable from preference heterogeneity.

As the response of consumption to permanent shocks is partly mitigated by labor supply, table 8 quantifies the role labor supply precisely plays. When male labor supply does not respond ($\eta_{h_1,w_1} = 0$), the pass-through rates of female shocks increases (0.326) compared to the baseline because women lose their husbands' added-worker insurance (à la Lundberg, 1985). However, the pass-through rates of male shocks declines (0.411) because male hours no longer respond positively to own shocks. Similarly, when female labor supply does not respond, the pass-through rates of male shocks increases (0.498) while that of female shocks declines.[35] When neither male nor female labor supply respond, the pass-through rates increase to 0.497 and 0.312 respectively reflecting the loss of insurance through family labor supply.

Overall, out of 54.8 percentage points (pp) of partial insurance to male shocks in the baseline, 4.5pp or 8.2% come from family labor supply; the remaining comes from self-insurance and the mere presence of two earners financing consumption at any time.[36] The analogous

[34]The pass-through rates of permanent shocks are also higher than in Blundell et al. (2008) at 0.31 at the household level (table 7 therein with earnings the closest variable to wages). Alan et al. (2018) find that the central 80% of the distribution of pass-through rates of income shocks falls in the interval 0.05-0.69. This is similar to the central 80% of the distribution of pass-through rates here. Alan et al. (2018) do not distinguish between permanent and transitory shocks making it likelier to find higher insurance as transitory shocks are on average fully insured. Both papers abstract from labor supply and higher moments of earnings and wages.

[35]Labor supply nonresponse implies $\eta_{c,w_j} = 0$ per case as the complementarity between consumption and hours is defined only when hours are variable.

[36]The degree of partial insurance expressed in pp is $(1 - \mathbb{E}(\partial \Delta c_{it}/\partial v_{jit})) * 100$. The fraction for which family

figures for women are 4.5pp and 6.1%. Compared to BPS, family labor supply plays a relatively smaller role here due to the drop in labor supply elasticities as soon as the model targets the skewness of earnings and consumption in the data.

Policy implications. What do the estimates of partial insurance imply in practice? Suppose that a steel company must cut wages of 5000 male employees permanently by 10% in an attempt to avert bankruptcy and layoffs. These men live together with their families in several towns around the company facilities. Assuming that they all have the same average age, consumption, hours, and wealth prior to the wage cut, and subject to partial equilibrium, consumption drops by 3.4% according to BPS, or by \$1,608 annually from average \$47,307 (table D.1) to \$45,699. Over 20 years until age 65, each household will have lost cumulatively about \$32,000 while jointly they will have lost \$160 million. Alternatively, if households have preferences according to figure 2, consumption drops on average by \$2,138 annually, \$43,000 cumulatively, or \$214 million in total. Compared to BPS, an insurer must pay an additional \$54 million in order to compensate this population for forgone consumption.[37]

I repeat WK's main welfare exercise and calculate the welfare loss from idiosyncratic wage risk. WK calculate this in terms of consumption equivalent variation CEV as the root of $W((1+CEV)\boldsymbol{C}^0, \boldsymbol{H}_1^0, \boldsymbol{H}_2^0) = W(\boldsymbol{C}^1, \boldsymbol{H}_1^1, \boldsymbol{H}_2^1)$ where $(\boldsymbol{C}^0, \boldsymbol{H}_1^0, \boldsymbol{H}_2^0)$ is the allocation of consumption and hours without wage risk and $(\boldsymbol{C}^1, \boldsymbol{H}_1^1, \boldsymbol{H}_2^1)$ is the allocation with risk. I calculate this for the *average* household as calculation over a distribution of households requires to calibrate, solve, and simulate WK for a distribution of preferences that reflects my baseline estimates. Results are in appendix table H.2; details on the model are in section 6.4 and appendix H. Using estimates from the preferred specification, the average household is willing to give up about 17% of lifetime consumption in order to be fully insulated from wage risk. This is 2pp more than WK (20% closer to the welfare loss with *exogenous* hours) as households here are more exposed to risk due to a limited insurance role of family labor supply.

6.3 Alternative Explanations for Preference Heterogeneity

Model misspecification may confound preference heterogeneity with the various factors behind the possible misspecification. For example, in the absence of a multi-agent modeling of the household, cross-household and intra-household heterogeneity may interact in the spirit of Lise and Seitz (2011). That environment is, however, not consistent with the empirical pattern of

labor supply is responsible is $\mathbb{E}(\partial\Delta c_{it}/\partial v_{jit})|_{\text{no labor supply}} - \mathbb{E}(\partial\Delta c_{it}/\partial v_{jit})|_{\text{baseline}}$. Even with fixed labor supply the presence of a spouse provides consumption insurance to shocks because of the additional income stream contributing to financing consumption.

[37] This is of course a stylized example. The link between household preferences and household consumption and assets prior to the cut matter for how consumption responds in each household and for the aggregation across households. The baseline model does not allow to study this; a formal treatment requires to simulate lifecycle consumption parametrically for a preference distribution that reflects my baseline parameter estimates.

heterogeneity: the finding that consumption elasticities exhibit heterogeneity but those of labor supply do not, allows me to reject intra-household heterogeneity as such heterogeneity would also be picked up by the variance of the labor supply elasticities (Theloudis, 2017a). A model without labor supply would generally not allow this. Below I discuss in detail the implications of progressive taxation, home production, liquidity constraints, wage heterogeneity, and time aggregation in the PSID.

Progressive joint taxation. Suppose the true sequential budget constraint of the household is $A_{it}(1+r_t)+T_{it}(\sum_{j=1}^{2} W_{jit}H_{jit}) = C_{it}+A_{it+1}$, where $T_{it}(\cdot)$ maps gross earnings to disposable income. As in BPS, $T_{it}(\cdot)$ may capture progressive joint taxation as well as aspects of the public benefits system, such as food stamps and the EITC. Following Heathcote et al. (2014) and BPS, I approximate this as $T_{it}(\sum_{j=1}^{2} W_{jit}H_{jit}) \approx (1-\chi_{it})(\sum_{j=1}^{2} W_{jit}H_{jit})^{1-\nu_{it}}$ with χ_{it} and ν_{it} household- and time-specific parameters that determine the proportionality and progressivity of the tax system. Different values for them give rise to different tax systems. This approximation is handy because it facilitates the analytical expressions of this paper.

The dynamics of consumption and hours differ from baseline expressions (5)-(7) as each equation is augmented by an additional term that reflects the disincentives joint taxation induces on family labor supply. I estimate the preferred specification allowing for progressive joint taxation; the results are in column 1 of appendix table E.9. Joint taxation makes the average labor supply elasticities a bit larger, exactly as in BPS. Importantly, taxation does not change the main pattern of preference heterogeneity: consumption preference heterogeneity remains substantial while heterogeneity in the labor supply elasticities is at least an order of magnitude smaller. Appendix E discusses the dynamics of consumption and hours with progressive joint taxation, as well as the details of the estimation.

Home production. Suppose the true utility function is $U_{it}(C_{it}, H_{1it}, H_{2it}, D_{1it}, D_{2it}; \mathbf{Z}_{it})$ where D_{jit} is the time spouse j devotes to home production. Blundell et al. (2018) define D as parental childcare and show that the household response to wage shocks depends on the substitutability of parental time in home production. Boerma and Karabarbounis (2019) show that heterogeneity in domestic productivity may be mistakenly classified as heterogeneity in preferences. As I abstract from home production here, preference heterogeneity may mask heterogeneity in the home production technology or in the levels of chores and childcare.

This is in principle an omitted variables issue. Consider a household who, as in Blundell et al. (2018), devotes time to childcare. The response of labor supply to a transitory shock reflects the standard substitution between hours and leisure (preferences), the substitution between hours and childcare (preferences and productivity), and the relative share of childcare in a person's overall time. The response of consumption follows a similar rationale. Neglected heterogeneity in productivity or time shares will bias any true preference heterogeneity.

Allowing for home production makes identification without parametric form assumptions particularly tedious, if not impossible. However, I can reestimate all model specifications *also* controlling for (1.) the age of the youngest child (the data do not allow me to control for childcare), and (2.) the hours spouses spend on chores. Assuming that cross-household heterogeneity in home production is either due to differences in children's age or captured by heterogeneity in chores, then accounting for these variables allows me to address the omitted variables issue. The results in appendix tables E.3-E.4 are virtually unchanged from the baseline. This is not surprising. The baseline controls for parental age and number of children in the household, both of which have large predictive power for the age of the youngest child. Moreover, it controls for a large number of observables, which Boerma and Karabarbounis (2019) do not, apparently capturing exhaustively heterogeneity in home production.

Liquidity constraints. The analytical equations for consumption and hours hinge on households being away from liquidity constraints. If liquidity constraints bind, the Euler equation becomes an inequality as households actively save away from the constraint (e.g. Deaton, 1991). This breaks down the neat distinction between permanent and transitory shocks in the present context, as transitory shocks induce substitution effects on outcomes *and* shift the marginal utility of wealth similarly to permanent shocks. As a result, the transmission of transitory shocks into consumption/hours no longer identifies Frisch elasticities only. For example, if the true preference relationship between consumption and labor supply is one of Frisch substitution ($\eta_{c,w_j} < 0$), binding liquidity constraints bias such substitution towards zero as constrained households tend to move consumption in the same direction with wages. The estimation in that case picks up a combination of η_{c,w_j} and wealth effects.

To easiest way to see this point while remaining in the analytical framework of this paper is to consider a modified problem where households can neither save nor borrow, and labor supply is fixed for simplicity (as if there are extreme adjustment costs to work). Keeping the notation similar to the baseline, households solve $\mathbb{E}_0 \sum_{t=0}^{T} U_{it}(C_{it}, \bar{H}_{1it}, \bar{H}_{2it}; \mathbf{Z}_{it})$ subject to the budget constraint $\sum_{j=1}^{2} W_{jit}\bar{H}_{jit} = C_{it}$, $\forall t$. Given the lack of assets, resources cannot be moved across periods and the problem is effectively static. Households are 'hand-to-mouth' (as if there are binding constraints) and consumption equals available income.

The solution to this problem is trivial. Earnings growth reflects wage growth as hours are fixed. The Taylor approximation to the budget constraint yields $\Delta c_{it} \approx q_{1it-1}(v_{1it} + \Delta u_{1it}) + q_{2it-1}(v_{2it} + \Delta u_{2it})$ where q_{jit-1} is spouse j's share of family earnings at $t-1$. All shocks, permanent and transitory, pass through to consumption. A generalization is

$$\Delta c_{it} \approx q_{1it-1}(\vartheta_{1it-1}v_{1it} + \theta_{1it-1}\Delta u_{1it}) + q_{2it-1}(\vartheta_{2it-1}v_{2it-1} + \theta_{2it-1}\Delta u_{2it}) \tag{10}$$

where each shock is associated with a unique loading factor reflecting the tightness of the constraints, i.e. the extent to which saving/borrowing is disallowed. In the extreme case without saving/borrowing, $\vartheta_{jit-1} = \theta_{jit-1} = 1$. On the contrary, if liquidity constraints do not bind and saving/borrowing is reinstated then $\theta_{jit-1} \approx 0$ and $\vartheta_{jit-1} \approx 1 - \pi_{it} > 0$ corresponding to the case of self-insurance with exogenous labor supply.[38]

Naturally, only a fraction of households is subject to the modified problem. Let a proportion ϱ_{t-1} of households solve the baseline problem while a proportion $1-\varrho_{t-1}$ solve the constrained problem above. Whether a household is constrained at the point of its consumption decision at t is determined by its asset position just before, i.e. at $t-1$. Of course households may switch sides from one period to another and I discuss this further in appendix E.

This environment has implications for the parameter estimates and these implications are testable among *unconstrained* households as I explain in appendix E. The difficulty lies in determining which households are unconstrained. Like other micro data, the PSID does not provide information on liquidity constraints. I thus take wealthy households as the empirical counterpart of the theoretically unconstrained but I try different definitions of 'wealthy'.

Appendix table E.9 estimates the preferred specification on wealthy households. The estimator recovers the transmission parameters of transitory shocks in the presence of liquidity constraints (Kaplan and Violante, 2010) so the target moments and all other estimation details remain as in the baseline. Overall, the results are not too different from the baseline and do not reveal a clear case for liquidity constraints.

Based on the specific sample and previous evidence, this is not entirely surprising. Kaplan et al. (2014) classify 1/3 of married and single households in the US as hand-to-mouth and show that their marginal propensity to consume out of income is markedly different from the rest. The difference, however, mostly disappears once they focus on continuously married couples as in here, suggesting the presence of additional insurance among the continuously married. Meyer et al. (2019) show that *singles* constitute the vast majority of the materially constrained. Finally, Aguiar et al. (2020) argue that asset holdings alone do not necessarily reflect truly constrained households; they show that most of those classified as hand-to-mouth in the PSID exhibit behavior that is consistent with preferences for low wealth rather than liquidity constraints. I return to liquidity constraints in the context of a quantitative model in the next section.

Wage process heterogeneity. Suppose that the true wage process is $ARMA(1,1)$ as in Alan et al. (2018) and Hryshko and Manovskii (2017). For each spouse log wage is given by $\ln W_{jit} = \mathbf{X}'_{jit}\boldsymbol{\alpha}_{W_j} + \ln W^p_{jit} + u_{jit}$, the persistent component by $\ln W^p_{jit} = \rho_{ji} \ln W^p_{jit-1} + v_{jit}$, and the transitory component by $u_{jit} = \zeta_{jit} + \tau_{ji}\zeta_{jit-1}$. Here ρ_{ji} is the $AR(1)$ parameter and τ_{ji} is the $MA(1)$ parameter, both heterogeneous across households and spouses. ζ_{jit} is now the serially uncorrelated transitory shock with $\mathbb{E}(\zeta^2_{jit}) \equiv \sigma^2_{\zeta_j(t)}$. This specification implies

$$\Delta w_{jit} = \widetilde{\rho}_{jit} + v_{jit} + \Delta\zeta_{jit} + \tau_{ji}\Delta\zeta_{jit-1} \tag{11}$$

where $\widetilde{\rho}_{jit} = (\rho_{ji}-1)\ln W^p_{jit-1}$. The permanent-transitory process has $\rho_{ji}=1$ and $\tau_{ji}=0$.

[38] When liquidity constraints do not bind, one obtains $\theta_{jit-1} \approx 0$ and $\vartheta_{jit-1} \approx 1-\pi_{it}$ from expression (5) when, due to fixed labor supply, all hours and consumption-wage elasticities are zero.

It helps momentarily to focus on one earner, say the female, keeping male wages constant. For the sake of illustration I also set $\tau_{2i} = 0$. None of these simplifications are crucial for the main result below. The dynamics of consumption are now given by

$$\Delta c_{it} \approx \eta_{c,w_2(i,t-1)}\widetilde{\rho}_{2it} + \eta_{c,w_2(i,t-1)}\left(v_{2it} + \Delta\zeta_{2it}\right) + \bar{\eta}_{c(i,t-1)}\widetilde{\varepsilon}_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}, \rho_{2i})v_{2it}.$$ [39]

The first term captures heterogeneity in consumption growth due to heterogeneity in the wage process, the second term captures the intertemporal substitution due to wage shocks and the last term captures income effects from shifts in the lifetime budget constraint.

In the simplest form, $\mathbb{E}(\eta^2_{c,w_2(i,t)})$ was previously identified by the ratio of the first-order consumption over the first-order wage autocovariance. Here, $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1}, \mathbf{O}_t) = -\mathbb{E}(\eta^2_{c,w_2(i,t)})\times \left\{\sigma^2_{\zeta_2(t)} - \mathbb{E}(\widetilde{\rho}_{2it}\widetilde{\rho}_{2it+1})\right\}$ while $\mathbb{E}(\Delta w_{2it}\Delta w_{2it+1}) = \mathbb{E}(\widetilde{\rho}_{2it}\widetilde{\rho}_{2it+1}) - \sigma^2_{\zeta_2(t)}$. Therefore (minus) the ratio of the two still identifies preference heterogeneity in spite of the heterogeneity in the wage process. Identification of preference heterogeneity is robust to the choice of wage process between the two prominent specifications here because it relies on ratios of intertemporal correlations that are insensitive to the wage specification. By contrast, if identification relied on contemporaneous moments (e.g. on the transmission of permanent shocks), the correct specification of the wage process would matter. The transmission of permanent shocks depends on ρ_{ji} and contemporaneous moments generally pick up preference as well as wage heterogeneity.

Time aggregation in the PSID. The first difference of the *sum* or *average* of a random walk (such as the permanent component of wages in the baseline) exhibits serial correlation even though the underlying variable, by definition, does not. This is an important remark made by Working (1960). The problem arises when a time series is available at a lower frequency (annual earnings) than the frequency at which the underlying variable operates (monthly salary). In the PSID, households report their wage over the last year, which is likely the average over several monthly or other installments. While the covariance between consecutive first differences identifies the variance of the transitory shock in my baseline, Working (1960) shows that such covariance is nonzero for reasons unrelated to the presence of transitory shocks. Crawley (2020) investigates the implications of such neglected time aggregation in the context of the response of consumption to income shocks as in Blundell et al. (2008). He finds that the pass-through of transitory shocks increases materially (and becomes comparable with estimates from quasi-experimental studies) while that of permanent shocks drops.

I highlight the implications of time aggregation using a stylized setting in which wages are paid on a monthly basis but observed annually as the sum over the previous $M = 12$ months. Following the baseline notation closely, let log wage of spouse j in year t and month m be

$$\ln W_{jit,m} = \mathbf{X}'_{jit,m}\boldsymbol{\alpha}_{W_j} + \ln W^p_{jit,m} + u_{jit,m}$$

[39] To obtain this, I plug the new wage process into the log-linearized first-order conditions (4) that are derived independently of the wage process. In addition, I re-approximate the budget constraint in order to map the innovation to λ into shocks to the new wage process. Details of this step appear in appendix A.

$$\ln W^p_{jit,m} = \ln W^p_{jit,m-1} + v_{jit,m}.$$

The wage process thus operates monthly but all properties of shocks, now defined on a monthly basis, are otherwise as in the baseline. Instead of $W_{jit,m}$, one observes the sum $\sum_{m=1}^{M} W_{jit,m}$. Moreover, the data is biennial so wage growth is a first difference over *two* years. I show in appendix G that observed wage growth net of observables is now given by

$$\Delta w_{jit} = \sum_{m=0}^{M-1}(M-m)v_{jit,1+m} + M\sum_{m=1}^{M} v_{jit-1,m} + \sum_{m=1}^{M-1}(M-m)v_{jit-2,M+1-m} + \sum_{m=1}^{M}\Delta u_{jit,m}$$

where $\Delta u_{jit,m} = u_{jit,m} - u_{jit-2,m}$. This expression, markedly different from the baseline process (3), highlights the issues that arise from time aggregation. Δw_{jit} depends not only on transitory shocks at $t-2$ (which the non-aggregated series also does) but also on permanent shocks at $t-2$. Consequently, the covariance between consecutive wage growths does not identify the variance of transitory shocks alone, but a mix of that and the variance of permanent shocks.

Recasting model time from years to months yields analogous expressions for consumption and hours growth.[40] I report these expressions in appendix G, along with expressions for the wage, earnings, and consumption moments that I target in the baseline. I then estimate the preferred specification of the model (appendix table G.1). The main conclusions remain unchanged. The average consumption elasticities are close to zero, the average labor supply elasticities are lower than average estimates in the literature (on average around 0.25-0.30), consumption preference heterogeneity remains substantial and numerically close to the baseline, while heterogeneity in the labor supply elasticities is about an order of magnitude lower. I discuss the estimation details, the results, and some extensions in appendix G.

6.4 Results from a Quantitative Model

WK set up a lifecycle two-earner model with endogenous labor supply and show that the implied partial insurance to male and female wages agrees with the empirical estimates in BPS. They also show that estimation based on the approximation to the problem's first-order conditions correctly recovers the underlying preferences under preference *homogeneity*. In this section I use the parametric model of WK to assess: (i.) the extent to which the baseline approximation and estimation method recovers preferences in the presence of heterogeneity, (ii.) the performance of moments for the transmission of transitory shocks, (iii.) the implications of heterogeneity in the discount factor and the consumption substitution elasticity, and (iv.) the bias induced by binding liquidity constraints.

[40] This assumes that hours and consumption are observed similarly to wages, that is, annually as the sum of the previous 12 months. This is not strictly true for consumption. Certain items are observed as a snapshot in a given month (food) while most items are observed on a varying basis, such as weekly, half-yearly, or yearly. The overall timing of consumption is often ambiguous and subject to alternative interpretations. As my focus here is on the implications of neglected time aggregation in wages, I maintain simplicity by treating consumption in a comparable way to wages. This is of course a crude assumption that future work should relax.

Table 9 – Quantitative Model: Estimates of Average Elasticities

		large sample		small sample	
	(1)	(2)	(3)	(4)	(5)
specification:	true	BPS	transitory	BPS	transitory
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.234	-0.212	-0.229	-0.211 (0.015)	-0.230 (0.024)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.105	-0.122	-0.094	-0.123 (0.014)	-0.097 (0.041)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	1.016	1.058	1.024	1.062 (0.058)	1.026 (0.043)
$\mathbb{E}(\eta_{h_1,w_2(i)})$	0.240	0.223	0.218	0.224 (0.021)	0.219 (0.018)
$\mathbb{E}(\eta_{h_2,w_1(i)})$	0.516	0.537	0.525	0.542 (0.051)	0.527 (0.043)
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.740	0.760	0.768	0.766 (0.052)	0.777 (0.073)

Notes: The table presents GMM estimates of first moments of wage elasticities based on simulated data for four sub-populations of households from the parametric model of WK. Column 1 reports the true average Frisch elasticities in the pooled population. Column 2 estimates preferences using moments of permanent and transitory shocks as in BPS. Column 3 estimates preferences using moments of transitory shocks only as in most of this paper. While columns 2 & 3 use data from four sub-populations of 50000 households each, columns 4 & 5 repeat the estimation drawing a sample of 6032 households only, comparable to my baseline sample. Standard errors are computed based on 500 replications over independent sample draws of similar size. All estimations target second moments of wages and outcomes (third moments are zero by construction) and use households aged 30-60 as in the baseline. Appendix H provides details on the parameterization.

The model in WK abstracts from heterogeneity. Given a sequence of wage shocks, households are all similar. In addition, shocks are normally distributed which is in stark contrast to my baseline that relies on skewness in order to identify preferences.

Extending the model to heterogeneous preferences is straightforward; one can simply simulate multiple sub-populations from the original model, each one based on a different vector of preferences. Pooling all sub-populations together results in a population of households that exhibit preference heterogeneity. Allowing for non-normal shocks is more complicated. While in principle one may replace the normal with a skewed distribution, the solution and simulation methods must be substantially adapted to provide sufficient support in the tails of the state space. This brings about a number of complications (see appendix H) so I do not attempt this extension here. In fact, only a few quantitative models allow for income skewness to date (see De Nardi et al., 2019, for a discussion). Sticking to normality allows me to estimate average elasticities in the presence of heterogeneity but not their second moments whose identification hinges on skewness of wages and outcomes. However, comparing inequality or partial insurance with and without preference heterogeneity is informative for whether the implications of heterogeneity in the model are broadly in line with what I found in the baseline.

All scenarios below are based on simulations of four independent sub-populations, each with different preferences and possibly other attributes described in appendix H. In all cases I compare true average elasticities in the pooled population (*true* because I control the data generating process) with estimates using the baseline system of closed-form equations.

Table 9 presents estimates of average Frisch elasticities using two specifications: BPS (col-

Table 10 – Quantitative Model: Additional Scenarios

	heterogeneity in discount factor		**heterogeneity in $\boldsymbol{\eta_{c,p}}$**		**liquidity constraints**	
	(1)	(2)	(3)	(4)	(5)	(6)
specification:	true	transitory	true	transitory	true	transitory
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.229	-0.178	-0.173	-0.178	-0.234	-0.229
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.104	-0.081	-0.075	-0.077	-0.104	-0.092
$\mathbb{E}(\eta_{h_1,w_1(i)})$	1.009	0.933	0.858	0.860	1.016	1.023
$\mathbb{E}(\eta_{h_1,w_2(i)})$	0.238	0.181	0.161	0.147	0.237	0.216
$\mathbb{E}(\eta_{h_2,w_1(i)})$	0.509	0.436	0.357	0.354	0.515	0.524
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.739	0.739	0.662	0.669	0.738	0.765

Notes: The table presents GMM estimates of first moments of wage elasticities based on simulated data for four sub-populations of households from the parametric model of WK. In addition to heterogeneity in the parameters governing the estimable Frisch elasticities, columns 1 & 2 feature heterogeneity in the discount factor, columns 3 & 4 feature heterogeneity in $\eta_{c,p}$, and columns 5 & 6 feature severe liquidity constraints. Appendix H provides details on the parameterization.

umn 2) and the specification that uses moments of transitory shocks only (column 3). The underlying sub-populations differ in the utility parameters that control the substitution between consumption, male hours, and female hours.[41] Two points emerge. First, both specifications recover average preferences very precisely in the presence of heterogeneity. The elasticities in the underlying sub-populations vary around WK's calibration (for example, η_{h_1,w_1} varies in $0.8-1.3$) and the method successfully recovers them also when applied to each sub-population separately. Second, estimation in samples of size comparable to my baseline is equally good (columns 4-5). This confirms that the baseline approximation and estimation method can successfully recover preferences in the presence of heterogeneity. Moreover, moments for the transmission of transitory shocks (which carry most of the estimation burden in this paper) convey sufficient information for identification.

Table 10 presents estimates from the latter specification under three additional scenarios: heterogeneity in the discount factor, heterogeneity in the consumption substitution elasticity, and when 30% of households have no access to credit *and* debt in the first 10 years of life equal to one quarter of average male earnings upon entry in the labor market (e.g. student debt).[42] The data in all three cases come from four sub-populations, each with different parameters governing the estimable Frisch elasticities (appendix H). The estimation recovers preferences

[41]The utility function in WK is given by $U_t = \beta^t(1-\sigma)^{-1}\big(\{\alpha C_{it}^{\gamma} + (1-\alpha)[\xi H_{1it}^{\theta} + (1-\xi)H_{2it}^{\theta}]^{-\frac{\gamma}{\theta}}\}^{\frac{1-\sigma}{\gamma}} - 1\big)$. Parameters γ and θ control the substitution between consumption, male hours, and female hours.

[42]The discount factor varies over $\beta = \{0.95; 0.97; 0.98; 0.99\}$ while the consumption substitution parameter varies over $\sigma = \{2.1; 2.24; 2.8; 3.0\}$. These values revolve around WK's default calibration, but the specific choices are ultimately arbitrary.

remarkably well in all cases. The largest bias appears in the male labor supply elasticity when households differ in their discount factors (the bias stems from those who are less patient). Appendix table H.1 confirms that the estimation recovers preferences also when I combine heterogeneity in the discount factor with heterogeneity in $\eta_{c,p}$.[43]

Overall, the results suggest that the quasi-reduced-form equations (5)-(7) successfully recover average elasticities even in the presence of preference heterogeneity. Depending on the specification, the contribution of heterogeneity to consumption inequality is about 5% in the early years of life and increases gradually to about 40% in the end. This is in the same order of magnitude as the back-of-the-envelope calculation in my baseline even though calibration of heterogeneity in the quantitative model is arbitrary, i.e. not informed by the data, and rather moderate. Heterogeneity also implies a sizable distribution of consumption insurance.

7 Conclusions

This paper studies the link between wage and consumption inequality using a lifecycle model of family labor supply, consumption and wealth. Although this is not a new topic, the paper distinctively brings a general preference heterogeneity into the nexus of family labor supply and consumption. By doing so, it formalizes an intuitive idea: inequality is not driven by wages and wealth only, but also by different preferences. I show identification of all moments of the joint distribution of consumption and labor supply wage elasticities without resorting to a specific utility function or preference distribution. Identification rests on the idea that cross-sectional dispersion in outcomes net of prices and observables reflects preference heterogeneity.

A summary of the key empirical takeaways is as follows. First, unexplained heterogeneity in consumption preferences is substantial while heterogeneity in labor supply preferences is not. A few recent papers obtain implicitly a similar result. Heathcote et al. (2014) estimate substantial preference heterogeneity when they use consumption data but not when they only use labor supply data. Attanasio et al. (2018) find that heterogeneity in female labor supply elasticities is limited after accounting for observables and hours across households.

Second, preference heterogeneity increases inequality in consumption growth. This implies that consumption inequality within a given cohort increases over time not only because of wage risk etc., but also because different households respond differently to changes in their environment. This is important for how inequality responds to redistributive policies or for how much redistribution may be desirable. Moreover, cross-sectional consumption data do not only reflect a distribution of prices, resources and needs, but also a distribution of preferences. Measuring preference heterogeneity appropriately and incorporating it into models of inequality or redistribution is thus important for the policy conclusions we draw.

[43] Liquidity constraints cause little bias to the parameters as Blundell et al. (2013), who study the approximation error, and WK confirm. The is because most households have accumulated assets away from the constraint by the first lifecycle period that enters the baseline estimation (age 30, corresponding to model period 10).

Third, matching third moments of outcomes and wages lowers the average labor supply elasticities and renders family labor supply less effective in insuring against shocks. So consumption insurance is lower when higher moments are targeted. Although there may be also other mechanisms through which higher moments affect partial insurance (e.g. De Nardi et al. (2019) find that a non-normal wage process results in a different degree of partial insurance compared to the permanent-transitory process), this confirms that moments *beyond* the workhorse covariance of consumption and income matter for partial insurance. Ghosh (2016) reaches a similar conclusion by adding third and fourth moments in Blundell et al. (2008).

Seen together, studies of inequality or partial insurance should allow a role for preference heterogeneity and higher moments. Previous attempts to allow for heterogeneity typically involve a scalar component in the utility function or the income process. The problem with this approach is that such component affects consumption and hours similarly, while heterogeneity is likely higher dimensional. In addition, higher moments matter for consumption insurance and it is instructive to measure the distribution of such insurance and assess how well alternative models can match it.

References

Abowd, J. M. and D. Card (1989). On the covariance structure of earnings and hours changes. *Econometrica 57*(2), 411–45.

Aguiar, M. and E. Hurst (2013). Deconstructing life cycle expenditure. *Journal of Political Economy 121*(3), 437–492.

Aguiar, M. A., M. Bils, and C. Boar (2020). Who are the hand-to-mouth? Working Paper 26643, National Bureau of Economic Research.

Alan, S. and M. Browning (2010). Estimating intertemporal allocation parameters using synthetic residual estimation. *Review of Economic Studies 77*(4), 1231–1261.

Alan, S., M. Browning, and M. Ejrnæs (2018). Income and consumption: A micro semistructural analysis with pervasive heterogeneity. *Journal of Political Economy 126*(5), 1827–1864.

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2008). Eliciting risk and time preferences. *Econometrica 76*(3), 583–618.

Andrews, D. W. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica 68*(2), 399–405.

Arellano, M., R. Blundell, and S. Bonhomme (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica 85*(3), 693–734.

Attanasio, O., G. Berloffa, R. Blundell, and I. Preston (2002). From earnings inequality to consumption inequality. *Economic Journal 112*(478), C52–C59.

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., P. Levell, H. Low, and V. Sanchez-Marcos (2018). Aggregating elasticities: Intensive and extensive margins of women's labor supply. *Econometrica 86*(6), 2049–2082.

Attanasio, O., H. Low, and V. Sánchez-Marcos (2005). Female labor supply as insurance against idiosyncratic risk. *Journal of the European Economic Association 3*(2/3), 755–764.

Attanasio, O., H. Low, and V. Sánchez-Marcos (2008). Explaining changes in female labor supply in a life-cycle model. *American Economic Review 98*(4), 1517–1552.

Azzalini, A. and A. D. Valle (1996). The multivariate skew-normal distribution. *Biometrika 83*(4), 715–726.

Blundell, R. and B. Etheridge (2010). Consumption, income and earnings inequality in Britain. *Review of Economic Dynamics 13*(1), 76–102.

Blundell, R., J. Horowitz, and M. Parey (2017). Nonparametric estimation of a nonseparable demand function under the Slutsky inequality restriction. *The Review of Economics and Statistics 99*(2), 291–304.

Blundell, R., H. Low, and I. Preston (2013). Decomposing changes in income risk using consumption data. *Quantitative Economics 4*(1), 1–37.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887–1921.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2016). Consumption inequality and family labor supply. *American Economic Review 106*(2), 387–435.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2018). Children, time allocation, and consumption insurance. *Journal of Political Economy 126*(S1), S73–S115.

Blundell, R. and I. Preston (1998). Consumption inequality and income uncertainty. *Quarterly Journal of Economics 113*(2), 603–640.

Boerma, J. and L. Karabarbounis (2019). Inferring inequality with home production. NBER Working Paper 24166, NBER.

Bohrnstedt, G. and A. Goldberger (1969). On the exact covariance of products of random variables. *Journal of the American Statistical Association 64*(328), 1439–1442.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics 12*(3), pp. 345–368.

Browning, M., M. Ejrnæs, and J. Alvarez (2010). Modelling income processes with lots of heterogeneity. *Review of Economic Studies 77*(4), 1353–1381.

Campbell, J. Y. (1993). Intertemporal asset pricing without consumption data. *American Economic Review 83*(3), pp. 487–512.

Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review 96*(5), 1821–1834.

Cohen, A. and L. Einav (2007). Estimating risk preferences from deductible choice. *American Economic Review 97*(3), 745–788.

Cosaert, S. and T. Demuynck (2018). Nonparametric welfare and demand analysis with unobserved individual heterogeneity. *The Review of Economics and Statistics 100*(2), 349–361.

Crawley, E. (2020). In search of lost time aggregation. *Economics Letters 189*.

De Nardi, M., G. Fella, and G. Paz-Pardo (2019). Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare. *Journal of the European Economic Association 18*(2), 890–926.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica 59*(5), 1221–1248.

Deaton, A. and C. Paxson (1994). Intertemporal choice and inequality. *Journal of Political Economy 102*(3), 437–67.

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies 72*(2), 395–427.

Gayle, G.-L. and A. Shephard (2019). Optimal taxation, marriage, home production, and family labor supply. *Econometrica 87*(1), 291–326.

Ghosh, A. (2016). Income versus consumption inequality: the role of time-varying higher moments. Unpublished manuscript.

Goodman, L. (1960). On the exact variance of products. *Journal of the American Statistical Association 55*(292), 708–713.

Gottschalk, P. and R. Moffitt (2009). The rising instability of us earnings. *Journal of Economic Perspectives 23*(4), 3–24.

Guiso, L. and M. Paiella (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic Association 6*(6), 1109–1150.

Guiso, L. and P. Sodini (2013). *Household Finance: An Emerging Field*, Volume 2 of *Handbook of the Economics of Finance*, pp. 1397–1532. Elsevier.

Guvenen, F. (2007). Learning your earning: Are labor income shocks really very persistent? *American Economic Review 97*(3), 687–712.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015). What do data on millions of U.S. workers reveal about life-cycle earnings risk? NBER Working Paper 20913, NBER.

Guvenen, F. and A. A. Smith (2014). Inferring labor income risk and partial insurance from economic choices. *Econometrica 82*(6), 2085–2129.

Heathcote, J., F. Perri, and G. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the United States. *Review of Economic Dynamics 13*(1), 15–51.

Heathcote, J., K. Storesletten, and G. L. Violante (2014). Consumption and labor supply with partial insurance: An analytical framework. *American Economic Review 104*(7), 2075–2126.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy 109*(4), 673–748.

Horowitz, J. L. (2001). The Bootstrap. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5 of *Handbook of Econometrics*, pp. 3159 – 3228. Elsevier.

Hryshko, D. and I. Manovskii (2017). How much consumption insurance in the US? NBER Summer Institute.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *American Economic Review 101*(7), 2923–54.

Hyslop, D. R. (2001). Rising U.S. earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *American Economic Review 91*(4), 755–777.

Jappelli, T. and L. Pistaferri (2010). The consumption response to income changes. *Annual Review of Economics 2*(1), 479–506.

Kaplan, G. and G. Violante (2010). How much consumption insurance beyond self-insurance? *American Economic Journal: Macroeconomics 2*(4), 53–87.

Kaplan, G., G. L. Violante, and J. Weidner (2014). The wealthy hand-to-mouth. *Brookings Papers on Economic Activity 45*, 77–153.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature 49*(4), 961–1075.

Kimball, M. S., C. R. Sahm, and M. D. Shapiro (2009). Risk preferences in the PSID: Individual imputations and family covariation. *American Economic Review 99*(2), 363–368.

Krueger, D. and F. Perri (2006). Does income inequality lead to consumption inequality? evidence and theory. *Review of Economic Studies 73*(1), 163–193.

Lewbel, A. (2001). Demand systems with and without errors. *American Economic Review 91*(3), 611–618.

Lewbel, A. and K. Pendakur (2017). Unobserved preference heterogeneity in demand using generalized random coefficients. *Journal of Political Economy 125*(4), 1100–1148.

Lise, J. and S. Seitz (2011). Consumption inequality and intra-household allocations. *Review of Economic Studies 78*(1), 328–355.

Low, H., C. Meghir, and L. Pistaferri (2010). Wage risk and employment risk over the life cycle. *American Economic Review 100*(4), 1432–67.

Lundberg, S. (1985). The added worker effect. *Journal of Labor Economics 3*(1), 11–37.

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.

Matzkin, R. L. (2003). Nonparametric estimation of nonadditive random functions. *Econometrica 71*(5), 1339–1375.

Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Meghir, C. and L. Pistaferri (2011). Earnings, consumption and life cycle choices. Volume 4 of *Handbook of Labor Economics*, pp. 773–854. Elsevier.

Meyer, B. D., D. Wu, V. R. Mooers, and C. Medalia (2019). The use and misuse of income data and extreme poverty in the united states. Working Paper 25907, National Bureau of Economic Research.

Phlips, L. (1974). *Applied Consumption Analysis.* Advanced textbooks in economics. North-Holland.

Primiceri, G. E. and T. van Rens (2009). Heterogeneous life-cycle profiles, income risk and consumption inequality. *Journal of Monetary Economics 56*(1), 20 – 39.

PSID. Panel Study of Income Dynamics, public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (final data downloaded in 2019).

Theloudis, A. (2017a). Consumption inequality across heterogeneous families. LISER Working Paper Series 2017-18, LISER.

Theloudis, A. (2017b). *Essays in Consumption Inequality and the Allocation of Household Resources.* Ph. D. thesis, University College London.

Working, H. (1960). Note on the correlation of first differences of averages in a random chain. *Econometrica 28*(4), 916–918.

Wu, C. and D. Krueger (2021). Consumption insurance against wage risk: Family labor supply and optimal progressive income taxation. *American Economic Journal: Macroeconomics 13*(1), 79–113.

Appendix

A Taylor Approximations of First-Order Conditions and Lifetime Budget Constraint

Suppose the household utility function takes the form

$$U_{it}(C_{it}, H_{1it}, H_{2it}; \mathbf{Z}_{it}) = \beta_{it}(\mathbf{Z}_i) U_i(C_{it}, H_{1it}, H_{2it}; \mathbf{Z}_{it}) \equiv \widetilde{\beta}_{it} \widetilde{U}_i(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it})$$

where $\widetilde{\beta}_{it} = \beta_{it}(\mathbf{Z}_i)$, $\widetilde{C}_{it} = C_{it} \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha}_C)$, $\widetilde{H}_{jit} = H_{jit} \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha}_{H_j})$ for $j = \{1, 2\}$, and $\mathbf{Z}_i$ is the time invariant portion of $\mathbf{Z}_{it}$. Normalizing the price of consumption at $P_t = 1\ \forall t$, and assuming an interior solution and geometric discounting ($\widetilde{\beta}_{it} = \widetilde{\beta}_i^t$), the first-order conditions of household problem (1) *s.t.* (2) are

$$\begin{aligned}
[C_{it}]: &\quad \widetilde{U}_{iC}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it}) \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha}_C) = \lambda_{it} \\
[H_{jit}]: &\quad -\widetilde{U}_{iH_j}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it}) \exp(-\mathbf{Z}'_{it}\boldsymbol{\alpha}_{H_j}) = \lambda_{it} W_{jit}, \quad j = \{1, 2\} \\
[A_{it+1}]: &\quad \widetilde{\beta}_i (1 + r_{t+1}) \mathbb{E}_t \lambda_{it+1} = \lambda_{it}.
\end{aligned}$$

$\widetilde{U}_{iC}$ is the marginal utility of consumption and $\widetilde{U}_{iH_j}$ the marginal utility of hours of spouse j; λ_{it} is the marginal utility of wealth (the Lagrange multiplier on the sequential budget constraint). In the remaining of this appendix I follow steps similar to Blundell et al. (2013) and BPS.

Approximation of intra-temporal first-order conditions. Applying logs to the intra-temporal first-order conditions and taking a first difference in time yields

$$\begin{aligned}
[C_{it}]: &\quad \Delta \ln \widetilde{U}_{iC}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it}) - \Delta\left(\mathbf{Z}'_{it}\boldsymbol{\alpha}_C\right) = \Delta \ln \lambda_{it} \\
[H_{jit}]: &\quad \Delta \ln\left(-\widetilde{U}_{iH_j}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it})\right) - \Delta\left(\mathbf{Z}'_{it}\boldsymbol{\alpha}_{H_j}\right) = \Delta \ln \lambda_{it} + \Delta \ln W_{jit}, \quad j = \{1, 2\}.
\end{aligned}$$

A first-order approximation of $\ln \widetilde{U}_{iC}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it})$ around $\widetilde{C}_{it-1}$, $\widetilde{H}_{1it-1}$, and $\widetilde{H}_{2it-1}$ yields

$$\begin{aligned}
\Delta \ln \widetilde{U}_{iC}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it}) \approx \widetilde{U}_{iC}^{-1}(\widetilde{C}_{it-1}, \widetilde{H}_{1it-1}, \widetilde{H}_{2it-1}) \times \Big(&\widetilde{U}_{iCC}(\widetilde{C}_{it-1}, \widetilde{H}_{1it-1}, \widetilde{H}_{2it-1}) \widetilde{C}_{it-1} \Delta \ln C_{it} \\
&+ \widetilde{U}_{iCH_1}(\widetilde{C}_{it-1}, \widetilde{H}_{1it-1}, \widetilde{H}_{2it-1}) \widetilde{H}_{1it-1} \Delta \ln H_{1it} \\
&+ \widetilde{U}_{iCH_2}(\widetilde{C}_{it-1}, \widetilde{H}_{1it-1}, \widetilde{H}_{2it-1}) \widetilde{H}_{2it-1} \Delta \ln H_{2it}\Big)
\end{aligned}$$

where $\widetilde{U}_{iCC}$ denotes the derivative of $\widetilde{U}_{iC}$ with respect to consumption C, etc. The approximation of $-\widetilde{U}_{iH_j}(\widetilde{C}_{it}, \widetilde{H}_{1it}, \widetilde{H}_{2it})$ follows similarly. Replacing $\Delta \ln \widetilde{U}_{iC}$ and $\Delta \ln\left(-\widetilde{U}_{iH_j}\right)$ with their approximation counterparts yields a system of 3 equations in the growth rates of outcomes $\Delta \ln C_{it}$, $\Delta \ln H_{1it}$, and $\Delta \ln H_{2it}$. Solving the system and rearranging so that all outcome variables and observable taste shifters are on the left results in system (4) in the text. Note that the Frisch elasticities are defined on the basis of first & second order derivatives of the utility function, as well as of levels of consumption and hours (see appendix B for definitions). Given the nature of the approximation, these derivatives and outcome levels concern period $t-1$. The relevant Frisch elasticities in this system thus also concern period $t-1$.

Approximation of Euler equation. The approximation of the inter-temporal first-order condition involves future expectations. Suppose $\exp(\Gamma_{it+1}) = {}^{1}/_{\widetilde{\beta}_i(1+r_{t+1})}$ for some Γ_{it+1}. I apply a second-order approximation of $\exp(\ln \lambda_{it+1})$ around $\ln \lambda_{it} + \Gamma_{it+1}$ to get

$$\exp(\ln \lambda_{it+1}) \approx \exp(\ln \lambda_{it} + \Gamma_{it+1}) \left(1 + (\Delta \ln \lambda_{it+1} - \Gamma_{it+1}) + \frac{1}{2}(\Delta \ln \lambda_{it+1} - \Gamma_{it+1})^2\right).$$

Taking expectations at time t and noting that $\mathbb{E}_t \lambda_{it+1} = \lambda_{it} \exp(\Gamma_{it+1})$ yields

$$\mathbb{E}_t \left(\Delta \ln \lambda_{it+1} - \Gamma_{it+1} + \frac{1}{2}(\Delta \ln \lambda_{it+1} - \Gamma_{it+1})^2\right) \approx 0$$

which in turn implies

$$\begin{aligned} \mathbb{E}_t \Delta \ln \lambda_{it+1} &\approx \Gamma_{it+1} - \frac{1}{2}\mathbb{E}_t(\Delta \ln \lambda_{it+1} - \Gamma_{it+1})^2 \\ \Delta \ln \lambda_{it+1} &\approx \omega_{it+1} + \varepsilon_{it+1}. \end{aligned} \tag{A.1}$$

The first term $\omega_{it+1} = \Gamma_{it+1} - \frac{1}{2}\mathbb{E}_t(\Delta \ln \lambda_{it+1} - \Gamma_{it+1})^2$ captures the impact of interest rate r_{t+1}, impatience $\widetilde{\beta}_i$ and precautionary motives on consumption growth (Blundell et al., 2013). To maintain tractability, I fix $\mathbb{E}_t(\Delta \ln \lambda_{it+1})^2$ (precautionary motive) in the cross-section. ω_{it} is heterogeneous across households due to Γ_{it+1} -the anticipated gradient of outcome growth- but non-stochastic. The second term is an expectation error with $\mathbb{E}_t(\varepsilon_{it+1}) = 0$; it captures idiosyncratic revisions to λ upon arrival of new information, namely of wage shocks.

Approximation of lifetime budget constraint (draws on Campbell (1993)'s log-linearization of the intertemporal budget constraint)**.** Normalizing the price of consumption at $P_t = 1$, $\forall t$, the general form of household i's lifetime budget constraint is

$$A_{it} + \mathbb{E}_t \sum_{s=0}^{T-t} \sum_{j=1}^{2} \frac{W_{jit+s} H_{jit+s}}{(1+r)^s} = \mathbb{E}_t \sum_{s=0}^{T-t} \frac{C_{it+s}}{(1+r)^s}.$$

To ease the notation I will temporarily suppress cross-sectional subscript i.

Let $G(\boldsymbol{\xi}) = \ln \sum_{s=0}^{T-t} \exp \xi_s$ for $\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_{T-t})'$. Applying a first-order approximation of $G(\boldsymbol{\xi})$ around a deterministic $\boldsymbol{\xi}^0$, and taking expectations conditional on some information set $\mathcal{I}$, yields

$$\mathbb{E}_{\mathcal{I}} G(\boldsymbol{\xi}) \approx G(\boldsymbol{\xi}^0) + \sum_{s=0}^{T-t} \frac{\exp \xi_s^0}{\sum_{\kappa=0}^{T-t} \exp \xi_\kappa^0} \left(\mathbb{E}_{\mathcal{I}} \xi_s - \xi_s^0\right). \tag{A.2}$$

The logarithm of the *right* hand side of the budget constraint, assuming expectations away, is

$$G^{RH}(\boldsymbol{\xi}) = \ln \sum_{s=0}^{T-t} \exp \left(\ln \frac{C_{t+s}}{(1+r)^s}\right)$$

for $\xi_s = \ln C_{t+s} - s\ln(1+r)$. Suppose $\xi_s^0 = \mathbb{E}_{t-1}\ln C_{t+s} - s\ln(1+r)$. Following (A.2) I write

$$\mathbb{E}_{\mathcal{I}} G^{RH}(\boldsymbol{\xi}) \approx G^{RH}(\boldsymbol{\xi}^0) + \sum_{s=0}^{T-t} \theta_{t+s}\left(\mathbb{E}_{\mathcal{I}} \ln C_{t+s} - \mathbb{E}_{t-1}\ln C_{t+s}\right)$$

where $\theta_{t+s} = \frac{\exp(\mathbb{E}_{t-1}\ln C_{t+s} - s\ln(1+r))}{\sum_{\kappa=0}^{T-t}\exp(\mathbb{E}_{t-1}\ln C_{t+\kappa} - \kappa\ln(1+r))}$ is approximately equal to the expected share of time $t+s$ consumption in total lifetime consumption. θ_{t+s} is known at any $t+s \geq t$ because it pertains to expectations at $t-1$ (before time t shocks realize).

Defining the information set to contain information known at time t, that is $\mathcal{I} := t$, and replacing $\ln C_{t+s}$ consecutively by the analytical expression in (4) yields

$$\sum_{s=0}^{T-t} \theta_{t+s}\left(\mathbb{E}_t \ln C_{t+s} - \mathbb{E}_{t-1}\ln C_{t+s}\right) \approx \bar{\eta}_{c(t-1)}\varepsilon_t + \sum_{j=1}^{2}\eta_{c,w_j(t-1)}v_{jt} + \sum_{j=1}^{2}\theta_t\eta_{c,w_j(t-1)}u_{jt}$$

where $\bar{\eta}_c = \eta_{c,p} + \eta_{c,w_1} + \eta_{c,w_2}$ (ω_{t+s} is non-stochastic and disappears in the first difference). If the share of a period's consumption in total lifetime consumption is negligible, i.e. $\theta_t \approx 0$, then taking a first difference in expectations between t and $t-1$ and reinstating i yields

$$\mathbb{E}_t G^{RH}(\boldsymbol{\xi}) - \mathbb{E}_{t-1}G^{RH}(\boldsymbol{\xi}) \approx \bar{\eta}_{c(i,t-1)}\varepsilon_{it} + \sum_{j=1}^{2}\eta_{c,w_j(i,t-1)}v_{jit}.$$

Applying similar arguments to the *left* hand side of the budget constraint yields

$$\mathbb{E}_t G^{LH}(\boldsymbol{\xi}) - \mathbb{E}_{t-1}G^{LH}(\boldsymbol{\xi}) \approx (1-\pi_{it})\sum_{j=1}^{2}\left\{s_{jit}\bar{\eta}_{h_j(i,t-1)}\varepsilon_{it} + \left(s_{jit}(1+\eta_{h_j,w_j(i,t-1)}) + s_{-jit}\eta_{h_{-j},w_j(i,t-1)}\right)v_{jit}\right\}$$

where $\bar{\eta}_{h_j} = \eta_{h_j,p} + \eta_{h_j,w_1} + \eta_{h_j,w_2}$, $-j$ denotes the spouse, and

$$G^{LH}(\boldsymbol{\xi}) = \ln\left(\exp\left(\ln A_t\right) + \sum_{s=1}^{T-t+1}\exp\left(\ln\sum_{j=1}^{2}\frac{W_{jt+s-1}H_{jt+s-1}}{(1+r)^{s-1}}\right)\right)$$

$$\xi_s = \begin{cases} \ln A_{t+s} & \text{for } s=0 \\ \ln\sum_{j=1}^{2}W_{jt+s-1}H_{jt+s-1} - (s-1)\ln(1+r) & \text{for } s=1,\dots,T-t+1 \end{cases}$$

$$\xi_s^0 = \begin{cases} \mathbb{E}_{t-1}\ln A_{t+s} & \text{for } s=0 \\ \mathbb{E}_{t-1}\ln\sum_{j=1}^{2}W_{jt+s-1}H_{jt+s-1} - (s-1)\ln(1+r) & \text{for } s=1,\dots,T-t+1. \end{cases}$$

The rest of the notation is as follows: $\pi_t = \frac{Q_{1t}}{Q_{1t}+Q_{2t}}$, with $Q_{1t} = \exp\left(\mathbb{E}_{t-1}\ln A_t\right)$ and $Q_{2t} = \sum_{\kappa=0}^{T-t}\exp\left(\mathbb{E}_{t-1}\ln\sum_j W_{jt+\kappa}H_{jt+\kappa} - \kappa\ln(1+r)\right)$, is the 'partial insurance' coefficient: it is approximately equal to the share of financial wealth in the household's total financial and human wealth at t. $s_{jt} = \sum_{s=0}^{T-t}\vartheta_{t+s}\widetilde{q}_{jt+s}$, with $\vartheta_{t+s} = \exp\left(\mathbb{E}_{t-1}\ln\sum_{j=1}^{2}W_{jt+s}H_{jt+s} - s\ln(1+r)\right)/Q_{2t}$ and $\widetilde{q}_{jt+s} = \frac{\mathbb{E}_{t-1}W_{jt+s}H_{jt+s}}{\sum_{\iota=1}^{2}\mathbb{E}_{t-1}W_{\iota t+s}H_{\iota t+s}}$, is approximately equal to the share of spouse j's human wealth (expected lifetime earnings) in the household's total human wealth at t. ϑ_{t+s} and $\widetilde{q}_{jt+s}$ are known at $t+s \geq t$ (they pertain to $t-1$ expectations before shocks at t realize) and $\vartheta_{t+s} \approx 0$ if the share of a period's earnings in the household's total lifetime earnings is negligible.

I bring the two sides together following Blundell et al. (2013) and solve for ε_{it} to get

$$\varepsilon_{it} \approx \varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}) v_{1it} + \varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}) v_{2it} \tag{A.3}$$

where

$$\begin{aligned}
\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}) &= \frac{(1-\pi_{it})\left(s_{1it}(1+\eta_{h_1,w_1(i,t-1)}) + s_{2it}\eta_{h_2,w_1(i,t-1)}\right) - \eta_{c,w_1(i,t-1)}}{\bar{\eta}_{c(i,t-1)} - (1-\pi_{it})\left(s_{1it}\bar{\eta}_{h_1(i,t-1)} + s_{2it}\bar{\eta}_{h_2(i,t-1)}\right)} \\
\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}) &= \frac{(1-\pi_{it})\left(s_{1it}\eta_{h_1,w_2(i,t-1)} + s_{2it}(1+\eta_{h_2,w_2(i,t-1)})\right) - \eta_{c,w_2(i,t-1)}}{\bar{\eta}_{c(i,t-1)} - (1-\pi_{it})\left(s_{1it}\bar{\eta}_{h_1(i,t-1)} + s_{2it}\bar{\eta}_{h_2(i,t-1)}\right)}
\end{aligned}$$

and $\mathbf{s}_{it} = (s_{1it}, s_{2it})'$ with $s_{1it} + s_{2it} = 1$ by construction. $\boldsymbol{\eta}$ is the 9×1 vector of household-specific Frisch elasticities presented in table 1 and defined in appendix B.

Estimable equations. Combining the approximations of the intra-temporal first-order conditions in (4), the Euler equation in (A.1), and the lifetime budget constraint in (A.3), I obtain

$$\begin{aligned}
\Delta c_{it} &\approx \bar{\eta}_{c(i,t-1)}\omega_{it} + \eta_{c,w_1(i,t-1)}\Delta u_{1it} + \eta_{c,w_2(i,t-1)}\Delta u_{2it} \\
&\quad + \left(\eta_{c,w_1(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{1it} + \left(\eta_{c,w_2(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{2it} \\
\Delta h_{jit} &\approx \bar{\eta}_{h_j(i,t-1)}\omega_{it} + \eta_{h_j,w_1(i,t-1)}\Delta u_{1it} + \eta_{h_j,w_2(i,t-1)}\Delta u_{2it} \\
&\quad + \left(\eta_{h_j,w_1(i,t-1)} + \bar{\eta}_{h_j(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{1it} + \left(\eta_{h_j,w_2(i,t-1)} + \bar{\eta}_{h_j(i,t-1)}\varepsilon_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})\right) v_{2it}.
\end{aligned}$$

The intercepts $\bar{\eta}_{c(i,t-1)}\omega_{it}$ and $\bar{\eta}_{h_j(i,t-1)}\omega_{it}$ reflect the gradient of outcome growth in the absence of shocks. As shown above, these are functions of the interest rate, the discount factor and precautionary motives, and they also depend on the past levels of consumption and hours through $\bar{\eta}_{c(i,t-1)}$ and $\bar{\eta}_{h_j(i,t-1)}$. Identification of moments of the intercepts is straightforward. However, this does not identify moments of the discount factor $\widetilde{\beta}_i$ because two consecutive moments of $\widetilde{\beta}_i$ appear additively in any moment of ω_{it} and separating them is impossible. Therefore I absorb $\bar{\eta}_{c(i,t-1)}\omega_{it}$ and $\bar{\eta}_{h_j(i,t-1)}\omega_{it}$ into the intercepts of the first-stage regression of consumption and hours growth on observables, including on past levels of consumption and hours, resulting in the main estimating equations (5)-(7).

Approximation with heterogeneous wage process. Following similar steps but replacing the permanent-transitory wage process (3) with the $ARMA(1,1)$ process (11) yields

$$\mathbb{E}_t G^{RH}(\boldsymbol{\xi}) - \mathbb{E}_{t-1} G^{RH}(\boldsymbol{\xi}) \approx \bar{\eta}_{c(i,t-1)}\varepsilon_{it} + \sum_{j=1}^{2} \eta_{c,w_j(i,t-1)} f_c(\rho_{ji}) v_{jit}$$

$$\begin{aligned}
\mathbb{E}_t G^{LH}(\boldsymbol{\xi}) - \mathbb{E}_{t-1} G^{LH}(\boldsymbol{\xi}) &\approx (1-\pi_{it})\left(s_{1it}\bar{\eta}_{h_1(i,t-1)} + s_{2it}\bar{\eta}_{h_2(i,t-1)}\right)\varepsilon_{it} \\
&\quad + (1-\pi_{it})\sum_{j=1}^{2}\left(f_{h_j}(s_{jit}, \rho_{ji})(1+\eta_{h_j,w_j(i,t-1)}) + f_{h_{-j}}(s_{-jit}, \rho_{ji})\eta_{h_{-j},w_j(i,t-1)}\right) v_{jit}
\end{aligned}$$

where $f_c(\rho_{ji}) = 1 + \sum_{s=1}^{T-t} \theta_{it+s}(\rho_{ji}^s - 1)$, $f_{h_{j'}}(s_{j'it}, \rho_{ji}) = s_{j'it} + \sum_{s=1}^{T-t} \vartheta_{it+s}\widetilde{q}_{j'it+s}(\rho_{ji}^s - 1)$ with $f_c(\rho_{ji} = 1) = 1$, $f_{h_{j'}}(s_{j'it}, \rho_{ji} = 1) = s_{j'it}$ and $j, j' = \{1, 2\}$. Bringing the two sides together yields

$$\varepsilon_{it} \approx \widetilde{\varepsilon}_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}, \rho_{1i})v_{1it} + \widetilde{\varepsilon}_2(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}, \rho_{2i})v_{2it}$$

where $\widetilde{\varepsilon}_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}, \rho_{ji})$ equals

$$\frac{(1-\pi_{it})\left(f_{h_j}(s_{jit}, \rho_{ji})(1+\eta_{h_j,w_j(i,t-1)}) + f_{h_{-j}}(s_{-jit}, \rho_{ji})\eta_{h_{-j},w_j(i,t-1)}\right) - f_c(\rho_{ji})\eta_{c,w_j(i,t-1)}}{\bar{\eta}_{c(i,t-1)} - (1-\pi_{it})\left(s_{1it}\bar{\eta}_{h_1(i,t-1)} + s_{2it}\bar{\eta}_{h_2(i,t-1)}\right)}$$

and $\widetilde{\varepsilon}_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1}, \rho_{ji} = 1) = \varepsilon_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})$. Combining this with previous results yields the equation for consumption growth in the presence of a heterogeneous wage process.

B Frisch Elasticities

Period preferences $U_i(C_{it}, H_{1it}, H_{2it})$ are described ordinally by 9 Frisch elasticities. There are 9 elasticities because there are 3 goods (C, H_1, H_2) and 3 prices (P, W_1, W_2); consequently there are 3 own-price and 6 cross-price elasticities. Their analytical expressions are

$$\begin{aligned}
\eta_{c,w_j(i,t)} &= \frac{\partial C}{\partial W_j}\frac{W_j}{C}\bigg|_{\lambda\text{-const.}}^{i,t} &&= Det^{-1}\frac{U_{H_j}}{C}(U_{CH_{-j}}U_{H_1H_2} - U_{CH_j}U_{H_{-j}H_{-j}}) \\
\eta_{c,p(i,t)} &= \frac{\partial C}{\partial P}\frac{P}{C}\bigg|_{\lambda\text{-const.}}^{i,t} &&= Det^{-1}\frac{U_C}{C}(U_{H_1H_1}U_{H_2H_2} - U_{H_1H_2}^2) \\
\eta_{h_j,w_j(i,t)} &= \frac{\partial H_j}{\partial W_j}\frac{W_j}{H_j}\bigg|_{\lambda\text{-const.}}^{i,t} &&= Det^{-1}\frac{U_{H_j}}{H_j}(U_{CC}U_{H_{-j}H_{-j}} - U_{CH_{-j}}^2) \\
\eta_{h_j,w_{-j}(i,t)} &= \frac{\partial H_j}{\partial W_{-j}}\frac{W_{-j}}{H_j}\bigg|_{\lambda\text{-const.}}^{i,t} &&= Det^{-1}\frac{U_{H_{-j}}}{H_j}(U_{CH_1}U_{CH_2} - U_{CC}U_{H_1H_2}) \\
\eta_{h_j,p(i,t)} &= \frac{\partial H_j}{\partial P}\frac{P}{H_j}\bigg|_{\lambda\text{-const.}}^{i,t} &&= Det^{-1}\frac{U_C}{H_j}(U_{CH_{-j}}U_{H_1H_2} - U_{CH_j}U_{H_{-j}H_{-j}})
\end{aligned}$$

where $j = \{1, 2\}$, $-j$ denotes the spouse, U_o the marginal utility with respect to outcome variable $o = \{C, H_1, H_2\}$ and $U_{o\chi}$ the derivative of U_o with respect to $\chi = \{C, H_1, H_2\}$. $Det = U_{CC}U_{H_1H_1}U_{H_2H_2} + 2U_{CH_1}U_{CH_2}U_{H_1H_2} - U_{CC}U_{H_1H_2}^2 - U_{H_1H_1}U_{CH_2}^2 - U_{H_2H_2}U_{CH_1}^2$ is the determinant of the Hessian matrix of preferences. All partial derivatives, outcome variables, and the determinant are i- and t-specific but I suppress these subscripts to ease the notation. The partial effects are calculated at the household level holding λ constant in expectation.

From Phlips (1974), section 2.4, the matrix of substitution effects after a marginal-utility-of-wealth-compensated price change is

$$\begin{pmatrix} \frac{dC}{dP} & -\frac{dC}{dW_1} & -\frac{dC}{dW_2} \\ \frac{dH_1}{dP} & -\frac{dH_1}{dW_1} & -\frac{dH_1}{dW_2} \\ \frac{dH_2}{dP} & -\frac{dH_2}{dW_1} & -\frac{dH_2}{dW_2} \end{pmatrix} = \lambda_i \mathbf{H}_i^{-1}\mathbf{I}_3 \quad \text{(B.1)}$$

where $\mathbf{H}_i$ is the Hessian of U_i and $\mathbf{I}_3$ is a 3×3 identity matrix. I obtain (B.1) by totally differentiating the intra-temporal first-order conditions of the household problem with respect to prices and noting that $\Delta\lambda_{it} = 0$ in expectation. As the right hand side of (B.1) is a symmetric matrix (the Hessian is symmetric by Young's theorem and standard regularity conditions on U_i), it follows that $\frac{dH_j}{dP} = -\frac{dC}{dW_j}$ and $\frac{dH_1}{dW_2} = \frac{dH_2}{dW_1}$. Simple manipulations of these restrictions translate into restrictions on the corresponding cross-price Frisch elasticities.

C Identification Details

C.1 Wage Process

Second moments. There are 6 parameters for the cross-sectional dispersion of shocks at t, namely $\sigma^2_{v_j(t)}$, $\sigma^2_{u_j(t)}$, $\sigma_{v_1v_2(t)}$ and $\sigma_{u_1u_2(t)}$ for $j = \{1, 2\}$. Identification requires second moments of the joint distribution of spouses' wages across households. Consider wage process (3) and assume measurement error away for now. The covariance between consecutive wage growths $\mathbb{E}\left(\Delta w_{jit}\Delta w_{jit+1}\right)$ identifies the variance of the transitory shock due to mean reversion. The covariance between contemporaneous wage growth and a sum of three consecutive wage growths $\mathbb{E}\left(\Delta w_{jit}\sum_{\varsigma=-1}^{\varsigma=1}\Delta w_{jit+\varsigma}\right)$ identifies the variance of the permanent shock as the sum strips Δw_{jit} of the time-t transitory shock. Formally

$$\begin{aligned}
\sigma^2_{v_j(t)} &= \mathbb{E}\left(\Delta w_{jit}(\Delta w_{jit-1} + \Delta w_{jit} + \Delta w_{jit+1})\right)\\
\sigma^2_{u_j(t)} &= -\mathbb{E}\left(\Delta w_{jit}\Delta w_{jit+1}\right)\\
\sigma_{v_1v_2(t)} &= \mathbb{E}\left(\Delta w_{1it}(\Delta w_{2it-1} + \Delta w_{2it} + \Delta w_{2it+1})\right)\\
\sigma_{u_1u_2(t)} &= -\mathbb{E}\left(\Delta w_{1it}\Delta w_{2it+1}\right).
\end{aligned}$$

Third moments. There are 8 parameters for the cross-sectional skewness of shocks at t, namely $\gamma_{v_j(t)}$, $\gamma_{u_j(t)}$, $\gamma_{v_1v_2^2(t)}$, $\gamma_{v_1^2v_2(t)}$, $\gamma_{u_1u_2^2(t)}$ and $\gamma_{u_1^2u_2(t)}$ for $j = \{1, 2\}$. Identification parallels that for the second moments; formally

$$\begin{aligned}
\gamma_{v_j(t)} &= \mathbb{E}\left((\Delta w_{jit})^2(\Delta w_{jit-1} + \Delta w_{jit} + \Delta w_{jit+1})\right)\\
\gamma_{u_j(t)} &= -\mathbb{E}\left((\Delta w_{jit})^2\Delta w_{jit+1}\right)\\
\gamma_{v_1v_2^2(t)} &= \mathbb{E}\left((\Delta w_{2it})^2(\Delta w_{1it-1} + \Delta w_{1it} + \Delta w_{1it+1})\right)\\
\gamma_{v_1^2v_2(t)} &= \mathbb{E}\left((\Delta w_{1it})^2(\Delta w_{2it-1} + \Delta w_{2it} + \Delta w_{2it+1})\right)\\
\gamma_{u_1u_2^2(t)} &= -\mathbb{E}\left((\Delta w_{2it})^2\Delta w_{1it+1}\right)\\
\gamma_{u_1^2u_2(t)} &= -\mathbb{E}\left((\Delta w_{1it})^2\Delta w_{2it+1}\right).
\end{aligned}$$

Higher moments. Generalization to the n^{th} moment ($n > 1$) is straightforward. The n^{th} moment of permanent shocks is identified through $\mathbb{E}\left((\Delta w_{jit})^{n-1}(\Delta w_{jit-1} + \Delta w_{jit} + \Delta w_{jit+1})\right)$

(own moments) and $\mathbb{E}\left((\Delta w_{2it})^{n-\nu}\left(\Delta w_{1it-1}+\Delta w_{1it}+\Delta w_{1it+1}\right)^{\nu}\right)$ (cross-moments) with $\nu=\{1,\dots,n-1\}$. These moments convey information on $\mathbb{E}(v^n_{jit})$ and $\mathbb{E}(v^{\nu}_{1it}v^{n-\nu}_{2it})$ respectively *plus* a sum of products of lower-order moments (up to $n-2\geq 2$) of the spouses' permanent and transitory shocks between $t-2$ and $t+1$. Such lower-order moments are identified sequentially relying on results for the variance and skewness and then moving up, if required, until reaching moments of order $n-2$.

The n^{th} moment of transitory shocks is identified through the n^{th} covariance of wages; namely $\mathbb{E}\left((\Delta w_{jit})^{n-1}\Delta w_{jit+1}\right)$ (own moments) and $\mathbb{E}\left((\Delta w_{2it})^{n-\nu}(\Delta w_{1it+1})^{\nu}\right)$ (cross-moments). The covariances convey information on $\mathbb{E}(u^n_{jit})$ and $\mathbb{E}(u^{\nu}_{1it}u^{n-\nu}_{2it})$ respectively *plus*, like previously, a sum of products of lower-order moments (order up to $n-2\geq 2$) of the spouses' permanent and transitory wage shocks between $t-1$ and $t+1$.

Three remarks are in order. First, unlike the variance and skewness, it is not possible to identify higher than third moments without previously identifying lower-order moments. Second, no generic formulas exist for $\mathbb{E}(v^n_{jit})$, $\mathbb{E}(u^n_{jit})$ etc, *for every n*. The reason is the accompanying sum of products of lower-order moments in each case. Such term depends on (and increases with) n; lower-order sums are not nested within higher-order sums thus ruling out a generic formula *for every n*. Third, over-identifying restrictions exist for all own- and cross-moments of order higher than 2.

Measurement error. Identification so far assumes measurement error away. If wages are contaminated with error, then identification requires moments of the error up to order n. In the permanent-transitory specification the variance of the transitory shock is not separately identified from the variance of measurement error (Meghir and Pistaferri, 2011). However, information on the variance of wage error in survey data is often available through validation studies (Bound et al., 1994, for the PSID). The previous identifying equations can be easily adapted for an external estimate of the error variance.

It is not possible to separate higher moments of wage measurement error from those of the transitory shock. Moreover, there are no validation studies (at least not for the PSID) with information on the error's distributional aspects. This thus necessitates additional assumptions, for example the Gaussian assumption in proposition 1 of section 3.1.

C.2 Preferences

Outline. There are 9 parameters for the first moment of the conditional joint distribution $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ of Frisch elasticities across households, given a value for $\mathbf{O}_t=(C_t,H_{1t},H_{2t})'$. These are $\mathbb{E}(\eta_{c,w_j(i,t)}|\mathbf{O}_t)$, $\mathbb{E}(\eta_{c,p(i,t)}|\mathbf{O}_t)$, $\mathbb{E}(\eta_{h_{j'},w_j(i,t)}|\mathbf{O}_t)$, and $\mathbb{E}(\eta_{h_{j'},p(i,t)}|\mathbf{O}_t)$ for $j,j'=\{1,2\}$. There are 45 parameters for the conditional second moment; table C.1 lists these parameters. In general, there are $\prod_{i=1}^{8}(n+i)/8!$ parameters for the conditional $n^{\text{th}}=\{1,2,3,\dots\}$ moment, assuming that such moment exists and is finite.

Table C.1 – Second Moments of Preference Distribution $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$

	Consumption elasticities			*Male labor supply elasticities*			*Female labor supply elasticities*		
	$\eta_{c,w_1(i,t)}$	$\eta_{c,w_2(i,t)}$	$\eta_{c,p(i,t)}$	$\eta_{h_1,w_1(i,t)}$	$\eta_{h_1,w_2(i,t)}$	$\eta_{h_1,p(i,t)}$	$\eta_{h_2,w_1(i,t)}$	$\eta_{h_2,w_2(i,t)}$	$\eta_{h_2,p(i,t)}$
$\eta_{c,w_1(i,t)}$	$\mathrm{V}(\eta_{c,w_1})$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{c,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{c,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_1,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_1,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_1,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_1}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{c,w_2(i,t)}$		$\mathrm{V}(\eta_{c,w_2})$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{c,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_1,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_1,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_1,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,w_2}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{c,p(i,t)}$			$\mathrm{V}(\eta_{c,p})$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_1,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_1,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_1,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{c,p}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_1,w_1(i,t)}$				$\mathrm{V}(\eta_{h_1,w_1})$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_1}\\\eta_{h_1,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_1}\\\eta_{h_1,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_1}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_1}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_1}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_1,w_2(i,t)}$					$\mathrm{V}(\eta_{h_1,w_2})$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_2}\\\eta_{h_1,p}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_2}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_2}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,w_2}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_1,p(i,t)}$						$\mathrm{V}(\eta_{h_1,p})$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,p}\\\eta_{h_2,w_1}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,p}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_1,p}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_2,w_1(i,t)}$							$\mathrm{V}(\eta_{h_2,w_1})$	$\mathrm{C}\begin{pmatrix}\eta_{h_2,w_1}\\\eta_{h_2,w_2}\end{pmatrix}$	$\mathrm{C}\begin{pmatrix}\eta_{h_2,w_1}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_2,w_2(i,t)}$								$\mathrm{V}(\eta_{h_2,w_2})$	$\mathrm{C}\begin{pmatrix}\eta_{h_2,w_2}\\\eta_{h_2,p}\end{pmatrix}$
$\eta_{h_2,p(i,t)}$									$\mathrm{V}(\eta_{h_2,p})$

Notes: The table lists the 45 parameters that characterize the second moment of the conditional distribution of preferences $F_{\boldsymbol{\eta}_t|\mathbf{O}_t}$ for a given value for $\mathbf{O}_t = (C_t, H_{1t}, H_{2t})'$. Parameters that refer exclusively to wage elasticities appear without shade. These are all identified from panel data on consumption, hours, and wages. Parameters that involve moments of the labor supply elasticities with respect to the price of consumption (own moments and cross-moments with wage elasticities) appear in light gray. These are linear transformations of the previous moments, thus identified. Moments of the consumption elasticity with respect to its price appear in dark gray; these are not identified. V denotes the cross-sectional variance and C the covariance.

I group these parameters into three categories. The first involves exclusively moments of *wage elasticities*. The second involves moments of the labor supply elasticities *with respect to the price of consumption* $\eta_{h_j,p}$, including their cross-moments with wage elasticities. The third includes all other parameters, namely moments of the consumption elasticity *with respect to its price* $\eta_{c,p}$. In table C.1, second moments belonging to the first category appear without shade, those in the second appear in light gray, and those in the third appear in dark gray.

Detailed derivation of consumption autocovariance. Abstracting from measurement error for now, the first-order consumption autocovariance that appears in the text is given by

$$\begin{aligned} m_{cc(t)} &= \mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t) \\ &= -\,\mathbb{E}(\eta^2_{c,w_1(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_1(t)} - \mathbb{E}(\eta^2_{c,w_2(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma^2_{u_2(t)} \\ &\quad - 2\mathbb{E}(\eta_{c,w_1(i,t)}\eta_{c,w_2(i,t)}|\mathbf{O}_{t-1},\mathbf{O}_t)\sigma_{u_1u_2(t)}. \end{aligned} \tag{C.1}$$

I condition this moment on both $\mathbf{O}_{t-1}$ and $\mathbf{O}_t$ because, by the nature of the approximation, Δc_{it} is a function of $\boldsymbol{\eta}_{it-1}$ while Δc_{it+1} is a function of $\boldsymbol{\eta}_{it}$. As these $\boldsymbol{\eta}$'s depend on outcome levels at $t-1$ and t, the appropriate conditioning variables given identifying *assumption 1b* are both $\mathbf{O}_{t-1}$ and $\mathbf{O}_t$.

I will now show in detail how I obtain expression (C.1). To simplify the illustration I assume temporarily that female shocks $v_{2it} = u_{2it} = 0, \forall t$, so as per (5), $\Delta c_{it} \approx \eta_{c,w_1(i,t-1)}\Delta u_{1it} +$

$g(\pi_{it}, \mathbf{s}_{it}, \boldsymbol{\eta}_{it-1})v_{1it}$ where $g(\pi_{it}, \mathbf{s}_{it}, \boldsymbol{\eta}_{it-1}) = \eta_{c,w_1(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_1(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})$. Using expressions (3) and (5), assumptions 1a and 1b, and results from Bohrnstedt and Goldberger (1969) for the covariance of products of random variables, one can show

$$
\begin{aligned}
m_{cc(t)} &= \mathbb{E}\big\{\big(\eta_{c,w_1(i,t-1)} \times \Delta u_{1it} + g(\pi_{it}, \mathbf{s}_{it}, \boldsymbol{\eta}_{it-1}) \times v_{1it}\big) \\
&\quad \times \big(\eta_{c,w_1(i,t)} \times \Delta u_{1it+1} + g(\pi_{it+1}, \mathbf{s}_{it+1}, \boldsymbol{\eta}_{it}) \times v_{1it+1}\big) \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big\} \\
&= \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times \Delta u_{1it} \times \eta_{c,w_1(i,t)} \times \Delta u_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 1)} \\
&\quad + \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times \Delta u_{1it} \times g(\pi_{it+1}, \mathbf{s}_{it+1}, \boldsymbol{\eta}_{it}) \times v_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 2)} \\
&\quad + \mathbb{E}\big(g(\pi_{it}, \mathbf{s}_{it}, \boldsymbol{\eta}_{it-1}) \times v_{1it} \times \eta_{c,w_1(i,t)} \times \Delta u_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 3)} \\
&\quad + \mathbb{E}\big(g(\pi_{it}, \mathbf{s}_{it}, \boldsymbol{\eta}_{it-1}) \times v_{1it} \times g(\pi_{it+1}, \mathbf{s}_{it+1}, \boldsymbol{\eta}_{it}) \times v_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 4)} \\
&= \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times u_{1it} \times \eta_{c,w_1(i,t)} \times u_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 1.1)} \\
&\quad - \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times u_{1it-1} \times \eta_{c,w_1(i,t)} \times u_{1it+1} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 1.2)} \\
&\quad - \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times u_{1it} \times \eta_{c,w_1(i,t)} \times u_{1it} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 1.3)} \\
&\quad + \mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times u_{1it-1} \times \eta_{c,w_1(i,t)} \times u_{1it} \mid \mathbf{O}_{t-1}, \mathbf{O}_t\big) && \text{(term 1.4)} \\
&= -\,\mathbb{E}\big(\eta_{c,w_1(i,t-1)} \times \eta_{c,w_1(i,t)} \mid \mathbf{O}_{t-1}, \mathbf{O}_t)\big)\sigma^2_{u_1(t)} \\
&= -\,\mathbb{E}(\eta^2_{c,w_1(i,t)} | \mathbf{O}_{t-1}, \mathbf{O}_t)\sigma^2_{u_1(t)}.
\end{aligned}
$$

In detail: term 2 drops out because shocks are exogenous (*assumption 1a*), thus past preferences do not vary with future shocks. Moreover, conditional on $\mathbf{O}_{t-1}$, $\eta_{c,w_1(i,t-1)}$ does not vary with $\Delta u_{1it} = u_{1it} - u_{1it-1}$ (*assumption 1b*). Elasticity $\eta_{c,w_1(i,t-1)}$ likely varies with $g(\pi_{it+1}, \mathbf{s}_{it+1}, \boldsymbol{\eta}_{it})$ through $\boldsymbol{\eta}_{it}$. However, transitory and permanent shocks are independent so the entire term equals zero. Similar arguments prove that terms 3 and 4 are also zero. Term 1 survives these arguments because Δu_{1it} and Δu_{1it+1} are related through u_{1it}.

Expanding term 1, term 1.1 drops out because transitory shocks are exogenous (*assumption 1a*; thus past preferences do not vary with future shocks) and serially uncorrelated. Terms 1.2 and 1.4 drop out because, conditional on the contemporaneous levels of consumption and hours, shocks and elasticities are independent (*assumption 1b*). Term 1.3 survives these arguments. The last equality is then true because I have assumed that conditional elasticities in immediately adjacent ages are approximately equal, i.e. $\eta_{c,w_j(i,t-1)}|\mathbf{O}_{t-1}, \mathbf{O}_t \approx \eta_{c,w_j(i,t)}|\mathbf{O}_{t-1}, \mathbf{O}_t$. This assumption simplifies the notation and is not needed for identification. The full expression for $m_{cc(t)}$ in (C.1) obtains by reinstating the female shocks which, for the sake of illustration, I had temporarily shut down.

Expressions for other moments. I derive expressions for all other moments similarly. To ease the notation in this section, I no longer include the conditioning statement $|\mathbf{O}_t$ (or $|\mathbf{O}_{t-1}, \mathbf{O}_t$ when applicable) in the elasticities; it is, however, implied that all elasticity moments are conditional on the same consumption and hours levels on which the data moments are also conditioned. Finally, to maintain consistency with BPS, I use earnings y_j in lieu of hours h_j.

Growth in log earnings net of observables is given by $\Delta y_{jit} = \Delta w_{jit} + \Delta h_{jit}$. The full list of identifying moments is given by

$$
\begin{aligned}
m_{w_j c(t)} &= \mathbb{E}\big(\Delta w_{jit+1}\Delta c_{it}|\mathbf{O}_{t-1}\big) = -\mathbb{E}(\eta_{c,w_j(i,t)})\sigma^2_{u_j(t)} - \mathbb{E}(\eta_{c,w_{j'}(i,t)})\sigma_{u_1u_2(t)} \\
m_{w_j c^2(t)} &= \mathbb{E}\big(\Delta w_{jit+1}(\Delta c_{it})^2|\mathbf{O}_{t-1})\big) \\
&= -\mathbb{E}(\eta^2_{c,w_j(i,t)})\gamma_{u_j(t)} - \mathbb{E}(\eta^2_{c,w_{j'}(i,t)})\gamma_{u_j u^2_{j'}(t)} - 2\mathbb{E}(\eta_{c,w_1(i,t)}\eta_{c,w_2(i,t)})\gamma_{u^2_j u_{j'}(t)} \\
m_{w_j y_j(t)} &= \mathbb{E}\big(\Delta w_{jit+1}\Delta y_{jit}|\mathbf{O}_{t-1}\big) = -\mathbb{E}(1+\eta_{h_j,w_j(i,t)})\sigma^2_{u_j(t)} - \mathbb{E}(\eta_{h_j,w_{j'}(i,t)})\sigma_{u_1u_2(t)} \\
m_{w_{j'} y_j(t)} &= \mathbb{E}\big(\Delta w_{j'it+1}\Delta y_{jit}|\mathbf{O}_{t-1}\big) = -\mathbb{E}(1+\eta_{h_j,w_j(i,t)})\sigma_{u_1u_2(t)} - \mathbb{E}(\eta_{h_j,w_{j'}(i,t)})\sigma^2_{u_{j'}(t)} \\
m_{y_j y_j(t)} &= \mathbb{E}\big(\Delta y_{jit}\Delta y_{jit+1}|\mathbf{O}_{t-1},\mathbf{O}_t\big) \\
&= -\mathbb{E}((1+\eta_{h_j,w_j(i,t)})^2)\sigma^2_{u_j(t)} - \mathbb{E}(\eta^2_{h_j,w_{j'}(i,t)})\sigma^2_{u_{j'}(t)} - 2\mathbb{E}((1+\eta_{h_j,w_j(i,t)})\eta_{h_j,w_{j'}(i,t)})\sigma_{u_1u_2(t)} \\
m_{w_j y^2_j(t)} &= \mathbb{E}\big(\Delta w_{jit+1}(\Delta y_{jit})^2|\mathbf{O}_{t-1}\big) \\
&= -\mathbb{E}((1+\eta_{h_j,w_j(i,t)})^2)\gamma_{u_j(t)} - \mathbb{E}(\eta^2_{h_j,w_{j'}(i,t)})\gamma_{u_j u^2_{j'}(t)} - 2\mathbb{E}((1+\eta_{h_j,w_j(i,t)})\eta_{h_j,w_{j'}(i,t)})\gamma_{u^2_j u_{j'}(t)} \\
m_{w_{j'} y^2_j(t)} &= \mathbb{E}\big(\Delta w_{j'it+1}(\Delta y_{jit})^2|\mathbf{O}_{t-1}\big) \\
&= -\mathbb{E}((1+\eta_{h_j,w_j(i,t)})^2)\gamma_{u^2_j u_{j'}(t)} - \mathbb{E}(\eta^2_{h_j,w_{j'}(i,t)})\gamma_{u_{j'}(t)} - 2\mathbb{E}((1+\eta_{h_j,w_j(i,t)})\eta_{h_j,w_{j'}(i,t)})\gamma_{u_j u^2_{j'}(t)} \\
m_{y_j c(t)} &= \mathbb{E}\big(\Delta y_{jit+1}\Delta c_{it}|\mathbf{O}_{t-1},\mathbf{O}_t\big) \\
&= -\mathbb{E}(\eta_{c,w_j(i,t)}(1+\eta_{h_j,w_j(i,t)}))\sigma^2_{u_j(t)} - \mathbb{E}(\eta_{c,w_{j'}(i,t)}\eta_{h_j,w_{j'}(i,t)})\sigma^2_{u_{j'}(t)} \\
&\quad - \big(\mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_j,w_{j'}(i,t)}) + \mathbb{E}(\eta_{c,w_{j'}(i,t)}(1+\eta_{h_j,w_j(i,t)}))\big)\sigma_{u_1u_2(t)} \\
m_{w_j y_j c(t)} &= \mathbb{E}\big(\Delta w_{jit+1}\Delta y_{jit}\Delta c_{it}|\mathbf{O}_{t-1}\big) \\
&= -\mathbb{E}(\eta_{c,w_j(i,t)}(1+\eta_{h_j,w_j(i,t)}))\gamma_{u_j(t)} - \mathbb{E}(\eta_{c,w_{j'}(i,t)}\eta_{h_j,w_{j'}(i,t)})\gamma_{u_j u^2_{j'}(t)} \\
&\quad - \big(\mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_j,w_{j'}(i,t)}) + \mathbb{E}(\eta_{c,w_{j'}(i,t)}(1+\eta_{h_j,w_j(i,t)}))\big)\gamma_{u^2_j u_{j'}(t)} \\
m_{w_{j'} y_j c(t)} &= \mathbb{E}\big(\Delta w_{j'it+1}\Delta y_{jit}\Delta c_{it}|\mathbf{O}_{t-1}\big) \\
&= -\mathbb{E}(\eta_{c,w_j(i,t)}(1+\eta_{h_j,w_j(i,t)}))\gamma_{u^2_j u_{j'}(t)} - \mathbb{E}(\eta_{c,w_{j'}(i,t)}\eta_{h_j,w_{j'}(i,t)})\gamma_{u_{j'}(t)} \\
&\quad - \big(\mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_j,w_{j'}(i,t)}) + \mathbb{E}(\eta_{c,w_{j'}(i,t)}(1+\eta_{h_j,w_j(i,t)}))\big)\gamma_{u_j u^2_{j'}(t)}
\end{aligned}
$$

where $j, j' = \{1,2\}$ and $j \neq j'$.

Now I show that functions of these empirical moments identify first and second moments of wage elasticities. Identification proceeds as follows.

The mean wage elasticities, given values for $\mathbf{O}_{t-1}$ and/or $\mathbf{O}_t$, are identified through a combination of wage moments and joint consumption-wage and earnings-wage moments; namely

$$
\begin{aligned}
\mathbb{E}(\eta_{c,w_j(i,t)}) &= \Big(m_{w_{j'}c(t)}\sigma_{u_1u_2(t)} - m_{w_j c(t)}\sigma^2_{u_{j'}(t)}\Big) \,/\, \big(\sigma^2_{u_1(t)}\sigma^2_{u_2(t)} - (\sigma_{u_1u_2(t)})^2\big) \\
\mathbb{E}(\eta_{h_j,w_j(i,t)}) &= \Big(m_{w_{j'}y_j(t)}\sigma_{u_1u_2(t)} - m_{w_j y_j(t)}\sigma^2_{u_{j'}(t)}\Big) \,/\, \big(\sigma^2_{u_1(t)}\sigma^2_{u_2(t)} - (\sigma_{u_1u_2(t)})^2\big) - 1 \\
\mathbb{E}(\eta_{h_j,w_{j'}(i,t)}) &= \Big(m_{w_j y_j(t)}\sigma_{u_1u_2(t)} - m_{w_{j'}y_j(t)}\sigma^2_{u_j(t)}\Big) \,/\, \big(\sigma^2_{u_1(t)}\sigma^2_{u_2(t)} - (\sigma_{u_1u_2(t)})^2\big)\,.
\end{aligned}
$$

These parameters are heavily over-identified by many additional moments. In addition, symmetry of the matrix of Frisch substitution effects (see appendix B) imposes linear restrictions between reciprocal cross-elasticities. As a result the following relation must hold: $\mathbb{E}(\eta_{h_2,w_1(i,t)}) = \mathbb{E}(\eta_{h_1,w_2(i,t)})\mathbb{E}(Y_{1it}/Y_{2it})$ where Y_{jit} is earnings of spouse j.

The second moments of the consumption-wage elasticities (upper left triangle in table C.1), given values for $\mathbf{O}_{t-1}$ and/or $\mathbf{O}_t$, are identified from wage, consumption, and joint consumption-wage moments; namely

$$\begin{pmatrix} \sigma^2_{u_1(t)} & \sigma^2_{u_2(t)} & 2\sigma_{u_1u_2(t)} \\ \gamma_{u_1(t)} & \gamma_{u_1u_2^2(t)} & 2\gamma_{u_1^2u_2(t)} \\ \gamma_{u_1^2u_2(t)} & \gamma_{u_2(t)} & 2\gamma_{u_1u_2^2(t)} \end{pmatrix} \begin{pmatrix} \mathbb{E}(\eta^2_{c,w_1(i,t)}) \\ \mathbb{E}(\eta^2_{c,w_2(i,t)}) \\ \mathbb{E}(\eta_{c,w_1(i,t)}\eta_{c,w_2(i,t)}) \end{pmatrix} = - \begin{pmatrix} m_{cc(t)} \\ m_{w_1c^2(t)} \\ m_{w_2c^2(t)} \end{pmatrix}.$$

The system is linear in the parameters. The matrix of coefficients is nonsingular if the distribution of shocks is asymmetric about the mean, that is, if shocks are skewed. The matrix is nonsingular also if $\sigma_{u_1u_2} = 0$ or $\gamma_{u_1^2u_2} = \gamma_{u_1u_2^2} = 0$. If all cross-moments of shocks are zero, the matrix is singular and the covariance of elasticities is not identified but the variances are.

The second moments of the male and female labor supply elasticities (bottom middle and right triangles in table C.1) are identified similarly from wage, earnings, and joint earnings-wage moments. Given values for $\mathbf{O}_{t-1}$ and/or $\mathbf{O}_t$, the identifying equations are given by

$$\begin{pmatrix} \sigma^2_{u_j(t)} & \sigma^2_{u_{j'}(t)} & 2\sigma_{u_1u_2(t)} \\ \gamma_{u_j(t)} & \gamma_{u_ju_{j'}^2(t)} & 2\gamma_{u_j^2u_{j'}(t)} \\ \gamma_{u_j^2u_{j'}(t)} & \gamma_{u_{j'}(t)} & 2\gamma_{u_ju_{j'}^2(t)} \end{pmatrix} \begin{pmatrix} \mathbb{E}((1+\eta_{h_j,w_j(i,t)})^2) \\ \mathbb{E}(\eta^2_{h_j,w_{j'}(i,t)}) \\ \mathbb{E}((1+\eta_{h_j,w_j(i,t)})\eta_{h_j,w_{j'}(i,t)}) \end{pmatrix} = - \begin{pmatrix} m_{y_jy_j(t)} \\ m_{w_jy_j^2(t)} \\ m_{w_{j'}y_j^2(t)} \end{pmatrix}.$$

Frisch symmetry imposes a restriction between $\mathrm{Var}(\eta_{h_1,w_2(i,t)})$ and $\mathrm{Var}(\eta_{h_2,w_1(i,t)})$.

The second cross-moments of consumption and hours elasticities (upper middle and right rectangles in table C.1) are identified as follows. Given values for $\mathbf{O}_{t-1}$ and/or $\mathbf{O}_t$, consider the linear system

$$\begin{pmatrix} \sigma^2_{u_j(t)} & \sigma^2_{u_{j'}(t)} & \sigma_{u_1u_2(t)} & \sigma_{u_1u_2(t)} \\ \gamma_{u_j(t)} & \gamma_{u_ju_{j'}^2(t)} & \gamma_{u_j^2u_{j'}(t)} & \gamma_{u_j^2u_{j'}(t)} \\ \gamma_{u_j^2u_{j'}(t)} & \gamma_{u_{j'}(t)} & \gamma_{u_ju_{j'}^2(t)} & \gamma_{u_ju_{j'}^2(t)} \end{pmatrix} \begin{pmatrix} \mathbb{E}(\eta_{c,w_j(i,t)}(1+\eta_{h_j,w_j(i,t)})) \\ \mathbb{E}(\eta_{c,w_{j'}(i,t)}\eta_{h_j,w_{j'}(i,t)}) \\ \mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_j,w_{j'}(i,t)}) \\ \mathbb{E}(\eta_{c,w_{j'}(i,t)}(1+\eta_{h_j,w_j(i,t)})) \end{pmatrix} = - \begin{pmatrix} m_{y_jc(t)} \\ m_{w_jy_jc(t)} \\ m_{w_{j'}y_jc(t)} \end{pmatrix}$$

repeated twice for $j = \{1,2\}$ while $j' = \{1,2\} \neq j$. This yields 6 equations in 8 parameters. In addition, symmetry of the matrix of Frisch substitution effects provides two linear restrictions $\mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_2,w_1(i,t)}) = \mathbb{E}(\eta_{c,w_j(i,t)}\eta_{h_1,w_2(i,t)})\mathbb{E}(Y_{1it}/Y_{2it})$, $j = \{1,2\}$; taken together these 8 equations just identify the parameters of interest. In practice the parameters are over-identified by at least as many additional equations. Finally, the second cross-moments of male and female labor supply elasticities (middle right rectangle in table C.1) are identified in a similar manner.

Higher moments. Identification of higher moments obeys a similar idea: consumption skewness reflects skewness in wage shocks as well as skewness in consumption preferences; earnings kurtosis reflects kurtosis in the distribution of shocks as well as in labor supply preferences, etc. Practically, however, identification is less parsimonious due to the ever-increasing number of parameters involved. As an illustration, the four skewness parameters for the $\eta_{c,w_j(i)}$'s are over-identified by $\mathbb{E}\big((\Delta c_{it})^2\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t\big)$, $\mathbb{E}\big(\Delta w_{jit-1}(\Delta c_{it})^2\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t\big)$ and $\mathbb{E}\big((\Delta w_{jit-1})^2(\Delta c_{it})^2\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t\big)$, $j=\{1,2\}$. Third and fourth moments alone do not suffice to identify the parameters without restricting co-skewness in the η_{c,w_j}'s. To avoid such restriction, the fifth moment above provides the additional identifying equation that completes identification. All other third moments of wage elasticities are identified in a similar way.

The discussion extends to the n^{th} moment of wage elasticities. This involves an ever-longer intertemporal product of consumption or earnings (for consumption: $\prod_{\tau=0}^{n-3}\Delta c_{it-\tau}\Delta c_{it}\Delta c_{it+1}$, $n\geq 3$) as well as empirical moments of order at least $n+2$. The data requirements quickly become tedious and restrictions on higher cross-moments only partly alleviate such requirements. This is the reason why I focus on the estimation of first and second moments only.

Measurement error. Identification so far assumes measurement error away. If earnings and consumption are contaminated with error, then identification requires moments of the respective earnings and consumption error up to order n. For example, the variance of earnings error enters $m_{y_jy_j(t)}$ additively while the covariance between the error in earnings and the error in wages enters $m_{w_jy_j(t)}$ because wages are constructed as earnings over hours (section 4.2). The variance of consumption error enters $m_{cc(t)}$ also additively. The structure is thus under-identified so one needs external estimates of the error variance or additional restrictions.

The former two moments of the error are available from the validation study of Bound et al. (1994), thus necessitate no additional restrictions. The consumption error variance is not available in the validation study so additional restrictions or an external estimate are needed for identification. Finally, there are no validation studies (at least not for the PSID) with information on higher moments of the earnings and consumption error. This necessitates additional assumptions, e.g. the Gaussian assumption in proposition 1 of section 3.1.

Consumption substitution elasticity. Section 3.3 explains that no moment of the elasticity of consumption with respect to its price is identified without collapsing its distribution to degenerate. While the lifecycle Marshallian elasticity of consumption κ_{c,v_j} identifies a homogeneous $\eta_{c,p}$ in BPS, here any moment of κ_{c,v_j} depends on *various* moments of $\eta_{c,p}$. One can see this by a Taylor expansion of κ_{c,v_1} around mean preferences. With preference heterogeneity, one cannot separate the first from, say, the second moment of $\eta_{c,p}$ and identification of this parameter fails unless one collapses its distribution to degenerate.

To understand why $\mathbb{E}(\kappa_{c,v_1(i,t)})$ depends on various moments of $\eta_{c,p}$, abstract from female labor supply and consider households who, on average, dislike fluctuations in consumption ($\mathbb{E}(\eta_{c,p(i,t)})\rightarrow 0^-$) and male labor supply ($\mathbb{E}(\eta_{h_1,w_1(i,t)})\rightarrow 0^+$). Furthermore, suppose that households *less* reluctant to intertemporal consumption fluctuations ($\eta_{c,p}$ more negative) also have *less* elastic labor supply (low η_{h_1,w_1}; correlates positively with $\eta_{c,p}$). Because those who barely use labor supply to smooth shocks (smallest η_{h_1,w_1}) are also those who do not resent passing such shocks into consumption (largest absolute $\eta_{c,p}$), average consumption across households is more responsive to permanent shocks (higher Marshallian elasticity) than without the preference correlation.[1] The Marshallian elasticity thus depends not only on $\mathbb{E}(\eta_{c,p(i,t)})$ but also on the covariance with the male labor supply elasticity (and in fact on the covariances with all other elasticities that characterize household preference). Ignoring the correlation results in overestimating $|\mathbb{E}(\eta_{c,p(i)})|$ and mistakenly deeming the average household more willing to trade consumption intertemporally.

D Data and Estimation Details

This appendix provides details on the PSID sample and the estimation of wealth shares s_{jit} and π_{it} . In addition, it reports a number of robustness checks as well as the empirical and fitted values of all moments targeted in the structural estimation.

Sample summary statistics. Table D.1 presents descriptive statistics for the baseline sample (columns 1-3) and for four subsamples of relatively wealthy households (columns W1-W4). In the baseline sample, average earnings of men are 86% higher than average earnings of women; men, however, work on average 524 hours more, almost a third more than women. Women are 71% likely to have had some college education; men are slightly less likely. Household consumption is, on average, a fraction of men's earnings but greater than women's. The largest single component within the consumption basket is housing, followed by vehicles (including gasoline) and food at home. On average, there is one child under 18 in the household.

First stage regressions on past outcome levels. The relationship between growth rates at t and outcome levels at $t-1$ is interesting in its own right, and table D.2 presents the results. Own past levels are negative and highly significant while those of the other variables are not.

Estimation of $\boldsymbol{s_{it}}$ and $\boldsymbol{\pi_{it}}$. Wealth shares $\mathbf{s}_{it} = (s_{1it}, s_{2it})'$ and π_{it} are used in the replication of BPS and the simulation of consumption growth after the moments of wage elasticities are estimated. The simulation of Δc_{it} is used in the estimation of $\eta_{c,p}$ and the discussion of inequality in section 6.1 and partial insurance in section 6.2. From appendix A, $s_{jit} \approx \overline{\overline{Y}}_{jit}/\overline{\overline{Y}}_{it}$ and $\pi_{it} \approx \text{Assets}_{it}/(\text{Assets}_{it} + \overline{\overline{Y}}_{it})$, where $\overline{\overline{Y}}_{jit} = Y_{jit} + \mathbb{E}_t \sum_{\varsigma=1}^{T} \frac{Y_{jit+\varsigma}}{(1+r)^{\varsigma}}$ is spouse j's *human wealth* at the beginning of t, namely the expected discounted stream of earnings Y_{jit} between t and the end of working life. $\overline{\overline{Y}}_{it} = \sum_j \overline{\overline{Y}}_{jit}$ is the sum of human wealth in the household.

The main difficulty in estimating s_{jit} is that human wealth conforms to expectations through $\mathbb{E}_t Y_{jit+\varsigma}$. I follow BPS and estimate this as follows. I pool earnings across all periods and

[1] Suppose $\mathbb{E}(\eta_{c,p(i,t)}) = -0.05$ and $\mathbb{E}(\eta_{h_j,w_j(i,t)}) = 0.2$. At these values $\mathbb{E}(\kappa_{c,v_1(i,t)})$ increases more than 1-to-1 with a positive correlation in preferences ($\partial\mathbb{E}(\kappa_{c,v_1(i,t)})/\partial\text{Cov}(\eta_{c,p(i,t)}, \eta_{h_1,w_1(i,t)}) = 1.28$).

Table D.1 – Sample Descriptive Statistics

	baseline sample			$A > \overline{C}$	$A > 2\overline{C}$	$A > \overline{C}$ no debt	$A > \overline{C}$ liquid
	(1) mean	(2) median	(3) st.dev	(W1) mean	(W2) mean	(W3) mean	(W4) mean
Male earner							
Earnings	77466	58181	121938	86403	92394	93570	104818
Hours of work	2254	2205	614	2260	2266	2232	2256
Hourly wage	35.4	26.3	72.9	39.6	42.4	41.2	45.7
Age	45.1	45.0	7.8	47.0	47.4	47.4	47.9
Some college %	67.3	100.0	46.9	71.8	75.4	70.9	76.5
Female earner							
Earnings	41663	34612	36166	44807	46442	44607	48333
Hours of work	1730	1880	667	1729	1719	1681	1675
Hourly wage	24.5	19.6	29.6	26.4	27.6	27.1	29.7
Age	43.4	43.0	7.6	45.2	45.7	45.6	46.1
Some college %	70.8	100.0	45.5	73.7	76.1	72.8	76.7
Household consumption							
Total consumption	47307	40850	26373	51246	54186	49618	53324
food at home	7102	6606	3550	7230	7354	6937	7056
food out	3012	2400	2680	3224	3364	3115	3353
vehicles[a]	7136	5298	6653	7400	7579	6668	6841
public transport	347	0	2628	416	473	372	429
childcare	970	0	3228	902	898	793	840
education	3517	0	9538	4371	4878	3848	4402
medical expenses[b]	3657	2547	4266	3799	3925	3543	3825
utilities	4835	4436	2761	4895	4956	4473	4576
housing[c]	16731	12980	13924	19008	20758	19870	22002
Household assets and debt [in thousands]							
Total wealth	408.5	180.7	878.6	539.5	631.1	697.4	842.3
Home equity[d]	131.5	80.8	179.9	171.4	195.8	206.6	234.2
Other debt	12.3	2.7	29.1	9.2	9.0	0.1	0.1
All other assets	289.3	84.3	791.5	377.3	444.4	490.9	608.2
other real estate	46.0	0.0	302.0	59.5	71.2	71.7	84.7
savings accounts	27.2	7.1	71.6	34.6	40.0	51.0	62.7
stocks-shares	46.7	0.0	198.4	62.1	74.2	93.6	119.8
# of children	1.0	1.0	1.1	0.9	0.9	0.9	0.8
Obs. [households × years]		8308		5718	4715	2391	1849

Notes: The table presents summary statistics for the baseline sample (columns 1-3) and for four subsamples of relatively wealthy households (columns W1-W4) over waves 1999-2011 of the PSID. All monetary amounts are in 2010 dollars. Earnings, hours, consumption and wealth/debt are annual. Column W1 is for households with wealth A_t at least as much as average consumption $\overline{C}_t$ in the baseline sample. Column W2 is for households with wealth at least twice as much as average consumption. Column W3 is like column W1 with the additional condition that households hold real debt that does not exceed \$2K. Column W4 is like column W3 but the relevant measure of wealth excludes home equity, thus proxies for liquid assets. [a]including gasoline; [b]including health insurance and prescriptions; [c]including home insurance, rent and rent equivalent for homeowners and people in alternative housing arrangements; [d]the present value of owned house net of outstanding mortgages.

Table D.2 – First Stage Regressions: Past Levels

	baseline sample		
	$\Delta \ln Y_{1it}$	$\Delta \ln Y_{2it}$	$\Delta \ln C_{it}$
H_{1it-1}	-9.30e-05 (-7.79)	1.56e-05 (1.16)	3.59e-06 (0.59)
H_{2it-1}	1.43e-05 (1.29)	-2.03e-04 (-16.27)	-5.29e-06 (-0.93)
C_{it-1}	-4.78e-07 (-1.53)	7.64e-09 (0.02)	-4.21e-06 (-26.31)
observables	yes	yes	yes
Obs. [households $\times \Delta_t$]		6071	

Notes: The table reports the coefficients from first stage regressions of earnings and consumption growth on past levels of hours and consumption. t-stats appear in brackets under the coefficients. Sample sizes are smaller than in summary statistics table D.1 because the variables here are in first differences.

regress them on a set of predictable characteristics including a cubic polynomial in age, year of birth, race and education dummies, as well as interactions of the polynomial with the race and education dummies. I summarize this regression as $Y_{jit} = \mathbf{Q}'_{jit}\boldsymbol{\delta}_j + \epsilon_{jit}$. I then obtain $\mathbb{E}_t Y_{jit+\varsigma}$ as the appropriate fitted value from this regression, i.e. $\mathbb{E}_t Y_{jit+\varsigma} = \mathbf{Q}'_{jit+\varsigma}\hat{\boldsymbol{\delta}}_j$. I set the discount rate at 2% annually and the end of working life at 65. This allows me to construct s_{jit}, $j = \{1, 2\}$, and π_{it}. Note that assets/wealth are directly observed in the PSID. Table D.3 presents summary statistics. I estimate $\mathbb{E}(s_{1it}) = 0.614$ and $\mathbb{E}(\pi_{it}) = 0.191$. The averages as well as the patterns of $\mathbf{s}_{it}$ and π_{it} by age are similar to BPS.

Targeted moments. The structural estimation targets 80 moments of the joint distribution of residual wages, earnings, and consumption. Tables D.4-D.6 list the targeted conditional moments alongside their empirical values at the sample average of consumption and hours (equivalent to the unconditional moments) and their fitted values in the preferred specification. Block bootstrap standard errors for the empirical moments are in parentheses based on 1,000 replications. The t-statistics are for the null hypothesis that the theoretical moment equals the empirical one. More than 90% of targeted moments are associated with $|t\text{-stat}|$ lower than the rule of thumb of 1.96. The magnitude and standard error of most of the 7 moments for which $|t\text{-stat}| > 1.96$ are very small making small and economically unimportant departures from the target easier to generate large t-statistics.

Consumption substitution elasticity. While estimation of the wage elasticities does not require parametric form restrictions on preferences and their distribution, estimation of $\eta_{c,p}$ is not possible without strong distributional assumptions (section 3.3). Such assumptions are

Table D.3 – Summary Statistics for Wealth Shares $\mathbf{s}$ and π

s_{1it}					π_{it}				
mean	med.	st.d.	min	max	mean	med.	st.d.	min	max
0.614	0.620	0.094	0.144	0.996	0.191	0.133	0.182	0.000	0.956

Notes: The table presents summary statistics for men's share of human wealth (s_{1it}) and for the partial insurance coefficient (π_{it}) in the baseline sample. Women's share of human wealth is $s_{2it} = 1 - s_{1it}$.

not needed when preferences are homogeneous so BPS estimate $\eta_{c,p}$ together with the wage elasticities. This is not possible here unless one restricts the distribution of preferences to a specific class. To avoid subjecting *all* parameters to such restriction, I estimate $\eta_{c,p}$ in a second stage, i.e. after having estimated the wage elasticities *without* distributional restrictions.

I use the variance of consumption growth $\text{Var}(\Delta c_{it})$ at the sample average of consumption and hours (i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$) to estimate a homogeneous $\eta_{c,p}$; this variance involves moments of the Marshallian elasticity $\kappa_{c,v_j(i,t)}$ that was shown in section 3.3 to depend on $\eta_{c,p}$. The Marshallian elasticity also depends on the *entire* distribution of wage elasticities thus necessitating assumptions on their joint distribution. A natural benchmark is to restrict that to joint normal and parameterize it at the first and second moments from the preferred specification. Normality does not contradict the previous use of third moments as consumption and hours can still exhibit skewness because wage shocks are skewed. While a full parametric approach would subject all parameters to distributional assumptions, this two-step approach only subjects $\eta_{c,p}$. Estimation of $\eta_{c,p}$ is thus conditional on (rather than joint with) the wage elasticities.

I operationalize this estimation using simulated minimum distance as $\text{Var}(\Delta c_{it})$ depends implicitly and nonlinearly on the parameters. Specifically, I simulate preferences for 10 million households, I calculate the variance of consumption growth across them, and I estimate $\eta_{c,p}$ minimizing the quadratic distance between the simulated variance and its empirical counterpart. This estimation requires information on wealth shares π_{it} and $\mathbf{s}_{it}$ for which I draw random values from their empirical distributions. The distance metric is minimized at $\eta_{c,p} = -0.598$ (simulated and empirical variances matched to the third decimal digit) implying a coefficient of relative risk aversion approximately equal to 1.67.[2]

[2]The coefficient of relative risk aversion is approximately equal to $-\eta_{c,p}^{-1}$ (exactly equal if preferences are separable). Chetty (2006) uses labor supply data and calculates an upper bound on this parameter at 0.97 when consumption and hours are complements. Abstracting from labor supply, Kimball et al. (2009) impute the coefficient of relative risk aversion from hypothetical gamble responses in the PSID and report a range of 1.4-6.7. Guiso and Sodini (2013) calculate household risk aversion based on portfolio risk shares in the US Survey of Consumer Finances. They report a median coefficient at 3.5 with the central 90% of the distribution lying in the range 1.6-30.8 skewed to the left. Cohen and Einav (2007) estimate risk preferences from deductible choices in the auto insurance market and find a median coefficient at 0.37 with the average being much higher.

Table D.4 – Targeted Wage Moments

	data	model	t-stat. diff.		data	model	t-stat. diff.
$\mathbb{E}\big((\Delta w_{1t})^2\big)$	0.129 (0.013)	0.129	0.000	$\mathbb{E}\big((\Delta w_{1t})^2\Delta w_{2t}\big)$	0.010 (0.005)	0.010	0.000
$\mathbb{E}\big(\Delta w_{1t}\Delta w_{1t^+}\big)$	-0.024 (0.006)	-0.024	-0.001	$\mathbb{E}\big((\Delta w_{1t})^2\Delta w_{2t^+}\big)$	-0.012 (0.014)	0.000	0.866
$\mathbb{E}\big((\Delta w_{2t})^2\big)$	0.098 (0.008)	0.098	0.000	$\mathbb{E}\big((\Delta w_{1t^+})^2\Delta w_{2t}\big)$	-0.006 (0.004)	0.000	1.663
$\mathbb{E}\big(\Delta w_{2t}\Delta w_{2t^+}\big)$	-0.012 (0.005)	-0.012	-0.002	$\mathbb{E}\big(\Delta w_{1t}(\Delta w_{2t})^2\big)$	0.001 (0.003)	0.001	0.000
$\mathbb{E}\big(\Delta w_{1t}\Delta w_{2t}\big)$	0.017 (0.003)	0.017	0.000	$\mathbb{E}\big(\Delta w_{1t}(\Delta w_{2t^+})^2\big)$	0.003 (0.003)	0.000	-1.067
$\mathbb{E}\big(\Delta w_{1t}\Delta w_{2t^+}\big)$	-0.003 (0.003)	-0.004	-0.094	$\mathbb{E}\big(\Delta w_{1t^+}(\Delta w_{2t})^2\big)$	0.003 (0.003)	0.000	-1.148
$\mathbb{E}\big(\Delta w_{1t^+}\Delta w_{2t}\big)$	-0.004 (0.003)	-0.004	0.105	$\mathbb{E}\big(\Delta w_{1t}\Delta w_{1t^+}\Delta w_{2t}\big)$	0.003 (0.002)	0.000	-1.443
$\mathbb{E}\big((\Delta w_{1t})^3\big)$	0.044 (0.052)	0.044	0.000	$\mathbb{E}\big(\Delta w_{1t}\Delta w_{1t^+}\Delta w_{2t^+}\big)$	-0.003 (0.002)	0.000	1.116
$\mathbb{E}\big(\Delta w_{1t}(\Delta w_{1t^+})^2\big)$	-0.009 (0.008)	-0.018	-1.112	$\mathbb{E}\big(\Delta w_{1t}\Delta w_{2t}\Delta w_{2t^+}\big)$	-0.001 (0.002)	0.000	0.271
$\mathbb{E}\big(\Delta w_{1t^+}(\Delta w_{1t})^2\big)$	0.027 (0.011)	0.018	-0.852	$\mathbb{E}\big(\Delta w_{1t^+}\Delta w_{2t}\Delta w_{2t^+}\big)$	0.000 (0.002)	0.000	0.344
$\mathbb{E}\big((\Delta w_{2t})^3\big)$	-0.029 (0.011)	-0.029	0.000				
$\mathbb{E}\big(\Delta w_{2t}(\Delta w_{2t^+})^2\big)$	-0.003 (0.006)	-0.008	-0.741				
$\mathbb{E}\big(\Delta w_{2t^+}(\Delta w_{2t})^2\big)$	0.012 (0.006)	0.008	-0.674				

Notes: The notation is as follows: $t^+ \equiv t+1$. The table presents the list of targeted wage moments alongside their empirical and theoretical values. Empirical moments are net of observables and wage measurement error. Block bootstrap standard errors for the empirical moments are in parentheses based on 1,000 bootstrap replications. The reported t-statistic is for the null hypothesis that *theoretical moment* $-$ *empirical moment* $= 0$.

Table D.5 – Targeted Earnings Moments at Sample Average of Consumption and Hours

	data	model	t-stat. diff.		data	model	t-stat. diff.
$\mathbb{E}(\Delta w_{1t}\Delta y_{1t^+}\|\overline{\mathbf{o}}_t)$	-0.036 (0.005)	-0.030	1.068	$\mathbb{E}((\Delta w_{1t})^2\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	-0.001 (0.005)	0.000	0.223
$\mathbb{E}(\Delta w_{1t^+}\Delta y_{1t}\|\overline{\mathbf{o}}_{t^-})$	-0.035 (0.006)	-0.030	0.835	$\mathbb{E}((\Delta w_{1t^+})^2\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	-0.001 (0.005)	0.000	0.165
$\mathbb{E}(\Delta w_{2t}\Delta y_{1t^+}\|\overline{\mathbf{o}}_t)$	-0.003 (0.003)	-0.005	-0.446	$\mathbb{E}((\Delta w_{2t})^2\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	0.009 (0.006)	0.009	0.010
$\mathbb{E}(\Delta w_{2t^+}\Delta y_{1t}\|\overline{\mathbf{o}}_{t^-})$	-0.002 (0.003)	-0.005	-0.864	$\mathbb{E}((\Delta w_{2t^+})^2\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	-0.001 (0.007)	-0.009	-1.239
$\mathbb{E}(\Delta w_{1t}\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	-0.004 (0.003)	-0.005	-0.240	$\mathbb{E}(\Delta w_{1t}(\Delta y_{1t^+})^2\|\overline{\mathbf{o}}_t)$	-0.021 (0.006)	-0.030	-1.445
$\mathbb{E}(\Delta w_{1t^+}\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	0.002 (0.003)	-0.005	-1.863	$\mathbb{E}(\Delta w_{1t^+}(\Delta y_{1t})^2\|\overline{\mathbf{o}}_{t^-})$	0.026 (0.008)	0.030	0.476
$\mathbb{E}(\Delta w_{2t}\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	-0.040 (0.005)	-0.015	4.823	$\mathbb{E}(\Delta w_{2t}(\Delta y_{1t^+})^2\|\overline{\mathbf{o}}_t)$	-0.005 (0.003)	0.000	1.550
$\mathbb{E}(\Delta w_{2t^+}\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	-0.028 (0.006)	-0.015	2.107	$\mathbb{E}(\Delta w_{2t^+}(\Delta y_{1t})^2\|\overline{\mathbf{o}}_{t^-})$	0.004 (0.003)	0.000	-1.226
$\mathbb{E}(\Delta y_{1t}\Delta y_{1t^+}\|\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t)$	-0.044 (0.006)	-0.039	0.948	$\mathbb{E}(\Delta w_{1t}(\Delta y_{2t^+})^2\|\overline{\mathbf{o}}_t)$	0.006 (0.004)	0.001	-1.431
$\mathbb{E}(\Delta y_{1t}\Delta y_{2t^+}\|\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t)$	-0.008 (0.003)	-0.006	0.616	$\mathbb{E}(\Delta w_{1t^+}(\Delta y_{2t})^2\|\overline{\mathbf{o}}_{t^-})$	-0.001 (0.005)	-0.001	0.038
$\mathbb{E}(\Delta y_{1t^+}\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t)$	0.000 (0.004)	-0.006	-1.510	$\mathbb{E}(\Delta w_{2t}(\Delta y_{2t^+})^2\|\overline{\mathbf{o}}_t)$	-0.006 (0.006)	-0.012	-0.967
$\mathbb{E}(\Delta y_{2t}\Delta y_{2t^+}\|\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t)$	-0.012 (0.006)	-0.018	-1.117	$\mathbb{E}(\Delta w_{2t^+}(\Delta y_{2t})^2\|\overline{\mathbf{o}}_{t^-})$	0.010 (0.007)	0.012	0.282
$\mathbb{E}((\Delta w_{1t})^2\Delta y_{1t^+}\|\overline{\mathbf{o}}_t)$	0.031 (0.014)	0.023	-0.571	$\mathbb{E}(\Delta w_{1t}\Delta y_{1t^+}\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	-0.005 (0.002)	0.000	2.311
$\mathbb{E}((\Delta w_{1t^+})^2\Delta y_{1t}\|\overline{\mathbf{o}}_{t^-})$	-0.017 (0.008)	-0.023	-0.755	$\mathbb{E}(\Delta w_{1t^+}\Delta y_{1t}\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	-0.001 (0.003)	0.000	0.534
$\mathbb{E}((\Delta w_{2t})^2\Delta y_{1t^+}\|\overline{\mathbf{o}}_t)$	0.000 (0.003)	-0.001	-0.107	$\mathbb{E}(\Delta w_{2t}\Delta y_{1t^+}\Delta y_{2t^+}\|\overline{\mathbf{o}}_t)$	-0.001 (0.002)	0.001	1.046
$\mathbb{E}((\Delta w_{2t^+})^2\Delta y_{1t}\|\overline{\mathbf{o}}_{t^-})$	0.003 (0.003)	0.001	-0.925	$\mathbb{E}(\Delta w_{2t^+}\Delta y_{1t}\Delta y_{2t}\|\overline{\mathbf{o}}_{t^-})$	0.001 (0.002)	-0.001	-0.993

Notes: The notation is as follows: $t^- \equiv t-1$, $t^+ \equiv t+1$ and $\overline{\mathbf{o}}_{t^-} \equiv (\mathbf{O}_{t-1} = \mathbb{E}(\mathbf{O}_{it-1}))$, $\overline{\mathbf{o}}_t \equiv (\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it}))$ indicate conditioning on sample averages of consumption and hours. The table presents the list of targeted earnings moments alongside their empirical and theoretical values from the preferred specification. Empirical moments are net of observables and measurement error. Block bootstrap standard errors for the empirical moments are in parentheses based on 1,000 bootstrap replications. The reported t-statistic is for the null hypothesis that *theoretical moment* $-$ *empirical moment* $= 0$.

Table D.6 – Targeted Consumption Moments at Sample Average of Consumption and Hours

	data	model	t-stat. diff.		data	model	t-stat. diff.
$\mathbb{E}\big(\Delta w_{1t}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.000 (0.002)	0.001	0.759	$\mathbb{E}\big(\Delta w_{1t}(\Delta c_{t^+})^2\vert\overline{\mathbf{o}}_t\big)$	0.000 (0.001)	-0.002	-3.208
$\mathbb{E}\big(\Delta w_{1t^+}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.001 (0.002)	0.001	0.213	$\mathbb{E}\big(\Delta w_{1t^+}(\Delta c_t)^2\vert\overline{\mathbf{o}}_{t^-}\big)$	0.000 (0.001)	0.002	2.460
$\mathbb{E}\big(\Delta w_{2t}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.000 (0.002)	0.001	0.776	$\mathbb{E}\big(\Delta w_{2t}(\Delta c_{t^+})^2\vert\overline{\mathbf{o}}_t\big)$	0.000 (0.001)	-0.002	-3.722
$\mathbb{E}\big(\Delta w_{2t^+}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.002 (0.002)	0.001	-0.545	$\mathbb{E}\big(\Delta w_{2t^+}(\Delta c_t)^2\vert\overline{\mathbf{o}}_{t^-}\big)$	0.000 (0.001)	0.002	2.884
$\mathbb{E}\big(\Delta y_{1t}\Delta c_{t^+}\vert\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t\big)$	-0.003 (0.002)	0.000	1.713	$\mathbb{E}\big(\Delta w_{1t}\Delta y_{1t^+}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	-0.002 (0.001)	0.000	1.475
$\mathbb{E}\big(\Delta y_{1t^+}\Delta c_t\vert\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t\big)$	0.002 (0.002)	0.000	-1.064	$\mathbb{E}\big(\Delta w_{1t^+}\Delta y_{1t}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.000 (0.001)	0.000	0.011
$\mathbb{E}\big(\Delta y_{2t}\Delta c_{t^+}\vert\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t\big)$	0.002 (0.002)	0.001	-0.496	$\mathbb{E}\big(\Delta w_{2t}\Delta y_{1t^+}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.001 (0.001)	0.000	-0.996
$\mathbb{E}\big(\Delta y_{2t^+}\Delta c_t\vert\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t\big)$	0.001 (0.002)	0.001	0.014	$\mathbb{E}\big(\Delta w_{2t^+}\Delta y_{1t}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	-0.001 (0.001)	0.000	0.863
$\mathbb{E}\big(\Delta c_t\Delta c_{t^+}\vert\overline{\mathbf{o}}_{t^-},\overline{\mathbf{o}}_t\big)$	-0.008 (0.001)	-0.006	1.764	$\mathbb{E}\big(\Delta w_{1t}\Delta y_{2t^+}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.001 (0.001)	0.000	-1.485
$\mathbb{E}\big((\Delta w_{1t})^2\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.000 (0.006)	-0.001	-0.028	$\mathbb{E}\big(\Delta w_{1t^+}\Delta y_{2t}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.002 (0.001)	0.000	-1.680
$\mathbb{E}\big((\Delta w_{1t^+})^2\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.009 (0.004)	0.001	-1.928	$\mathbb{E}\big(\Delta w_{2t}\Delta y_{2t^+}\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.001 (0.001)	0.001	-0.238
$\mathbb{E}\big((\Delta w_{2t})^2\Delta c_{t^+}\vert\overline{\mathbf{o}}_t\big)$	0.001 (0.002)	0.000	-0.911	$\mathbb{E}\big(\Delta w_{2t^+}\Delta y_{2t}\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	-0.001 (0.001)	-0.001	0.165
$\mathbb{E}\big((\Delta w_{2t^+})^2\Delta c_t\vert\overline{\mathbf{o}}_{t^-}\big)$	0.000 (0.002)	0.000	0.247				

Notes: The notation is as follows: $t^- \equiv t-1$, $t^+ \equiv t+1$ and $\overline{\mathbf{o}}_{t^-} \equiv (\mathbf{O}_{t-1} = \mathbb{E}(\mathbf{O}_{it-1}))$, $\overline{\mathbf{o}}_t \equiv (\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it}))$ indicate conditioning on sample averages of consumption and hours. The table presents the list of targeted consumption moments alongside their empirical and theoretical values from the preferred specification. Empirical moments are net of observables and measurement error. Block bootstrap standard errors for the empirical moments are in parentheses based on 1,000 bootstrap replications. The reported t-statistic is for the null hypothesis that *theoretical moment − empirical moment* = 0.

E Additional Results and Robustness

Alternative 'preferred' specifications. Table E.1 presents alternative 'preferred' specifications, namely results from a number of restrictions that are alternative to the mild homogeneity assumption in the preferred specification. I start from column 7 of tables 4-5 and seek restrictions that improve efficiency of the parameters. Column 1 shuts down all heterogeneity in labor supply elasticities; column 2 shuts down joint heterogeneity in labor supply elasticities; column 3 posits $\eta_{c,w_2(i)} = \alpha_i \times \eta_{c,w_1(i)}$ with $\alpha_i = \alpha \; \forall i$ and estimates α together with all other parameters; column 4 is like column 3 but also equalizes the labor supply elasticities between spouses, i.e. $\eta_{h_1,w_1(i)} = \eta_{h_2,w_2(i)}$; column 5 imposes equal consumption-wage elasticities, i.e. $\eta_{c,w_1(i)} = \eta_{c,w_2(i)}$. The parameters across all columns are estimated at the sample average of consumption and hours. The results are similar to the preferred specification of column 8 in tables 4-5, although efficiency varies with the restrictions imposed. Interestingly, when I equalize the labor supply elasticities between spouses, efficiency of the elasticities improves dramatically (as expected) but the pattern of heterogeneity does not change: these elasticities seem to not exhibit economically meaningful heterogeneity across households.

Extreme observations. Table E.2 presents parameter estimates in the preferred specification after removing extreme observations of wages, earnings, consumption. Extreme observations may matter for the third and other moments so it is meaningful to check the robustness of the results to removing those. Column 1 shows the baseline results from tables 4-5. Column 2 trims the bottom and top 0.5% of the distribution of wages. This drops 3.2% of the baseline sample of table D.1 because the condition applies separately to male and female wages and drops additionally households that now appear in the sample only once. Column 3 trims the bottom and top 2% of the distribution of wages (drops 11.8% of the baseline sample) while column 4 trims the bottom and top 0.5% of the distributions of wages, earnings and consumption (drops 5.1% of the baseline sample). The main conclusions remain unchanged. Consumption preference heterogeneity remains substantial and by an order of magnitude greater than heterogeneity in the labor supply elasticities. The largest change from the baseline appears in column 3 where the magnitudes of $\mathrm{Var}(\eta_{c,w_j(i)})$ and $\mathbb{E}(\eta_{h_2,w_2(i)})$ increase considerably.

Additional first-stage controls: chores and children's age. Tables E.3-E.4 replicate the main results, this time controlling for the age of the youngest child and each spouse's past levels of chores in the first stage regressions. Blundell et al. (2018) show that the household response to wage shocks depends on the substitutability of parental time in home production, while Boerma and Karabarbounis (2019) show that heterogeneity in home production may be mistakenly classified as heterogeneity in preferences. The data do not allow me to control for childcare but I can control for the age of the youngest child, as well as for time spent on chores. Accounting for these variables enables me to check the extent to which preference heterogeneity masks heterogeneity in home production, assuming that the latter is either due to differences in children's age or captured by heterogeneity in chores. Unsurprisingly, the results are unchanged. The baseline results already control for parental age *and* the number of children in the household, both of which have large predictive power for the age of the youngest child. The baseline also controls for a long list of observables, thus apparently exhaustingly capturing heterogeneity in home production.

Diagonally weighted GMM. While equally weighted GMM weighs all target moments equally and thus favors numerically large moments, diagonally weighted GMM favors moments that are more precise. It uses the diagonal of the inverse covariance matrix of moments as weighting matrix, thus emphasizing moments whose variance in the data is lowest. Unless the model is misspecified, the two weighting schemes should deliver unbiased estimates of the parameters. In practice, however, possible misspecification or small-sample issues may produce materially different results so it is important to check for this possibility.

Tables E.5-E.6 replicate the main set of results using diagonally weighted GMM.[3] Five observations emerge. First, in replicating BPS in column 2, the average female labor supply elasticity turns negative but insignificant. A negative value contradicts most evidence in the literature (Keane, 2011), suggesting some form of possible misspecification in the moments used by BPS. Second, in targeting moments of transitory shocks in column 3, this elasticity turns again positive and all parameters become remarkably similar to the baseline. Third, adding third moments in column 4 reduces the size of the elasticities exactly as in the baseline; the reduction is in the magnitude of 25%, versus 35% in the baseline, reflecting the *relative* imprecision of third moments. Fourth, introducing marginal preference heterogeneity in column 5 produces conceptually similar results to the baseline: consumption elasticities exhibit substantial heterogeneity (although slightly less pronounced than in the baseline), while labor supply elasticities do not. Fifth, introducing joint heterogeneity (column 6), shutting down the labor supply cross-elasticities (column 7), and imposing a mild homogeneity assumption among the η_{c,w_j}'s results in: average labor supply elasticities drop further, consumption preference heterogeneity remains substantial and significant at the 10% level (column 8), and labor supply heterogeneity rises to become comparable, albeit statistically insignificant, to consumption heterogeneity. This last result is not observed in the baseline.

While consumption preference heterogeneity is consistently large across all specifications, heterogeneity in labor supply preferences fluctuates between zero and amounts comparable to consumption heterogeneity. Labor supply heterogeneity is never statistically significant while consumption heterogeneity is; however, this may be an artifact of the mild homogeneity assumption I impose on the η_{c,w_j}'s but not on the labor supply elasticities. Could a similar restriction on the labor supply elasticities result in statistically significant labor supply hetero-

[3]The set of target moments appears in appendix tables D.4-D.6. I estimate the covariance matrix of moments via block bootstrap based on 1,000 replications.

geneity comparable to consumption heterogeneity? The answer is a firm no. Table E.7 presents results from a number of restrictions alternative to those in the preferred specification, mimicking table E.1 for the baseline results. Columns 4 therein attempts to improve efficiency of labor supply preferences by imposing $\eta_{h_1,w_1(i)} = a_i \times \eta_{h_2,w_2(i)}$ for $a_i = 1, \forall i$. As soon as this is done, $\text{Var}(\eta_{h_1,w_1(i)})$ and $\text{Var}(\eta_{h_2,w_2(i)})$ both drop to a statistically significant 0.001 while consumption preference heterogeneity remains at least an order of magnitude larger.[4] I therefore conclude that the results from diagonally weighted GMM are qualitatively and in most cases quantitatively similar to the baseline. There is substantial heterogeneity in consumption preferences but not in preferences over labor supply; the average labor supply elasticities drop as soon as the model targets third moments of consumption and earnings in the data. At face value, consumption preference heterogeneity in the preferred specification (column 8 of tables E.5-E.6) implies that one st.d. of η_{c,w_1} about its mean falls approximately in the range $(-0.21; 0.17)$ while one st.d. of η_{c,w_2} in $(-0.39; 0.33)$.

Consumption measurement error. The variance of consumption measurement error is identified in certain specifications only after imposing additional restrictions (section 3.2). To avoid building in inconsistencies across specifications (namely impose restrictions in some specifications, e.g. when allowing for preference heterogeneity, but not others, e.g. in replicating BPS), I opt to fix this at 15% of the variance of consumption growth. While this number is higher than error in wages (13%) or earnings (4%) in Bound et al. (1994), it is ultimately arbitrary so I check here how the results change with a different amount of consumption error.

Figure E.1 plots the variances of consumption and labor supply elasticities against the variance of consumption error. Consumption preference heterogeneity remains substantial for amounts of consumption error up to about 40-45% of the variance of consumption growth; heterogeneity in labor supply elasticities remains always negligible. Most other parameters, including $\mathbb{E}(\eta_{c,w_1(i)})$ and $\mathbb{E}(\eta_{c,w_2(i)})$, do not change materially from the baseline. An exception is $\eta_{c,p}$ and figure E.2 illustrates this.

Naturally, a bigger amount of consumption error reduces the scope of preferences in the model especially as the amounts of earnings & wage error do not move in parallel. A larger consumption error reduces the 'true' variance of consumption; the model then gradually attributes a larger share of the consumption variation to wage variation, which is constant throughout the counterfactual of figure E.1. To see this, consider expression (9) for consumption inequality and focus for now on the first part of it: $\text{Var}(\Delta c_{it}) \approx \sum_{j=1}^{2} \mathbb{E}(\eta^2_{c,w_j(i,t-1)}) \times \left(\sigma^2_{u_j(t)} + \sigma^2_{u_j(t-1)}\right)$. As the left hand side become smaller with a larger amount of consumption error but variation in wages (i.e. $\sigma^2_{u_j}$) remains unchanged, the model pushes consumption preference heterogeneity (i.e. $\mathbb{E}(\eta^2_{c,w_j(i)})$) gradually towards zero. This is, however, not true if the amount of wage error

[4]Neither $\text{Var}(\eta_{c,w_1(i)})$ nor $\text{Var}(\eta_{c,w_2(i)})$ is statistically significant in column 4 of table E.1 as the proportionality α_i of the underlying elasticities is not fixed in this specification. As soon as I fix α_i, consumption preference heterogeneity becomes statistically significant exactly as in the preferred and other specifications.

increases proportionally to the amount of consumption error, thus reducing $\sigma^2_{u_j}$ in parallel to $\text{Var}(\Delta c_{it})$. In such case, consumption preference heterogeneity retains its relative importance; moreover, for most values of consumption, wage, and earnings error, heterogeneity in labor supply preferences remains negligible (related figures not shown for brevity). For completeness, table E.8 presents the decomposition of consumption inequality and figure E.5 visualizes the distributions of pass-through rates of shocks for different amounts of consumption error.

Progressive joint taxation. With progressive joint taxation in the family, the dynamics of consumption and hours are different from baseline expressions (5)-(7). Each equation is augmented by an additional term that is a function of *all* shocks and reflects the disincentives joint taxation induces on family labor supply. Without taxation, a transitory shock shifts own labor supply reflecting the substitution between hours and leisure, i.e. the Frisch elasticity of labor supply. With joint taxation, the labor supply response is mitigated because part of the shock is taxed or compensated. This depends *nonlinearly* on the tax system progressivity, the partner's hours response to the tax (earnings are taxed jointly so the tax disincentivizes *family* labor supply), and the relative shares of spouses' earnings. Consequently, the response of labor supply (or of consumption) to the transitory shock no longer picks up a single Frisch elasticity but, instead, a plethora of elasticities and tax parameters. BPS derive analytical expressions under joint taxation and preference homogeneity, and I refer to them for the details. The extension of their derivations to heterogeneous preferences is straightforward.

With joint taxation, the intertemporal moments of consumption, earnings, and wages that have so far identified preference heterogeneity are no longer *linear* functions of moments of the elasticities. The reason is that the response to transitory shocks, which the said moments reflect, is a nonlinear function of elasticities and tax parameters as explained above. For example, the transmission parameter of Δu_{1it} into Δc_{it} ($\eta_{c,w_1(i,t-1)}$ in the baseline) now becomes $(\eta_{c,w_1(i,t-1)} - \nu_{it} s_{1it} \sum_j \eta_{c,w_j(i,t-1)})/(1 + \nu_{it} s_{1it} \sum_j \eta_{h_1,w_j(i,t-1)} + \nu_{it} s_{2it} \sum_j \eta_{h_2,w_j(i,t-1)})$ where ν_{it} determines the progressivity of the tax system. The denominator reflects the labor supply disincentives of joint taxation while the extra term in the numerator reflects the feedback on consumption due to non-separability with labor supply. Consequently, identification of heterogeneity can no longer be shown analytically, while estimation requires parametric restrictions on the distribution of preferences. Given that I do not want to a priori restrict preferences to a specific distribution, my baseline model so far has been one without joint taxation.

To understand the implications of joint taxation for preference heterogeneity, I draw preferences for 10 million households from the multivariate normal distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and I construct all second and third moments that I fit in the baseline. I fix $\nu_{it} = 0.185$ as in Heathcote et al. (2014) and I draw human wealth shares from the empirical distribution of $\mathbf{s}_{it}$; the proportionality parameter χ_{it} drops out because of the nature of the approximation. I match these moments to their empirical counterparts keeping all estimation details as in the baseline; now $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the parameters of interest. Results from the preferred specification (column 1, table E.9) show that joint taxation makes the *average* labor supply elasticities a bit larger, exactly as in BPS, but does not change the baseline pattern of preference heterogeneity. Results for the unrestricted model specifications produce similar conclusions.

Three remarks are due here. First, the results keep ν_{it} fixed; a more detailed treatment should have this vary across households. However, any such variation is justified on the basis of varying demographics and earnings/hours levels, both of which I control for empirically. Second, any neglected heterogeneity in ν_{it} would also be picked up by $\text{Var}(\eta_{h_j,w_j(i)})$ which serves as an upper bound to $\text{Var}(\nu_{it})$; see Theloudis (2017a). As $\text{Var}(\eta_{h_j,w_j(i)})$ is close to zero, conditional heterogeneity in ν is practically small. Third, joint taxation may produce different results from the baseline *also* due to the normality assumption. I reestimate the baseline model (i.e. without taxation) imposing joint normality and I find that the results are not very different from the baseline.[5] Joint normality of preferences seems not too strong an assumption *ex post*.

Liquidity constraints and adjustment costs of work. To understand the implications of consumption equation (10), it helps to focus on the transmission of women's transitory shock u_2 and abstract from male wages as if $q_{2it-1} = 1$.

- The average loading factor of u_2 into consumption, captured by $\mathbb{E}(\Delta w_{2it}\Delta c_{it+1}|\mathbf{O}_t)$, identifies $\mathbb{E}(\eta_{c,w_2(i,t)})$ in the baseline but now reflects the degree of liquidity tightness among constrained households. Pooling both types of households together, the average transmission parameter identifies $\varrho_t\mathbb{E}(\eta_{c,w_2(i,t)}) + (1-\varrho_t)\mathbb{E}(\theta_{2it})$, thus biases $\mathbb{E}(\eta_{c,w_2(i,t)})$ upwards if consumption and hours are Frisch substitutes, as in BPS, and the measure of tightness positive (in principle $\theta_j \in [0,1]$ when labor supply is fixed).

- The second moment of this loading factor, captured by $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t)$, identifies $\mathbb{E}(\eta^2_{c,w_2(i,t)})$ in the baseline but $\mathbb{E}(\theta^2_{2it})$ among the constrained. Pooling the two together, and assuming households do not switch status from $t-1$ to $t+1$, this moment identifies $\varrho_t\mathbb{E}(\eta^2_{c,w_2(i,t)})+(1-\varrho_t)\mathbb{E}(\theta^2_{2it})$ and picks up heterogeneity in η_{c,w_2} and variability in liquidity tightness. This biases $\text{Var}(\eta_{c,w_2(i,t)})$ upwards overstating true preference heterogeneity.[6]

Similar expressions can be obtained for all other moments and transmission parameters (though not all offer testable conditions). For example, $\mathbb{E}(\Delta w_{2it}\Delta y_{2it+1}|\mathbf{O}_t)$identifies $\varrho_t\mathbb{E}(\eta_{h_2,w_2(i,t)})$ and understates the true labor supply elasticity. This last point is an implication of fixed labor supply and does not hold if labor supply is flexible.

These implications are testable among *unconstrained* households. Assuming household membership in each group is fixed over time: (1.) average consumption elasticities are smaller (more negative) than in the pool of constrained and unconstrained; (2.) the variance of the consumption elasticities is lower; (3.) the average labor supply elasticities are larger. These

[5]These results, and the results with taxation for all model specifications, are available upon request.

[6]This holds if heterogeneity in θ_j is larger than preference heterogeneity.

implications should hold also if household membership changes over the span of the sample. The following discussion is only an approximation to true behavior (the model does not allow me to formally characterize the consumption dynamics when households switch between constrained and unconstrained) but it highlights the main issues at play.

Suppose a household is unconstrained at $t-1$, constrained at t (which matters for how c_t changes to c_{t+1} once shocks hit at the start of $t+1$), but then again unconstrained at $t+1$. Moments involving one outcome (earnings, consumption) only, such as $\mathbb{E}(\Delta w_{2it}\Delta c_{it+1}|\mathbf{O}_t)$, are unaffected. All that matters for them is the snapshot classification as constrained/unconstrained at t as this is what determines the change Δc_{it+1}. So if ϱ_t households are unconstrained at t and $1-\varrho_t$ constrained, then $\mathbb{E}(\Delta w_{2it}\Delta c_{it+1}|\mathbf{O}_t)$ still identifies $\varrho_t\mathbb{E}(\eta_{c,w_2(i,t)}) + (1-\varrho_t)\mathbb{E}(\theta_{2it})$. Moments involving two outcomes (or the same outcome twice) are affected. In the case of $\mathbb{E}(\Delta c_{it}\Delta c_{it+1}|\mathbf{O}_{t-1},\mathbf{O}_t)$, Δc_{it} is given by (5) (the hypothetical household is unconstrained at $t-1$) and Δc_{it+1} by (10) (they are constrained at t). The covariance identifies the product $\mathbb{E}(\eta_{c,w_2(i,t-1)}\theta_{2it})$ among those who switch. This is negative (assuming consumption and hours are Frisch substitutes as in BPS), mitigating the upward bias from those continuously constrained. The mitigation depends on the fraction of households who switch at t.

I take wealthy households as the empirical counterpart of the theoretically *unconstrained*. Table D.1 presents descriptive statistics for four subsamples of gradually more stringently defined wealthy households (see also table 2). Column W1 describes households whose annual wealth A_t is at least as much as average annual consumption $\overline{C}_t$ in the baseline sample. These are households who can fund at least a year's consumption without working. Column W2 is for households whose annual wealth is at least twice as much as average annual consumption. Column W3 is like W1 with the additional condition that households hold real debt lower than \$2K in annual terms. Finally, column W4 is like W3 but wealth now excludes home equity (house value net of mortgages), therefore better proxies for liquid assets. Nearly 80% of the baseline sample is excluded (which clearly has implications for statistical power).

Table E.9 (columns W1-W4) estimates the preferred specification on the wealthy. Three observations emerge. First, the average consumption elasticities remain unchanged upon departure from the baseline, thus contradicting the first testable implication. Second, the variances of the consumption elasticities still reveal substantial consumption preference heterogeneity across households. These parameters are mostly smaller than the baseline so this pattern seems consistent with the second testable implication: consumption preference heterogeneity is partly eaten away as one moves towards households who are likelier to be less constrained. This is, however, expected to a certain extent as the most stringent samples are also more homogeneous. Third, most average labor supply elasticities get smaller the more stringent wealth is, thus invalidating the third testable implication. However, these households are richer so shocks likely induce strong income effects on their labor supply. Overall, the evidence does not provide clear support for liquidity constraints or adjustment costs of work.

Table E.1 – Robustness: Alternative Restrictions for Preferred Specification

	2nd & 3rd moments; labor supply cross-elasticities set to 0				
	(1)	(2)	(3)	(4)	(5)
restrictions:	fix α, no variation in η_{h_j,w_j}	fix α, no joint variation in η_{h_j,w_j}	estimate α	estimate α & $\eta_{h_1,w_1} = \eta_{h_2,w_2}$	$\eta_{c,w_1} = \eta_{c,w_2}$
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.023	-0.023	-0.033	-0.055	-0.043
	(0.030)	(0.030)	(0.040)	(0.072)	(0.042)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.036	-0.036	-0.053	-0.080	-0.043
	(0.048)	(0.047)	(0.047)	(0.057)	(0.042)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.281	0.281	0.276	0.275	0.274
	(0.123)	(0.122)	(0.119)	(0.106)	(0.121)
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.248	0.248	0.247	0.275	0.247
	(0.195)	(0.195)	(0.196)	(0.106)	(0.196)
$\mathrm{Var}(\eta_{c,w_1(i)})^{\#}$	0.088	0.088	0.091	0.094	0.129
	[0.016]	[0.016]	[0.091]	[0.105]	[0.011]
$\mathrm{Var}(\eta_{c,w_2(i)})^{\#}$	0.225	0.226	0.233	0.201	0.129
	[0.016]	[0.016]	[0.073]	[0.097]	[0.011]
$\mathrm{Var}(\eta_{h_1,w_1(i)})^{\#}$	0	0.002	0.018	0.001	0.022
		[0.149]	[0.212]	[0.092]	[0.201]
$\mathrm{Var}(\eta_{h_2,w_2(i)})^{\#}$	0	0.001	0.003	0.001	0.003
		[0.060]	[0.156]	[0.092]	[0.156]
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$	0.141	0.141	0.146	0.138	0.129
	(0.050)	(0.050)	(0.056)	(0.077)	(0.043)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$	0	0	0.036	0.065	0.049
			(0.028)	(0.087)	(0.028)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$	0	0	-0.008	0.065	-0.010
			(0.019)	(0.087)	(0.020)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$	0	0	0.058	0.094	0.049
			(0.040)	(0.066)	(0.028)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$	0	0	-0.013	0.094	-0.010
			(0.023)	(0.066)	(0.020)
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$	0	0	-0.003	0.001	-0.004
			(0.019)	(0.001)	(0.018)
Obs. [households $\times \Delta_t$]	6071	6071	6071	6071	6071

Notes: The table presents GMM estimates of first and second moments of wage elasticities under a number of alternative restrictions on preferences. The restrictions serve the purpose to improve efficiency of estimates in column 7 of tables 4-5 and can be seen as alternatives to the preferred specification in column 8 therein. All columns posit $\eta_{c,w_2(i)} = \alpha_i \times \eta_{c,w_1(i)}$. Column 1 fixes $\alpha_i = 1.6\ \forall i$ and shuts down heterogeneity in labor supply elasticities; column 2 fixes $\alpha_i = 1.6\ \forall i$ and shuts down joint heterogeneity in labor supply elasticities; column 3 estimates $\alpha_i = \alpha\ \forall i$ together with all other parameters; column 4 is like column 3 but also restricts the own labor supply elasticities to equal between spouses, i.e. $\eta_{h_1,w_1(i)} = \eta_{h_2,w_2(i)}$; column 5 imposes equal consumption-wage elasticities, i.e. $\alpha_i = 1\ \forall i$. Across all columns, the labor supply cross-elasticities are set to zero. All moments are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. $^{\#}p$-value in square brackets for the one-sided test that the respective parameter equals zero; standard errors appear in parentheses for all other parameters.

Table E.2 – Robustness: Remove Extreme Observations

	preferred specification			
	(1)	(2)	(3)	(4)
	baseline	distribution of wages central 99%	distribution of wages central 96%	distribution of W_j, Y_j, C central 99%
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.033	-0.006	-0.012	-0.011
	(0.032)	(0.021)	(0.019)	(0.020)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.053	-0.010	-0.019	-0.017
	(0.052)	(0.033)	(0.030)	(0.033)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.276	0.219	0.266	0.200
	(0.120)	(0.174)	(0.171)	(0.151)
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.247	0.194	0.542	0.271
	(0.196)	(0.109)	(0.101)	(0.109)
$\text{Var}(\eta_{c,w_1(i)})^{\#}$	0.091	0.082	0.151	0.080
	[0.014]	[0.000]	[0.000]	[0.000]
$\text{Var}(\eta_{c,w_2(i)})^{\#}$	0.233	0.209	0.387	0.206
	[0.014]	[0.000]	[0.000]	[0.000]
$\text{Var}(\eta_{h_1,w_1(i)})^{\#}$	0.018	0.014	0.008	0.018
	[0.209]	[0.461]	[0.491]	[0.405]
$\text{Var}(\eta_{h_2,w_2(i)})^{\#}$	0.003	0.003	0.005	0.014
	[0.152]	[0.426]	[0.378]	[0.443]
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$	0.146	0.131	0.242	0.129
	(0.050)	(0.016)	(0.016)	(0.015)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$	0.036	0.009	0.006	0.019
	(0.024)	(0.024)	(0.026)	(0.019)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.008	0.000	-0.009	-0.018
	(0.017)	(0.018)	(0.013)	(0.020)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$	0.058	0.014	0.010	0.030
	(0.038)	(0.039)	(0.041)	(0.030)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$	-0.013	0.000	-0.015	-0.028
	(0.027)	(0.028)	(0.021)	(0.033)
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.003	-0.002	-0.001	-0.012
	(0.019)	(0.110)	(0.078)	(0.112)
Obs. [households $\times\Delta_t$]	6071	5876	5357	5764

Notes: The table presents GMM estimates of first and second moments of wage elasticities in the preferred specification. Column 1 shows the baseline results of tables 4-5. Column 2 removes extreme observations of male and female wages trimming the bottom 0.5% and top 0.5% of the respective distributions. Column 3 extends this by trimming the bottom 2% and top 2% of the wage distributions. Column 4 removes extreme observations of wages, earnings, and consumption trimming the bottom 0.5% and top 0.5% of the respective distributions. All moments are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. $^{\#}p$-value in square brackets for the one-sided test that the respective parameter equals zero; standard errors appear in parentheses for all other parameters.

Table E.3 – Robustness: Estimates of Preferences net of Age of Youngest Child & Chores, Means and Variances

	(1)*	(2)	(3)	(4)	(5)	(6)	(7)	(8)
specification/moments:	BPS	BPS-repl.	2^{nd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}	2^{nd} & 3^{rd}
heterogeneity:	no	no	no	no	marginal	joint	joint-restr.	preferred
change from column on the left:		2011 data, past levels	only transitory	third moments	marg. heterogeneity	joint heterogeneity	$\eta_{h_j,w_{j'}} = 0$	mild homogeneity
$\mathbb{E}(\eta_{c,w_1(i)})$		-0.236	0.034	-0.018	-0.018	-0.040	-0.042	-0.033
		(0.058)	(0.064)	(0.060)	(0.055)	(0.062)	(0.063)	(0.033)
$\mathbb{E}(\eta_{c,w_2(i)})$		0.051	-0.117	-0.055	-0.045	-0.030	-0.030	-0.052
		(0.050)	(0.113)	(0.085)	(0.071)	(0.075)	(0.075)	(0.052)
$\mathbb{E}(\eta_{h_1,w_1(i)})$		0.591	0.419	0.281	0.281	0.274	0.271	0.273
		(0.176)	(0.152)	(0.128)	(0.127)	(0.129)	(0.122)	(0.123)
$\mathbb{E}(\eta_{h_1,w_2(i)})$		0.087	-0.056	-0.015	-0.015	-0.014		
		(0.054)	(0.061)	(0.048)	(0.047)	(0.047)		
$\mathbb{E}(\eta_{h_2,w_1(i)})$		0.180	-0.117	-0.031	-0.031	-0.030		
		(0.112)	(0.125)	(0.100)	(0.096)	(0.100)		
$\mathbb{E}(\eta_{h_2,w_2(i)})$		0.753	0.446	0.270	0.269	0.268	0.257	0.258
		(0.149)	(0.294)	(0.215)	(0.222)	(0.226)	(0.196)	(0.194)
$\text{Var}(\eta_{c,w_1(i)})^{\#}$					0.074	0.093	0.086	0.091
					[0.155]	[0.107]	[0.120]	[0.012]
$\text{Var}(\eta_{c,w_2(i)})^{\#}$					0.283	0.269	0.255	0.233
					[0.127]	[0.119]	[0.114]	[0.012]
$\text{Var}(\eta_{h_1,w_1(i)})^{\#}$					0.002	0.022	0.024	0.020
					[0.129]	[0.215]	[0.221]	[0.213]
$\text{Var}(\eta_{h_1,w_2(i)})^{\#}$					0.000	0.000		
					[0.190]	[0.243]		
$\text{Var}(\eta_{h_2,w_1(i)})^{\#}$					0.002	0.003		
					[0.201]	[0.244]		
$\text{Var}(\eta_{h_2,w_2(i)})^{\#}$					0.001	0.004	0.004	0.003
					[0.056]	[0.164]	[0.180]	[0.165]

Notes: See notes for tables 4-5 on page 27. Observations: 6057 (households $\times \Delta_t$); these are slightly fewer than the baseline because of some missing data on chores. *BPS do not report similar results. #p-value in square brackets for the one-sided test that the respective parameter equals zero. All moments are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$.

Table E.4 – Robustness *(continues from previous table)*: Estimates of Preferences net of Age of Youngest Child & Chores, Covariances

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$						0.001	0.127	0.146
						(0.010)	(0.129)	(0.048)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$						0.041	0.042	0.039
						(0.035)	(0.036)	(0.024)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_2(i)})$						-0.001		
						(0.005)		
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_1(i)})$						-0.002		
						(0.009)		
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.002	-0.010	-0.009
						(0.013)	(0.022)	(0.018)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$						0.017	0.061	0.063
						(0.026)	(0.050)	(0.039)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_2(i)})$						0.001		
						(0.005)		
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_1(i)})$						0.003		
						(0.011)		
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$						-0.021	-0.021	-0.014
						(0.029)	(0.035)	(0.028)
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_1,w_2(i)})$						0.000		
						(0.003)		
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_1(i)})$						0.000		
						(0.007)		
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.003	-0.005	-0.004
						(0.019)	(0.030)	(0.023)
$\text{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_1(i)})$					0.000	0.000		
					(0.002)	(0.001)		
$\text{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_2(i)})$						0.000		
						(0.001)		
$\text{Cov}(\eta_{h_2,w_1(i)}, \eta_{h_2,w_2(i)})$						0.000		
						(0.003)		

Notes: See notes for tables 4-5 on page 27. Observations: 6057 (households $\times \Delta_t$); these are slightly fewer than the baseline because of some missing data on chores. All moments are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$.

Table E.5 – Estimates of Preferences: Means and Variances, Diagonally Weighted GMM

	(1)*	(2)	(3)	(4)	(5)	(6)	(7)	(8)
specification/moments:	BPS	BPS-repl.	2nd	2nd & 3rd	2nd & 3rd	2nd & 3rd	2nd & 3rd	2nd & 3rd
heterogeneity:	no	no	no	no	marginal	joint	joint-restr.	preferred
change from column on the left:		2011 data, past levels	only transitory	third moments	marg. heterogeneity	joint heterogeneity	$\eta_{h_j,w_{j'}} = 0$	mild homogeneity
$\mathbb{E}(\eta_{c,w_1(i)})$		-0.402	0.047	0.018	0.013	-0.017	-0.027	-0.017
		(0.103)	(0.115)	(0.049)	(0.038)	(0.051)	(0.048)	(0.026)
$\mathbb{E}(\eta_{c,w_2(i)})$		-0.198	-0.178	-0.081	-0.059	-0.023	-0.011	-0.032
		(0.078)	(0.309)	(0.088)	(0.058)	(0.066)	(0.064)	(0.049)
$\mathbb{E}(\eta_{h_1,w_1(i)})$		0.273	0.401	0.324	0.323	0.307	0.260	0.287
		(0.107)	(0.139)	(0.104)	(0.112)	(0.109)	(0.110)	(0.109)
$\mathbb{E}(\eta_{h_1,w_2(i)})$		-0.054	-0.048	-0.037	-0.037	-0.033		
		(0.037)	(0.058)	(0.040)	(0.040)	(0.041)		
$\mathbb{E}(\eta_{h_2,w_1(i)})$		-0.113	-0.100	-0.077	-0.077	-0.068		
		(0.077)	(0.123)	(0.084)	(0.084)	(0.086)		
$\mathbb{E}(\eta_{h_2,w_2(i)})$		-0.262	0.437	0.298	0.301	0.286	0.213	0.210
		(0.213)	(0.282)	(0.208)	(0.217)	(0.215)	(0.177)	(0.178)
$\mathrm{Var}(\eta_{c,w_1(i)})^{\#}$					0.061	0.057	0.038	0.036
					[0.179]	[0.182]	[0.250]	[0.086]
$\mathrm{Var}(\eta_{c,w_2(i)})^{\#}$					0.115	0.127	0.137	0.130
					[0.237]	[0.221]	[0.221]	[0.086]
$\mathrm{Var}(\eta_{h_1,w_1(i)})^{\#}$					0.001	0.057	0.159	0.068
					[0.251]	[0.286]	[0.205]	[0.289]
$\mathrm{Var}(\eta_{h_1,w_2(i)})^{\#}$					0.000	0.002		
					[0.095]	[0.270]		
$\mathrm{Var}(\eta_{h_2,w_1(i)})^{\#}$					0.003	0.013		
					[0.143]	[0.258]		
$\mathrm{Var}(\eta_{h_2,w_2(i)})^{\#}$					0.000	0.037	0.064	0.037
					[0.074]	[0.228]	[0.243]	[0.263]

Notes: See notes for tables 4-5 on page 27. Observations: 6071 (households $\times\Delta_t$). *BPS do not report results from diagonally weighted GMM. #p-value in square brackets for the one-sided test that the respective parameter equals zero.

Table E.6 – Estimates of Preferences *(continues from previous table)*: Covariances, Diagonally Weighted GMM

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$						0.004	0.052	0.068
						(0.017)	(0.093)	(0.042)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$						0.034	0.032	0.039
						(0.040)	(0.046)	(0.030)
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_2(i)})$						0.002		
						(0.007)		
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_1(i)})$						0.005		
						(0.015)		
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.002	-0.020	-0.033
						(0.024)	(0.048)	(0.027)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$						0.070	0.136	0.074
						(0.091)	(0.091)	(0.057)
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_2(i)})$						-0.014		
						(0.012)		
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_1(i)})$						-0.029		
						(0.026)		
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$						-0.068	-0.086	-0.062
						(0.063)	(0.063)	(0.052)
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_1,w_2(i)})$						-0.006		
						(0.018)		
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_1(i)})$						-0.013		
						(0.037)		
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$						-0.038	-0.100	-0.048
						(0.130)	(0.166)	(0.155)
$\text{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_1(i)})$					0.000	0.003		
					(0.000)	(0.011)		
$\text{Cov}(\eta_{h_1,w_2(i)}, \eta_{h_2,w_2(i)})$						0.007		
						(0.012)		
$\text{Cov}(\eta_{h_2,w_1(i)}, \eta_{h_2,w_2(i)})$						0.016		
						(0.025)		

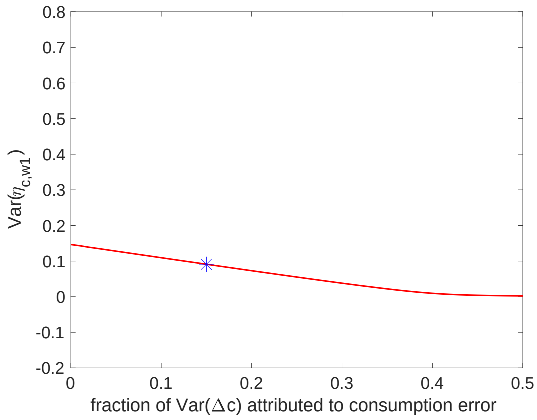
Notes: See notes for tables 4-5 on page 27. Observations: 6071 (households $\times \Delta_t$).

Table E.7 – Robustness: Alternative Restrictions, Diagonally Weighted GMM

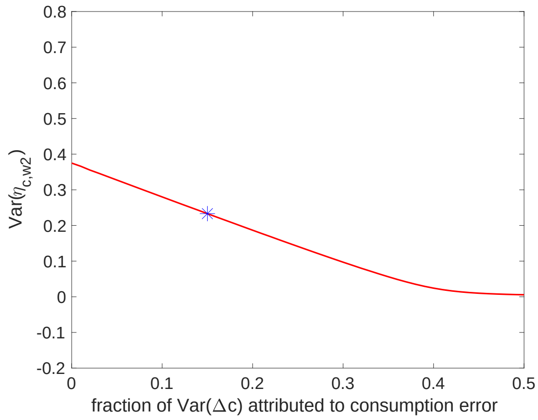
	2nd & 3rd moments; labor supply cross-elasticities set to 0				
	(1)	(2)	(3)	(4)	(5)
restrictions:	fix α, no variation in η_{h_j,w_j}	fix α, no joint variation in η_{h_j,w_j}	estimate α	estimate α & $\eta_{h_1,w_1} = \eta_{h_2,w_2}$	$\eta_{c,w_1} = \eta_{c,w_2}$
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.012	-0.012	-0.016	-0.013	-0.027
	(0.020)	(0.021)	(0.032)	(0.036)	(0.035)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.022	-0.022	-0.032	-0.027	-0.027
	(0.039)	(0.040)	(0.038)	(0.042)	(0.035)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.304	0.302	0.286	0.282	0.287
	(0.102)	(0.114)	(0.109)	(0.094)	(0.110)
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.215	0.216	0.210	0.282	0.208
	(0.181)	(0.185)	(0.179)	(0.094)	(0.178)
$\mathrm{Var}(\eta_{c,w_1(i)})^{\#}$	0.036	0.036	0.033	0.031	0.062
	[0.086]	[0.086]	[0.200]	[0.208]	[0.075]
$\mathrm{Var}(\eta_{c,w_2(i)})^{\#}$	0.130	0.130	0.137	0.140	0.062
	[0.086]	[0.086]	[0.143]	[0.141]	[0.075]
$\mathrm{Var}(\eta_{h_1,w_1(i)})^{\#}$		0.005	0.070	0.001	0.068
		[0.310]	[0.280]	[0.047]	[0.279]
$\mathrm{Var}(\eta_{h_2,w_2(i)})^{\#}$		0.000	0.037	0.001	0.049
		[0.150]	[0.266]	[0.047]	[0.255]
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$	0.069	0.069	0.067	0.066	0.062
	(0.042)	(0.042)	(0.049)	(0.052)	(0.035)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$			0.037	0.002	0.052
			(0.036)	(0.042)	(0.039)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$			-0.031	0.002	-0.051
			(0.039)	(0.042)	(0.041)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$			0.076	0.004	0.052
			(0.055)	(0.045)	(0.039)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$			-0.063	0.004	-0.051
			(0.050)	(0.045)	(0.041)
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$			-0.049	0.001	-0.055
			(0.149)	(0.000)	(0.154)
Obs. [households $\times \Delta_t$]	6071	6071	6071	6071	6071

Notes: See notes for table E.1 on appendix page 26. The table presents diagonally weighted GMM estimates of first and second moments of wage elasticities under a number of alternative restrictions on preferences. All moments are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. ${}^{\#}p$-value in square brackets for the one-sided test that the respective parameter equals zero; standard errors appear in parentheses for all other parameters.

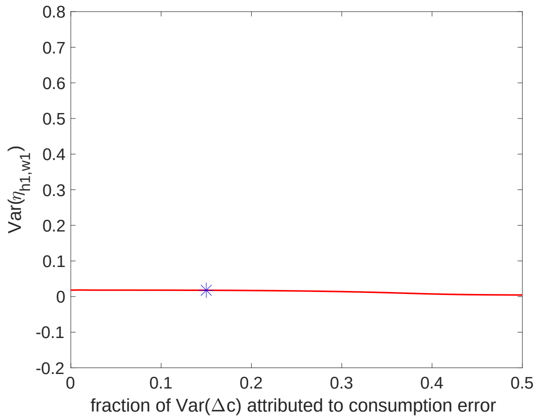
Figure E.1 – Preference Heterogeneity against Consumption Measurement Error



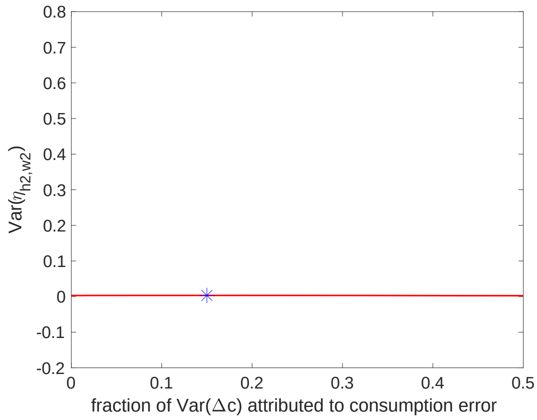
(a) marginal heterogeneity in η_{c,w_1}



(b) marginal heterogeneity in η_{c,w_2}



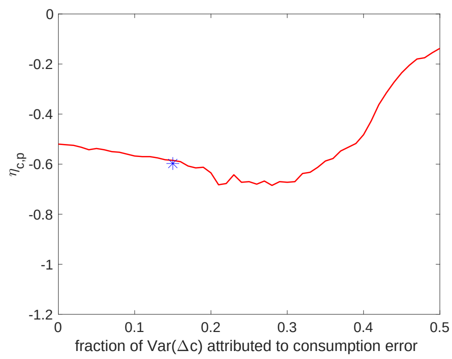
(c) marginal heterogeneity in η_{h_1,w_1}



(d) marginal heterogeneity in η_{h_2,w_2}

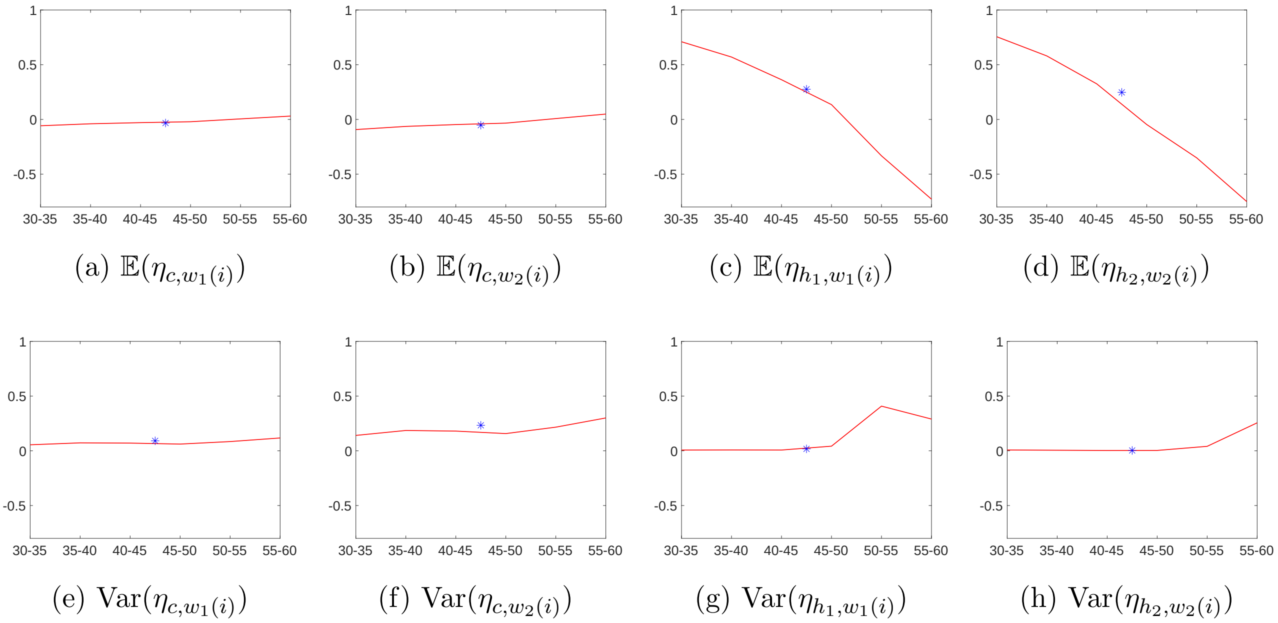
Notes: The figures present the evolution of the variances of consumption and labor supply elasticities in the preferred specification against values for the variance of consumption measurement error between zero and $50\% \times \mathrm{Var}(\Delta c_{it})$. Measurement error in wages and earnings remains fixed at the baseline values of 13% of the variance of wages and 4% of the variance of earnings respectively. The blue asterisk marks the baseline parameter estimate when consumption measurement error is fixed at $15\% \times \mathrm{Var}(\Delta c_{it})$. All parameters are estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$.

Figure E.2 – Consumption Substitution Elasticity against Consumption Measurement Error



Notes: The figure presents the evolution of the consumption substitution elasticity against values for the variance of consumption measurement error between zero and $50\% \times \mathrm{Var}(\Delta c_{it})$. Measurement error in wages and earnings remains fixed at the baseline values of 13% of the variance of wages and 4% of the variance of earnings respectively. The blue asterisk marks the baseline parameter estimate when consumption measurement error is fixed at $15\% \times \mathrm{Var}(\Delta c_{it})$. The elasticity is estimated at the sample average of consumption and hours, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$, using simulated minimum distance and assuming that all other parameters are jointly normally distributed, parameterized at the consumption error specific preference estimates underlying figure E.1.

Figure E.3 – Parameter Estimates over Age Brackets

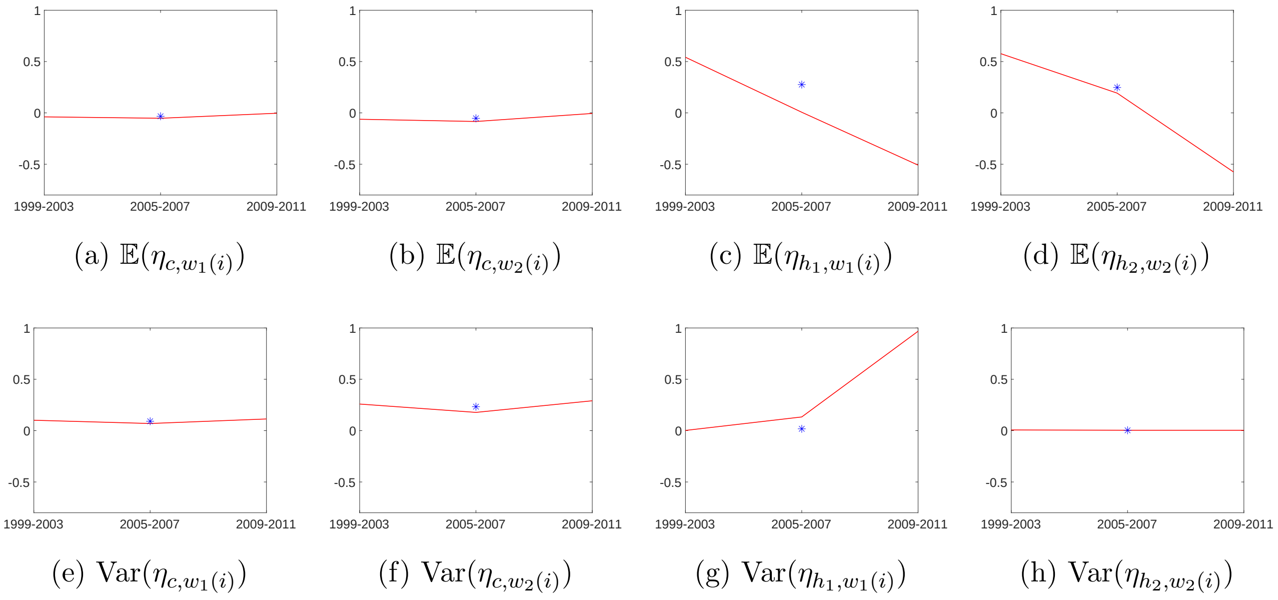


(a) $\mathbb{E}(\eta_{c,w_1(i)})$ (b) $\mathbb{E}(\eta_{c,w_2(i)})$ (c) $\mathbb{E}(\eta_{h_1,w_1(i)})$ (d) $\mathbb{E}(\eta_{h_2,w_2(i)})$

(e) $\text{Var}(\eta_{c,w_1(i)})$ (f) $\text{Var}(\eta_{c,w_2(i)})$ (g) $\text{Var}(\eta_{h_1,w_1(i)})$ (h) $\text{Var}(\eta_{h_2,w_2(i)})$

Notes: This figure plots selected parameter estimates in the preferred specification, over six brackets for the age of the male spouse: 30-35, 35-40, 40-45, 45-50, 50-55, 55-60. I condition on the sample average of consumption and hours within each age bracket. The blue asterisk marks the parameter estimate in the baseline of tables 4-5.

Figure E.4 – Parameter Estimates over Calendar Years



(a) $\mathbb{E}(\eta_{c,w_1(i)})$ (b) $\mathbb{E}(\eta_{c,w_2(i)})$ (c) $\mathbb{E}(\eta_{h_1,w_1(i)})$ (d) $\mathbb{E}(\eta_{h_2,w_2(i)})$

(e) $\text{Var}(\eta_{c,w_1(i)})$ (f) $\text{Var}(\eta_{c,w_2(i)})$ (g) $\text{Var}(\eta_{h_1,w_1(i)})$ (h) $\text{Var}(\eta_{h_2,w_2(i)})$

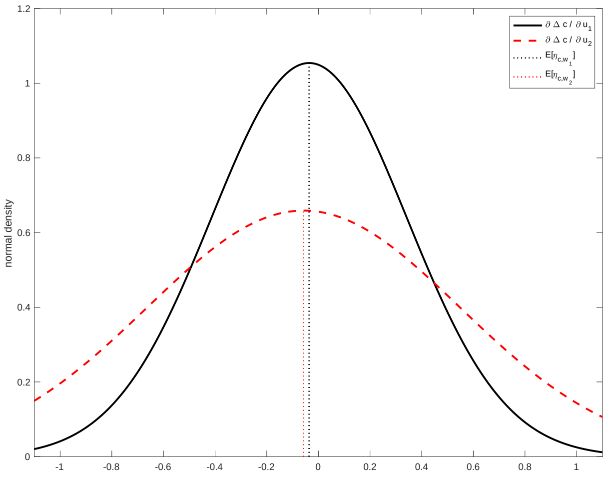
Notes: This figure plots selected parameter estimates in the preferred specification, over different calendar periods: 1999-2003 (3 waves of PSID data), 2005-2007 (two waves), 2009-2011 (two waves). I condition on the sample average of consumption and hours within each calendar period. The blue asterisk marks the parameter estimate in the baseline of tables 4-5.

Table E.8 – Accounting Decomposition of Consumption Inequality for varying Consumption Measurement Error

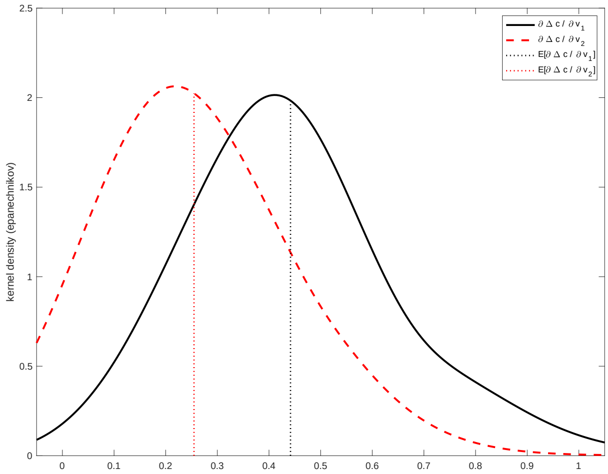
		% of cons. inequality	% of cons. instability	% of perm. inequality
		A. no consumption error		
$\text{Var}(\Delta c_{it})$	0.065	100%		
consumption instability	0.019	29.1%	100%	
without pref. heterogeneity	0.000		0.9%	
pref. heterogeneity induced	0.019		99.1%	
permanent inequality	0.046	70.9%		100%
without heterogeneity	0.024			51.4%
heterogen. in preferences only	0.034			73.2%
heterogen. in preferences, π_{it}, $\mathbf{s}_{it}$	0.046			100%
		B. consumption error at 40% × Var($\boldsymbol{\Delta c_{it}}$)		
$\text{Var}(\Delta c_{it})$	0.040	100%		
consumption instability	0.002	3.9%	100%	
without pref. heterogeneity	0.000		5.1%	
pref. heterogeneity induced	0.001		94.9%	
permanent inequality	0.038	96.3%		100%
without heterogeneity	0.024			63.2%
heterogen. in preferences only	0.029			76.1%
heterogen. in preferences, π_{it}, $\mathbf{s}_{it}$	0.038			100%

Notes: The table presents the decomposition of consumption inequality into consumption instability and permanent inequality for two different amounts of consumption measurement error: no error in panel A, and consumption error at 40% of the variance of consumption growth in panel B. $\text{Var}(\Delta c_{it})$ is estimated at the sample average of consumption and hours (i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$). Simulations of permanent inequality assume preferences are jointly normal, parameterized at the consumption error specific preference estimates underlying figure E.1. The consumption substitution elasticity $\eta_{c,p}$ is also consumption error specific, reflecting figure E.2. Wealth shares π_{it} and $\mathbf{s}_{it}$ are drawn from their empirical distributions. The sum of consumption instability and permanent inequality slightly overshoots the empirical amount of inequality in panel B reflecting that the model is unable to perfectly match its empirical target in this case. The decomposition suggests that preference heterogeneity accounts for up to 44% of consumption inequality when consumption error is absent (99.1% of consumption instability and 21.8% of permanent inequality in panel A) and as little as 16% of consumption inequality when consumption error is prevalent (94.9% of consumption instability and 12.9% of permanent inequality in panel B).

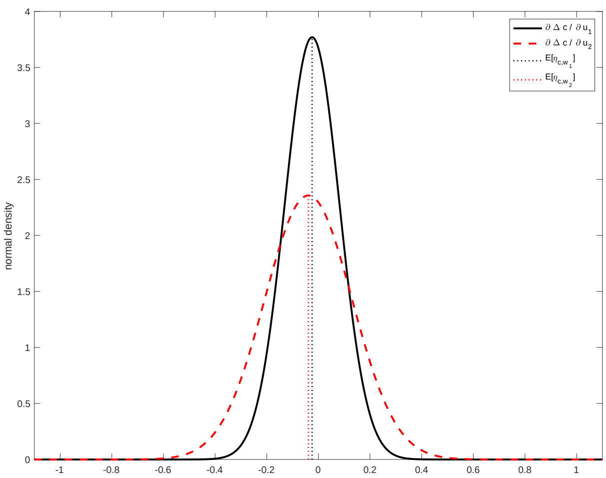
Figure E.5 – Distributions of Pass-Through Rates of Shocks into Consumption for varying Consumption Measurement Error



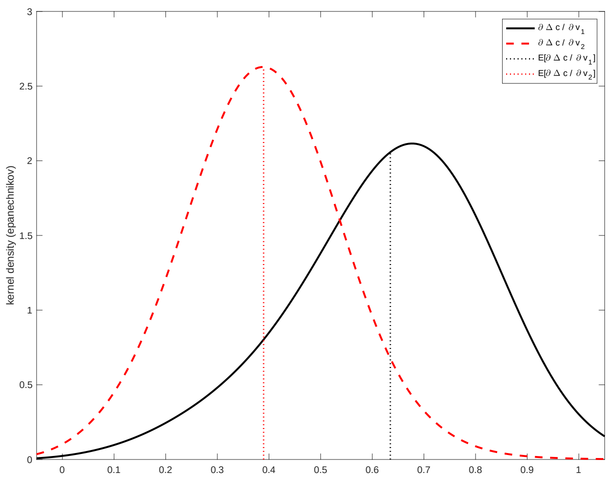
(a) transitory shocks, no consumption error



(b) permanent shocks, no consumption error



(c) transitory shocks, consumption error at $40\% \times \text{Var}(\Delta c_{it})$



(d) permanent shocks, consumption error at $40\% \times \text{Var}(\Delta c_{it})$

Notes: The figures visualize the distributions of pass-through rates of shocks for two different amounts of consumption measurement error: no error in graphs (a) and (b), and consumption error at 40% of the variance of consumption growth in graphs (c) and (d). The distributions are across 10 million households whose preferences are drawn from the joint normal, parameterized at the consumption error specific preference estimates underlying figure E.1. The consumption substitution elasticity $\eta_{c,p}$ is also consumption error specific (figure E.2), while π_{it} and $\mathbf{s}_{it}$ are drawn from their empirical distributions. The mass placed over specific pass-through rates is arbitrary following the normality assumption. Consumption measurement error reduces the scope of preference heterogeneity in the model and this is best reflected in graph (c) vis-à-vis (a). However, a limited preference heterogeneity when consumption error is at $40\% \times \text{Var}(\Delta c_{it})$ can still produce substantial heterogeneity in the consumption response to permanent shocks in graph (d). This is because the limited preference heterogeneity interacts with the underlying heterogeneity in wealth shares π_{it} and $\mathbf{s}_{it}$. The implied *average* pass-through of permanent shocks is as large as 0.6 in this case; given that the evidence in Blundell et al. (2008) and BPS suggests a lower empirical pass-through of permanent shocks, an amount of consumption error as large as $40\% \times \text{Var}(\Delta c_{it})$ is likely counterfactual given this model.

Table E.9 – Robustness: Joint Taxation and Wealthy Households

	preferred specification				
	(1)	(W1)	(W2)	(W3)	(W4)
taxation:	joint	no joint taxation			
sample:	baseline	$A > \overline{C}$	$A > 2\overline{C}$	$A > \overline{C}$ no debt	$A > \overline{C}$ liquid
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.037	-0.035	-0.028	0.028	0.083
	(0.021)	(0.025)	(0.026)	(0.038)	(0.036)
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.058	-0.055	-0.044	0.044	0.133
	(0.034)	(0.040)	(0.041)	(0.061)	(0.057)
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.326	0.184	0.171	0.398	0.214
	(0.103)	(0.203)	(0.219)	(0.439)	(0.443)
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.277	0.138	0.059	-0.055	-0.431
	(0.350)	(0.133)	(0.146)	(0.168)	(0.174)
$\mathrm{Var}(\eta_{c,w_1(i)})^{\#}$	0.124	0.071	0.051	0.087	0.022
	[0.000]	[0.000]	[0.000]	[0.000]	[0.034]
$\mathrm{Var}(\eta_{c,w_2(i)})^{\#}$	0.318	0.181	0.129	0.223	0.057
	[0.000]	[0.000]	[0.000]	[0.000]	[0.034]
$\mathrm{Var}(\eta_{h_1,w_1(i)})^{\#}$	0.028	0.005	0.004	0.009	0.047
	[0.233]	[0.435]	[0.464]	[0.390]	[0.323]
$\mathrm{Var}(\eta_{h_2,w_2(i)})^{\#}$	0.005	0.003	0.002	0.045	0.035
	[0.122]	[0.424]	[0.429]	[0.147]	[0.097]
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$	0.199	0.113	0.081	0.139	0.036
	(0.001)	(0.015)	(0.016)	(0.022)	(0.019)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$	0.053	0.016	0.010	0.013	0.019
	(0.051)	(0.025)	(0.027)	(0.039)	(0.036)
$\mathrm{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.013	0.001	0.001	0.057	0.014
	(0.014)	(0.019)	(0.020)	(0.015)	(0.009)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$	0.085	0.025	0.016	0.021	0.031
	(0.082)	(0.040)	(0.043)	(0.062)	(0.057)
$\mathrm{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$	-0.021	0.001	0.001	0.091	0.023
	(0.023)	(0.030)	(0.033)	(0.024)	(0.014)
$\mathrm{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.006	0.000	0.000	0.006	-0.011
	(0.012)	(0.137)	(0.141)	(0.062)	(0.042)
Obs. [households $\times \Delta_t$]	6071	4614	3789	1806	1374

Notes: The table presents GMM estimates of first and second moments of wage elasticities in the preferred specification. Column 1 uses the baseline sample and estimates a model with progressive joint taxation. The remaining columns estimate the baseline model, i.e. without taxation. Column W1 is for households with wealth A_t at least as much as average consumption $\overline{C}_t$ in the baseline sample. Column W2 is for households with wealth at least twice as much as average annual consumption. Column W3 is like column W1 with the additional condition that households hold real debt that does not exceed \$2K. Column W4 is like column W3 but the relevant measure of wealth excludes home equity (the value of one's home net of outstanding mortgages), therefore it proxies for liquid assets. All parameters are estimated at the average of consumption and hours within each sample, i.e. $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$. $^{\#}p$-value in square brackets for the one-sided test that the respective parameter equals zero; standard errors appear in parentheses for all other parameters.

F Consumption, Earnings, and Hours Inequality

To simplify the illustration of the derivation of consumption inequality, I write consumption growth as $\Delta c_{it} \approx \eta_{c,w_1(i,t-1)}\Delta u_{1it} + \eta_{c,w_2(i,t-1)}\Delta u_{2it} + \phi_{1it}v_{1it} + \phi_{2it}v_{2it}$. This mimics expression (5) in the main text with $\phi_{jit} = \eta_{c,w_j(i,t-1)} + \bar{\eta}_{c(i,t-1)}\varepsilon_j(\pi_{it}, \mathbf{s}_{it}; \boldsymbol{\eta}_{it-1})$. I also define $t^- \equiv t-1$. From the properties of the variance operator it follows that

$$\begin{aligned}
\mathrm{Var}(\Delta c_{it}) \approx\ & \mathrm{Var}(\eta_{c,w_1(i,t^-)}\Delta u_{1it}) + \mathrm{Var}(\eta_{c,w_2(i,t^-)}\Delta u_{2it}) + \mathrm{Var}(\phi_{1it}v_{1it}) + \mathrm{Var}(\phi_{2it}v_{2it}) \\
& + 2\mathrm{Cov}(\eta_{c,w_1(i,t^-)}\Delta u_{1it}, \eta_{c,w_2(i,t^-)}\Delta u_{2it}) + 2\mathrm{Cov}(\eta_{c,w_1(i,t^-)}\Delta u_{1it}, \phi_{1it}v_{1it}) \\
& + 2\mathrm{Cov}(\eta_{c,w_1(i,t^-)}\Delta u_{1it}, \phi_{2it}v_{2it}) + 2\mathrm{Cov}(\eta_{c,w_2(i,t^-)}\Delta u_{2it}, \phi_{1it}v_{1it}) \\
& + 2\mathrm{Cov}(\eta_{c,w_2(i,t^-)}\Delta u_{2it}, \phi_{2it}v_{2it}) + 2\mathrm{Cov}(\phi_{1it}v_{1it}, \phi_{2it}v_{2it}).
\end{aligned}$$

Given assumption 1 and results from Goodman (1960), $\mathrm{Var}(\eta_{c,w_1(i,t^-)}\Delta u_{1it})$ becomes equal to $\mathbb{E}(\eta^2_{c,w_1(i,t^-)})\mathrm{Var}(\Delta u_{1it})$. Results for the other variances are similar. From Bohrnstedt and Goldberger (1969) one can show that most covariances are zero except those involving exclusively transitory shocks or exclusively permanent shocks. Most covariances are zero because of assumption 1, shocks have zero mean, and the assumption that permanent and transitory shocks are independent. The non-zero covariances are $\mathrm{Cov}(\eta_{c,w_1(i,t^-)}\Delta u_{1it}, \eta_{c,w_2(i,t^-)}\Delta u_{2it}) = \mathbb{E}(\eta_{c,w_1(i,t^-)}\eta_{c,w_2(i,t^-)})\mathrm{Cov}(\Delta u_{1it}, \Delta u_{2it})$ and $\mathrm{Cov}(\phi_{1it}v_{1it}, \phi_{2it}v_{2it})$.

By similar arguments, the analytical expression for hours inequality is given by

$$\begin{aligned}
\mathrm{Var}(\Delta h_{jit}) \approx\ & \sum_{j'=1}^{2} \mathbb{E}(\eta^2_{h_j,w_{j'}(i,t^-)}) \times \left(\sigma^2_{u_{j'}(t)} + \sigma^2_{u_{j'}(t^-)}\right) \\
& + 2\mathbb{E}(\eta_{h_j,w_1(i,t^-)}\eta_{h_j,w_2(i,t^-)}) \times \left(\sigma_{u_1u_2(t)} + \sigma_{u_1u_2(t^-)}\right) \\
& + \sum_{j'=1}^{2} \mathbb{E}\left(\left(\eta_{h_j,w_{j'}(i,t^-)} + \bar{\eta}_{h_j(i,t^-)}\varepsilon_{j'}(\cdot;\boldsymbol{\eta}_{it^-})\right)^2\right) \times \sigma^2_{v_{j'}(t)} \\
& + 2\mathbb{E}\left(\left(\eta_{h_j,w_1(i,t^-)} + \bar{\eta}_{h_j(i,t^-)}\varepsilon_1(\cdot;\boldsymbol{\eta}_{it^-})\right)\left(\eta_{h_j,w_2(i,t^-)} + \bar{\eta}_{h_j(i,t^-)}\varepsilon_2(\cdot;\boldsymbol{\eta}_{it^-})\right)\right) \times \sigma_{v_1v_2(t)}.
\end{aligned}$$

The expression for $\mathrm{Var}(\Delta y_{jit})$ follows from the identity $\Delta y_{jit} = \Delta h_{jit} + \Delta w_{jit}$.

Two remarks are due. First, as consumption (or hours/earning) inequality increases with preference heterogeneity, neglecting such heterogeneity results in understating inequality or biasing mean preferences upwards. Second, positive assortative matching between spousal wages, captured by a positive correlation between shocks, may not always increase inequality. Consider the loading factor of the *covariance* of transitory shocks into consumption inequality, given by $\mathbb{E}(\eta_{c,w_1(i,t^-)}\eta_{c,w_2(i,t^-)}) = \mathbb{E}(\eta_{c,w_1(i,t^-)})\mathbb{E}(\eta_{c,w_2(i,t^-)}) + \mathrm{Cov}(\eta_{c,w_1(i,t^-)}, \eta_{c,w_2(i,t^-)})$. What (also) matters for loading such covariance is the correlation between consumption-wage elasticities. If $\mathbb{E}(\eta_{c,w_1(i,t^-)}\eta_{c,w_2(i,t^-)})$ is negative due to a strong negative correlation between elasticities, positive assortative matching on wages actually decreases inequality.

G Time Aggregation in the PSID

This appendix highlights the implications of time aggregation using a stylized setting in which wages are paid on a monthly basis but are observed on a yearly basis as the sum over the previous twelve months. Therefore, *model* time now refers to months rather than years. I discuss extensions to this stylized setting in the end. Like Working (1960) but unlike Crawley (2020), I maintain that time is discrete; otherwise I follow these two papers closely.

Following the baseline notation, let residual log wage of spouse $j = \{1, 2\}$ in year t and month m be $w_{jit,m} \equiv \ln W_{jit,m} - \mathbf{X}'_{jit,m}\boldsymbol{\alpha}_{W_j}$. In what follows, I drop the cross-sectional subscript i in order to ease the notation. The monthly wage process is given by

$$\begin{aligned} w_{jt,m} &= w^p_{jt,m} + u_{jt,m} \\ w^p_{jt,m} &= w^p_{jt,m-1} + v_{jt,m}, \end{aligned} \tag{G.1}$$

where $w^p_{jit,m} \equiv \ln W^p_{jit,m}$ is the permanent component. All properties of shocks remain as in the baseline so, among other things, shocks $v_{jt,m}$ and $u_{jt,m}$ are serially uncorrelated.

We do not observe $w_{jt,m}$ but we do observe the sum of wages over $M = 12$ calendar months in the year,[7] namely

$$w^{\text{year}}_{jt} = \sum_{m=1}^{M} w_{jt,m} = \sum_{m=1}^{M} w^p_{jt,m} + \sum_{m=1}^{M} u_{jt,m}. \tag{G.2}$$

Following Working (1960), I use (G.1) to express the first sum in (G.2) in terms of $w^p_{jt,1}$ (the permanent wage in January of year t) and the permanent shocks since then; this is given by $\sum_{m=1}^{M} w^p_{jt,m} = M \times w^p_{jt,1} + \sum_{m=1}^{M-1}(M-m)v_{jt,1+m}$. In a similar way, I express the sum of permanent wages in year $t-1$ in terms of $w^p_{jt,1}$ and the permanent shocks until then; namely $\sum_{m=1}^{M} w^p_{jt-1,m} = M \times w^p_{jt,1} - M \times v_{jt,1} - \sum_{m=1}^{M-1}(M-m)v_{jt-1,M+1-m}$. One can then obtain the expression for yearly wage growth $\Delta^{\text{year}} w^{\text{year}}_{jt} = w^{\text{year}}_{jt} - w^{\text{year}}_{jt-1}$.

The PSID is biennial after 1997 so observed wage growth is a first difference over *two years*, namely $\Delta^{\text{psid}} w^{\text{year}}_{jt} = w^{\text{year}}_{jt} - w^{\text{year}}_{jt-2}$. This is equivalent to the sum of two consecutive yearly growth rates, i.e. $\Delta^{\text{psid}} w^{\text{year}}_{jt} = \Delta^{\text{year}} w^{\text{year}}_{jt} + \Delta^{\text{year}} w^{\text{year}}_{jt-1}$, given by

$$\Delta^{\text{psid}} w^{\text{year}}_{jt} = \sum_{m=0}^{M-1}(M-m)v_{jt,1+m} + M\sum_{m=1}^{M} v_{jt-1,m} + \sum_{m=1}^{M-1}(M-m)v_{jt-2,M+1-m} + \sum_{m=1}^{M} \Delta^{\text{psid}} u_{jt,m}$$

where $\Delta^{\text{psid}} u_{jt,m} = u_{jt,m} - u_{jt-2,m}$. The expressions for $\Delta^{\text{psid}} w^{\text{year}}_{jt-2}$ etc. are analogous.

The baseline model remains effectively unchanged, except that households make consumption and hours decisions on a month-by-month basis rather than yearly. All the approximations

[7] We observe the annual *level* of wages, not the *logarithm*, and the logarithm of the sum is not the same as the sum of the logarithms. I refer to Crawley (2020) for a discussion of this point. Another problem arises from the fact that wages in the PSID are earnings over hours, and the denominator itself is subject to time aggregation. I do not attempt to address this point here but it certainly deserves attention in future work.

and derivations still hold, but they are now recast on a monthly basis. The system of consumption and hours equations, analogous to (5)-(7), is now given by

$$\Delta c_{t,m} \approx \eta_{c,w_1(t,m-1)}\Delta u_{1t,m} + \eta_{c,w_2(t,m-1)}\Delta u_{2t,m} + \kappa_{cv_1(t,m-1)}v_{1t,m} + \kappa_{cv_2(t,m-1)}v_{2t,m}$$
$$\Delta h_{1t,m} \approx \eta_{h_1,w_1(t,m-1)}\Delta u_{1t,m} + \eta_{h_1,w_2(t,m-1)}\Delta u_{2t,m} + \kappa_{h_1v_1(t,m-1)}v_{1t,m} + \kappa_{h_1v_2(t,m-1)}v_{2t,m}$$
$$\Delta h_{2t,m} \approx \eta_{h_2,w_1(t,m-1)}\Delta u_{1t,m} + \eta_{h_2,w_2(t,m-1)}\Delta u_{2t,m} + \kappa_{h_2v_1(t,m-1)}v_{1t,m} + \kappa_{h_2v_2(t,m-1)}v_{2t,m},$$

where the lifecycle Marshallian elasticities are given by

$$\kappa_{cv_j(t,m-1)} = \eta_{c,w_j(t,m-1)} + \bar{\eta}_{c(t,m-1)}\varepsilon_j(\pi_{t,m}, \mathbf{s}_{t,m}; \boldsymbol{\eta}_{t,m-1})$$
$$\kappa_{h_1v_j(t,m-1)} = \eta_{h_1,w_j(t,m-1)} + \bar{\eta}_{h_1(t,m-1)}\varepsilon_j(\pi_{t,m}, \mathbf{s}_{t,m}; \boldsymbol{\eta}_{t,m-1})$$
$$\kappa_{h_2v_j(t,m-1)} = \eta_{h_2,w_j(t,m-1)} + \bar{\eta}_{h_2(t,m-1)}\varepsilon_j(\pi_{t,m}, \mathbf{s}_{t,m}; \boldsymbol{\eta}_{t,m-1}),$$

and growth here refers to months (model time). As in the case of wages, we do not observe $c_{t,m}$ and $h_{jt,m}$; instead we observe $c_t^{\text{year}} = \sum_{m=1}^M c_{t,m}$ and $h_{jt}^{\text{year}} = \sum_{m=1}^M h_{jt,m}$ for $j = \{1,2\}$.[8]

Following Working (1960), I use the expression for $\Delta c_{t,m}$ to obtain yearly consumption $c_t^{\text{year}} \approx M \times c_{t,1} + \sum_j \sum_{m=1}^{M-1}(M-m)\eta_{c,w_j(t,m)}\Delta u_{jt,1+m} + \sum_j \sum_{m=1}^{M-1}(M-m)\kappa_{cv_j(t,m)}v_{jt,1+m}$. In a similar way, one can show that $c_{t-1}^{\text{year}} \approx M \times c_{t,1} - M\sum_j \eta_{c,w_j(t-1,M)}\Delta u_{jt,1} - M\sum_j \kappa_{cv_j(t-1,M)}v_{jt,1} - \sum_j \sum_{m=1}^{M-1}(M-m)\eta_{c,w_j(t-1,M-m)}\Delta u_{jt-1,M+1-m} - \sum_j \sum_{m=1}^{M-1}(M-m)\kappa_{cv_j(t-1,M-m)}v_{jt-1,M+1-m}$. One can then obtain the expression for yearly consumption growth $\Delta^{\text{year}}c_t^{\text{year}} = c_t^{\text{year}} - c_{t-1}^{\text{year}}$. Observed consumption growth is, however, a first difference over *two years*, namely $\Delta^{\text{psid}}c_t^{\text{year}} = c_t^{\text{year}} - c_{t-2}^{\text{year}} = \Delta^{\text{year}}c_t^{\text{year}} + \Delta^{\text{year}}c_{t-1}^{\text{year}}$, given by

$$\begin{aligned}\Delta^{\text{psid}}c_t^{\text{year}} \approx & \sum_j \sum_{m=0}^{M-1}(M-m)\kappa_{cv_j(t,m)}v_{jt,1+m} + M\sum_j \sum_{m=1}^{M}\kappa_{cv_j(t-1,m-1)}v_{jt-1,m} \\ & + \sum_j \sum_{m=1}^{M-1}(M-m)\kappa_{cv_j(t-2,M-m)}v_{jt-2,M+1-m} \\ & + \sum_j \sum_{m=1}^{M}\left(\eta_{c,w_j(t,m-1)}u_{jt,m} - \eta_{c,w_j(t-2,m-1)}u_{jt-2,m}\right).\end{aligned}$$

To obtain the last expression, I assume for simplicity that $\eta_{c,w_j(t,m)} \approx \eta_{c,w_j(t,m-1)}$ and similarly for $t-1$ etc. Then multiple terms involving transitory shocks cancel out due to mean reversion. One can obtain analogous expressions for $\Delta^{\text{psid}}h_{jt}^{\text{year}}$ in a similar way. Detailed steps for all these derivations (admittedly presented here only at a high level) are available upon request.

[8]This is not strictly true for consumption. Certain consumption items are observed as a snapshot in a given calendar month (food) while most items are observed on a varying basis, such as weekly, biweekly, half-yearly, or yearly. The overall timing of consumption is often ambiguous and subject to alternative interpretations. As my focus here is on the implications of neglected time aggregation in wages, I maintain simplicity by treating consumption in a comparable way to wages and hours. I assume that the snapshot of consumption that agents report is a monthly equivalent of their underlying purchases. In the data, I convert consumption to annual amounts; thus consumption is the sum of monthly snapshots over $M = 12$ calendar months. This allows me to address lost time aggregation without confounding it with the timing of consumption in the PSID.

Wage moments. Here I show the main wage moments I target, adjusted for time aggregation.

$$\begin{aligned}
\mathbb{E}\left((\Delta^{\text{psid}} w_{jt}^{\text{year}})^2\right) &= \frac{1}{6}(2M^3 + 3M^2 + M)\sigma^2_{v_j(t)} + M^3\sigma^2_{v_j(t-1)} \\
&+ \frac{1}{6}(2M^3 - 3M^2 + M)\sigma^2_{v_j(t-2)} + M(\sigma^2_{u_j(t)} + \sigma^2_{u_j(t-2)}) \\
\mathbb{E}\left(\Delta^{\text{psid}} w_{jt}^{\text{year}} \times \Delta^{\text{psid}} w_{jt-2}^{\text{year}}\right) &= \frac{1}{6}(M^3 - M)\sigma^2_{v_j(t-2)} - M\sigma^2_{u_j(t-2)} \\
\mathbb{E}\left((\Delta^{\text{psid}} w_{jt}^{\text{year}})^3\right) &= \frac{1}{4}M^2(M+1)^2\gamma_{v_j(t)} + M^4\gamma_{v_j(t-1)} + \frac{1}{4}M^2(M-1)^2\gamma_{v_j(t-2)} \\
&+ M\gamma_{u_j(t)} - M\gamma_{u_j(t-2)} \\
\mathbb{E}\left(\Delta^{\text{psid}} w_{jt}^{\text{year}} \times (\Delta^{\text{psid}} w_{jt+2}^{\text{year}})^2\right) &= \frac{1}{12}(M^4 - M^2)\gamma_{v_j(t)} + M\gamma_{u_j(t)}.
\end{aligned}$$

I use Faulhaber's formula to express the sum of integers. To obtain these expressions, recall that shock increments are serially uncorrelated. It is easy to see how time aggregation causes serial correlation even in the absence of mean-reverting transitory shocks. The covariance between consecutive wage growths reflects the variance of transitory shocks *and* that of permanent shocks. The expressions for all other own- and cross-moments of wages are analogous.

Consumption and earnings moments. Here I show the main earnings and consumption moments, adjusted for time aggregation. To ease the notation, I do not include the conditioning statement $|\mathbf{O}_t$ (or $|\mathbf{O}_{t-2}, \mathbf{O}_t$ when applicable) in the Frisch or Marshallian elasticities; it is implied that all elasticity moments are conditional on the same consumption and hours levels on which the data moments are also conditioned.

$$\begin{aligned}
\mathbb{E}(\Delta^{\text{psid}} w_{1t}^{\text{year}} \times \Delta^{\text{psid}} c_{t+2}^{\text{year}} \mid \mathbf{O}_t) &= \frac{1}{6}(M^3 - M)\sigma^2_{v_1(t)}\mathbb{E}(\kappa_{cv_1(t)}) + \frac{1}{6}(M^3 - M)\sigma_{v_1v_2(t)}\mathbb{E}(\kappa_{cv_2(t)}) \\
&- M\sigma^2_{u_1(t)}\mathbb{E}(\eta_{c,w_1(t)}) - M\sigma_{u_1u_2(t)}\mathbb{E}(\eta_{c,w_2(t)}) \\
\mathbb{E}(\Delta^{\text{psid}} w_{2t}^{\text{year}} \times \Delta^{\text{psid}} c_{t+2}^{\text{year}} \mid \mathbf{O}_t) &= \frac{1}{6}(M^3 - M)\sigma_{v_1v_2(t)}\mathbb{E}(\kappa_{cv_1(t)}) + \frac{1}{6}(M^3 - M)\sigma^2_{v_2(t)}\mathbb{E}(\kappa_{cv_2(t)}) \\
&- M\sigma_{u_1u_2(t)}\mathbb{E}(\eta_{c,w_1(t)}) - M\sigma^2_{u_2(t)}\mathbb{E}(\eta_{c,w_2(t)}) \\
\mathbb{E}(\Delta^{\text{psid}} w_{1t}^{\text{year}} \times \Delta^{\text{psid}} y_{jt+2}^{\text{year}} \mid \mathbf{O}_t) &= \frac{1}{6}(M^3 - M)\sigma^2_{v_1(t)}\mathbb{E}(\kappa_{y_jv_1(t)}) + \frac{1}{6}(M^3 - M)\sigma_{v_1v_2(t)}\mathbb{E}(\kappa_{y_jv_2(t)}) \\
&- M\sigma^2_{u_1(t)}\mathbb{E}(\kappa_{y_ju_1(t)}) - M\sigma_{u_1u_2(t)}\mathbb{E}(\kappa_{y_ju_2(t)}) \\
\mathbb{E}(\Delta^{\text{psid}} w_{2t}^{\text{year}} \times \Delta^{\text{psid}} y_{jt+2}^{\text{year}} \mid \mathbf{O}_t) &= \frac{1}{6}(M^3 - M)\sigma_{v_1v_2(t)}\mathbb{E}(\kappa_{y_jv_1(t)}) + \frac{1}{6}(M^3 - M)\sigma^2_{v_2(t)}\mathbb{E}(\kappa_{y_jv_2(t)}) \\
&- M\sigma_{u_1u_2(t)}\mathbb{E}(\kappa_{y_ju_1(t)}) - M\sigma^2_{u_2(t)}\mathbb{E}(\kappa_{y_ju_2(t)}) \\
\mathbb{E}(\Delta^{\text{psid}} c_{t}^{\text{year}} \times \Delta^{\text{psid}} c_{t+2}^{\text{year}} \mid \mathbf{O}_{t-2}, \mathbf{O}_t) &= \frac{1}{6}(M^3 - M)\sigma^2_{v_1(t)}\mathbb{E}(\kappa^2_{cv_1(t)}) + \frac{1}{6}(M^3 - M)\sigma^2_{v_2(t)}\mathbb{E}(\kappa^2_{cv_2(t)}) \\
&+ \frac{1}{3}(M^3 - M)\sigma_{v_1v_2(t)}\mathbb{E}(\kappa_{cv_1(t)}\kappa_{cv_2(t)}) \\
&- M\sigma^2_{u_1(t)}\mathbb{E}(\eta^2_{c,w_1(t)}) - M\sigma^2_{u_2(t)}\mathbb{E}(\eta^2_{c,w_2(t)}) \\
&- 2M\sigma_{u_1u_2(t)}\mathbb{E}(\eta_{c,w_1(t)}\eta_{c,w_2(t)})
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}\left(\Delta^{\text{psid}} w_{1t}^{\text{obs}} \times (\Delta^{\text{psid}} c_{t+2}^{\text{obs}})^2 \mid \mathbf{O}_t\right) &= \frac{1}{12}(M^4 - M^2)\gamma_{v_1(t)}\mathbb{E}(\kappa_{cv_1(t)}^2) + \frac{1}{12}(M^4 - M^2)\gamma_{v_1 v_2^2(t)}\mathbb{E}(\kappa_{cv_2(t)}^2) \\
&+ \frac{1}{6}(M^4 - M^2)\gamma_{v_1^2 v_2(t)}\mathbb{E}(\kappa_{cv_1(t)}\kappa_{cv_2(t)}) \\
&+ M\gamma_{u_1(t)}\mathbb{E}(\eta_{c,w_1(t)}^2) + M\gamma_{u_1 u_2^2(t)}\mathbb{E}(\eta_{c,w_2(t)}^2) \\
&+ 2M\gamma_{u_1^2 u_2(t)}\mathbb{E}(\eta_{c,w_1(t)}\eta_{c,w_2(t)}) \\
\mathbb{E}\left(\Delta^{\text{psid}} w_{2t}^{\text{obs}} \times (\Delta^{\text{psid}} c_{t+2}^{\text{obs}})^2 \mid \mathbf{O}_t\right) &= \frac{1}{12}(M^4 - M^2)\gamma_{v_1^2 v_2(t)}\mathbb{E}(\kappa_{cv_1(t)}^2) + \frac{1}{12}(M^4 - M^2)\gamma_{v_2(t)}\mathbb{E}(\kappa_{cv_2(t)}^2) \\
&+ \frac{1}{6}(M^4 - M^2)\gamma_{v_1 v_2^2(t)}\mathbb{E}(\kappa_{cv_1(t)}\kappa_{cv_2(t)}) \\
&+ M\gamma_{u_1^2 u_2(t)}\mathbb{E}(\eta_{c,w_1(t)}^2) + M\gamma_{u_2(t)}\mathbb{E}(\eta_{c,w_2(t)}^2) \\
&+ 2M\gamma_{u_1 u_2^2(t)}\mathbb{E}(\eta_{c,w_1(t)}\eta_{c,w_2(t)}) \\
\mathbb{E}\left(\Delta^{\text{psid}} w_{1t}^{\text{obs}} \times (\Delta^{\text{psid}} y_{jt+2}^{\text{obs}})^2 \mid \mathbf{O}_t\right) &= \frac{1}{12}(M^4 - M^2)\gamma_{v_1(t)}\mathbb{E}(\kappa_{y_j v_1(t)}^2) + \frac{1}{12}(M^4 - M^2)\gamma_{v_1 v_2^2(t)}\mathbb{E}(\kappa_{y_j v_2(t)}^2) \\
&+ \frac{1}{6}(M^4 - M^2)\gamma_{v_1^2 v_2(t)}\mathbb{E}(\kappa_{y_j v_1(t)}\kappa_{y_j v_2(t)}) \\
&+ M\gamma_{u_1(t)}\mathbb{E}(\kappa_{y_j u_1(t)}^2) + M\gamma_{u_1 u_2^2(t)}\mathbb{E}(\kappa_{y_j u_2(t)}^2) \\
&+ 2M\gamma_{u_1^2 u_2(t)}\mathbb{E}(\kappa_{y_j u_1(t)}\kappa_{y_j u_2(t)}) \\
\mathbb{E}\left(\Delta^{\text{psid}} w_{2t}^{\text{obs}} \times (\Delta^{\text{psid}} y_{jt+2}^{\text{obs}})^2 \mid \mathbf{O}_t\right) &= \frac{1}{12}(M^4 - M^2)\gamma_{v_1^2 v_2(t)}\mathbb{E}(\kappa_{y_j v_1(t)}^2) + \frac{1}{12}(M^4 - M^2)\gamma_{v_2(t)}\mathbb{E}(\kappa_{y_j v_2(t)}^2) \\
&+ \frac{1}{6}(M^4 - M^2)\gamma_{v_1 v_2^2(t)}\mathbb{E}(\kappa_{y_j v_1(t)}\kappa_{y_j v_2(t)}) \\
&+ M\gamma_{u_1^2 u_2(t)}\mathbb{E}(\kappa_{y_j u_1(t)}^2) + M\gamma_{u_2(t)}\mathbb{E}(\kappa_{y_j u_2(t)}^2) \\
&+ 2M\gamma_{u_1 u_2^2(t)}\mathbb{E}(\kappa_{y_j u_1(t)}\kappa_{y_j u_2(t)}),
\end{aligned}
$$

where the Marshallian elasticities of earnings (the $\kappa_{y_j v_j}$'s) are linear transformations of the Marshallian elasticities of hours given that $\Delta y_{jt} = \Delta w_{jt} + \Delta h_{jt}$; in addition $\kappa_{y_j u_j(t)} = 1 + \eta_{h_j, w_j(t)}$ and $\kappa_{y_j u_{-j}(t)} = \eta_{h_j, w_{-j}(t)}$. All other joint moments are analogous.

Identification. Identification of the wage parameters is similar to the baseline: the covariance between current wage growth and the sum of three consecutive growth rates identifies the variance of permanent shocks. Given this, the covariance between consecutive wage growths identifies the variance of transitory shocks (see Crawley, 2020, for a similar income process).

Identification of preferences is different from the baseline because the targeted earnings and consumption moments now also depend on moments of the transmission parameters of *permanent* shocks (i.e. the lifecycle Marshallian elasticities) that the baseline does not depend on. One can see this from $\mathbb{E}(\Delta^{\text{psid}} c_t^{\text{year}} \times \Delta^{\text{psid}} c_{t+2}^{\text{year}} \mid \mathbf{O}_{t-2}, \mathbf{O}_t)$ or other moments. While this identifies second moments of the consumption Frisch elasticities in the baseline (the η_{c,w_j}'s), here this depends also on moments of the Marshallian elasticities (the κ_{cv_j}'s). Of course, moments of the Marshallian elasticities are identified from concurrent wage, earnings, and consumption moments. However, such moments are not useful for identification of the underlying preferences under heterogeneity (sections 3.3 and 4.2), so the baseline does not target them.

There are two ways in which identification can proceed here. One is to also target *concurrent* wage, earnings, and consumption moments, obtain moments of the Marshallian elasticities,[9] and then identify moments of Frisch elasticities from *consecutive* wage, earnings, and consumption growth rates. The downside is that the empirical target of this estimation is markedly different from the baseline, so differences between the parameter estimates here and the baseline cannot be attributed exclusively to time aggregation. An alternative is to fix the Marshallian elasticities at some reasonable value, and identify the moments of Frisch elasticities from consecutive wage and outcome growth rates. The empirical target of the latter estimation remains as in the baseline.

Estimation and results. I follow the second approach above and fix the Marshallian elasticities. The value at which one fixes the κ's likely matters for the results so one must check the robustness of the results to alternative values.

Estimation proceeds in two steps. In the first step, I estimate the parameters of the wage process; these are independent of the Marshallian elasticities. In the second step, and conditional on the wage parameters, I estimate the moments of Frisch elasticities over a *grid* of alternative values for the κ's. There are six lifecycle Marshallian elasticities; these are κ_{cv1}, κ_{cv2}, κ_{y1v1}, κ_{y1v2}, κ_{y2v1}, and κ_{y2v2}. All first and second moments of these elasticities appear in the estimation amounting to 6 first moments, 6 variances, and 15 covariances. This is a total of 27 parameters for which I need to loop over possible values. Even with a conservative grid of, say, 5 values for each parameter, the curse of dimensionality is insurmountable.

To overcome this, I fix several covariances to zero and thus mostly loop over values for the averages and variances only. Specifically, I loop over four different values for each *average* Marshallian elasticity; in each case the grid contains equidistant points over an interval around the corresponding elasticity from BPS. For example, BPS find that $\kappa_{cv1} = 0.34$ homogeneously; my grid contains the values 0.05, 0.22, 0.38, 0.55. Neither the density of the grid, nor small deviations from the endpoints matter. There are no external estimates that can inform the grid for the *variance* of the Marshallian elasticities, so in this case I use arbitrary values mostly between zero and 0.4. The upper bound is higher than the variance of any Frisch elasticity in the baseline. Finally, I loop over certain covariances, namely $\mathbb{E}(\kappa_{cv_1}\kappa_{cv_2})$ and $\mathbb{E}(\kappa_{y_jv_1}\kappa_{y_jv_2})$; these are most likely positive because of the underlying structure of the problem (the Marshallian in each pair is a function of the same underlying Frisch elasticities).

Each instance of this large nested loop has values for the moments of the Marshallian elasticities that appear in the model. Given said values, I estimate the preferred specification of the model in order to obtain estimates for the Frisch elasticities. It is computationally very expensive to bootstrap each instance of the loop, so I do not calculate standard errors; however, all other estimation details are exactly as in the baseline.

[9] One can obtain moments of the Marshallian elasticities in a way analogous to the identification of moments of permanent shocks; see Crawley (2020) for the first moment.

Table G.1 – Robustness: Addressing Time Aggregation

	preferred specification				
parameter:	at lowest $fval$	average over loops	st. dev. over loops	25^{th} percentile	75^{th} percentile
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.020	0.008	0.017	-0.007	0.022
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.032	0.013	0.027	-0.011	0.035
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.228	0.238	0.015	0.225	0.250
$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.260	0.305	0.036	0.279	0.328
$\text{Var}(\eta_{c,w_1(i)})$	0.086	0.120	0.019	0.104	0.133
$\text{Var}(\eta_{c,w_2(i)})$	0.220	0.306	0.049	0.267	0.341
$\text{Var}(\eta_{h_1,w_1(i)})$	0.021	0.020	0.008	0.014	0.025
$\text{Var}(\eta_{h_2,w_2(i)})$	0.003	0.009	0.007	0.004	0.012
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{c,w_2(i)})$	0.137	0.191	0.031	0.167	0.213
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_1,w_1(i)})$	0.037	0.037	0.012	0.028	0.045
$\text{Cov}(\eta_{c,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.007	-0.023	0.016	-0.034	-0.010
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_1,w_1(i)})$	0.058	0.059	0.019	0.045	0.071
$\text{Cov}(\eta_{c,w_2(i)}, \eta_{h_2,w_2(i)})$	-0.012	-0.037	0.026	-0.054	-0.017
$\text{Cov}(\eta_{h_1,w_1(i)}, \eta_{h_2,w_2(i)})$	-0.004	-0.010	0.007	-0.013	-0.004

Notes: Observations: 6071 (households $\times\Delta_t$). The table presents summary statistics of estimates from the preferred specification over various values for the moments of Marshallian elasticities. All parameters are conditional on $\mathbf{O}_t = \mathbb{E}(\mathbf{O}_{it})$.

Given the loop, I obtain a large number of estimates for the Frisch elasticities. Here I report only a summary of the results. Specifically, I report the estimates associated with the lowest value of the GMM distance metric, the average and standard deviation of estimates over all loops, and the 25^{th} and 75^{th} percentile of estimates across the loops.

The results are in table G.1. Despite time aggregation changing the target moments substantially, the baseline conclusions remain unchanged. The average consumption elasticities are near zero, the average labor supply elasticities are close to the baseline (around 0.25-0.30), consumption preference heterogeneity remains substantial, heterogeneity in the labor supply elasticities is about an order of magnitude lower. The standard deviation of the estimates for $\text{Var}(\eta_{h_j,w_j(i)})$ is low, suggesting that this pattern of heterogeneity is present across different values for the moments of the Marshallian elasticities. Overall, the results are in line with those obtained from the non-aggregated wage process in the baseline.

Discussion. Time aggregation here is subject to certain limitations. The underlying wage shocks are assumed to hit the household only in 12 instances in a year; while this seems more realistic than one that assumes wage shocks appear on January 1^{st} only, it is ultimately ad hoc. Moreover, transitory shocks are assumed to last for only one month, while they most certainly last longer. Crawley (2020) reports that the time aggregated results in his setting are not very sensitive to this assumption. A more serious issue may be the exact timing of consumption in the PSID. Of course, lack of additional information restricts what one can ultimately do and there is a point where one must make choices balancing generality and tractability.

H Simulation of Quantitative Model

The model in WK is a parametric version of the lifecycle model for consumption and family labor supply that I developed in section 2. Their baseline parametrization imposes separability between consumption and family labor supply. Given heterogeneity in the consumption-wage elasticities η_{c,w_j} in my baseline, I work directly with the non-separable parametrization that WK present as an extension. The utility function is given by

$$U_t(C_{it}, H_{1it}, H_{2it}) = \beta^t(1-\sigma)^{-1}\left(\{\alpha C_{it}^{\gamma} + (1-\alpha)[\xi H_{1it}^{\theta} + (1-\xi)H_{2it}^{\theta}]^{-\frac{\gamma}{\theta}}\}^{\frac{1-\sigma}{\gamma}} - 1\right) \tag{H.1}$$

where β is the discount factor, σ governs the intertemporal substitution of consumption, γ governs the substitution between consumption and labor supply, and θ governs the substitution between spousal hours. This utility function allows for flexible substitution between consumption, male hours, and female hours, although it is *not* flexible enough to generate independent variation in all 9 underlying Frisch elasticities of consumption and hours.

A first difference between my model and WK is that the specification above abstracts from preference heterogeneity. All utility parameters are common across households. However, one can accommodate heterogeneity indirectly, namely by solving the model for different sets of parameters, simulating random households for each set, and pooling all households together to form one large population that exhibits preference heterogeneity.

A second difference is that WK restrict shocks to joint normal. Therefore, $\gamma_{v_j(t)} = 0$ and $\gamma_{u_j(t)} = 0$, and similarly for all other third moments. In my baseline estimation, I have found substantial left skewness in wage shocks. Abstracting from such skewness, the model in WK is unable to generate skewness in earnings and consumption (thus contradicting the data – table 2). Estimation of the variances and covariances of Frisch elasticities relies on the third moments of outcomes conveying information about preferences. With those moments restricted to zero, I cannot use WK's setting to estimate the second moments of preferences.

Allowing for skewness in wage shocks is not simply a matter of replacing the model's normal distribution by, say, the skew normal or a mixture of normals. Several complications arise. First, one needs the statistical tools to calculate multivariate skewed distribution functions (cdf and pdf). In the absence of analytical expressions for them, one must resort to numerical integration which, given the state space, may become prohibitively expensive. Second, a long tail in income requires sufficient support in the state space for the model to be able to 'visit' such extreme events. Replacing WK's joint normal with the bivariate skew normal (Azzalini and Valle, 1996) failed to generate *any* skewness in earnings or consumption without

Table H.1 – Quantitative Model: Heterogeneity in β and $\eta_{c,p}$

	true	transitory		true	transitory
$\mathbb{E}(\eta_{c,w_1(i)})$	-0.151	-0.146	$\mathbb{E}(\eta_{h_1,w_2(i)})$	0.145	0.125
$\mathbb{E}(\eta_{c,w_2(i)})$	-0.065	-0.060	$\mathbb{E}(\eta_{h_2,w_1(i)})$	0.318	0.301
$\mathbb{E}(\eta_{h_1,w_1(i)})$	0.818	0.806	$\mathbb{E}(\eta_{h_2,w_2(i)})$	0.646	0.649

Notes: The table presents GMM estimates of first moments of wage elasticities based on simulated data for four sub-populations of households from the parametric model of WK. In addition to heterogeneity in the parameters governing the estimable Frisch elasticities, households are heterogeneous in the discount factor β and the consumption substitution elasticity $\eta_{c,p}$.

substantially modifying the support of the state space. Third, the solution algorithm must be made proof to realizations in the tails of the state space, often associated with numerical instability. In practice, and despite increasing evidence for skewness in earnings dynamics, few only papers have deviated from normality to date (for a discussion see De Nardi et al., 2019). I do not attempt this extension here so I can only use the simulated data to estimate *average* elasticities under heterogeneity (but not second moments). I can also check if the implications of heterogeneity for inequality or partial insurance are broadly in line with my baseline.

Unless explicitly said otherwise, I keep the calibration of the non-separable model precisely as in the original paper. WK's budget constraint allows for joint taxation; my baseline abstracts from this so I set all tax parameters to zero. Except cols. 4 & 5 in table 9, I simulate four sub-populations of 50000 households each, and I pool them together to form a larger population of households. In cols. 4 & 5, I draw a sample of 1508 observations from each sub-population to create a dataset of 6032 household $\times \Delta t$ observations as in the baseline sample.

The four sub-populations differ in parameters γ and θ that control the substitution between consumption, male hours, and female hours, thus contributing to heterogeneity in the consumption-wage and labor supply elasticities. The first sub-population in all cases has $\gamma = -3$ and $\theta = 3$ (the choice in WK), the second has $\gamma = -2$ and $\theta = 4$, the third $\gamma = -3.5$ and $\theta = 2.3$, and the fourth has $\gamma = -3.9$ and $\theta = 3.5$. While these choices are ultimately arbitrary, note that it has not been possible to try parameter values that are too far away from WK's calibration without running into numerical instability.

In addition to this heterogeneity, the sub-populations for columns 1 & 2 in table 10 differ over $\beta = \{0.99; 0.98; 0.97; 0.95\}$. Those in columns 3 & 4 differ over $\sigma = \{2.24; 3; 2.1; 2.8\}$. Heterogeneity in σ induces heterogeneity in all Frisch elasticities. Finally, table H.1 has $(\beta, \sigma) = \{(0.99, 2.24); (0.99, 3), (0.98, 2.24), (0.98, 3)\}$. All choices are again arbitrary, subject to being 'close enough' to WK's default calibration at $(0.99, 2.24)$ in order to avoid numerical instability.

Welfare loss from idiosyncratic wage risk. I calculate this as the root CEV of the equation $W((1+CEV)\boldsymbol{C}^0, \boldsymbol{H}_1^0, \boldsymbol{H}_2^0) = W(\boldsymbol{C}^1, \boldsymbol{H}_1^1, \boldsymbol{H}_2^1)$ where $(\boldsymbol{C}^0, \boldsymbol{H}_1^0, \boldsymbol{H}_2^0)$ is the allocation of outcomes without wage risk and $(\boldsymbol{C}^1, \boldsymbol{H}_1^1, \boldsymbol{H}_2^1)$ is the allocation with risk. CEV measures

Table H.2 – Welfare Loss from Idiosyncratic Wage Risk

specification:	preferred	BPS
Total welfare change	-16.6%	-14.6%
A. Consumption (level, distribution)	-15.4% (2.1%, -17.1%)	-12.5% (4.1%, -15.9%)
B. Male hours (level, distribution)	-0.6% (0.8%, -1.4%)	0.0% (2.6%, -2.5%)
C. Female hours (level, distribution)	-0.8% (0.5%, -1.3%)	-2.4% (0.2%, -2.6%)

Notes: The table presents welfare losses from idiosyncratic wage risk in the average household (one with average preferences) and a decomposition to consumption and hours. See appendix B in WK and text for details on these calculation.

the % of lifetime consumption that households are willing to give up in order to maintain the same expected lifetime utility in a world without male and female wage risk.

With preference heterogeneity, the welfare function W and aggregation across households matter. The best treatment would be to solve and simulate the model for an underlying distribution of households that reflects my baseline parameter estimates, namely the preferences at which I want to calculate the welfare losses from wage risk. That would enable me to obtain the *distribution* of welfare losses across households. This is, however, hard to do for two reasons: (1.) solving the model for a realistic distribution of preferences is computationally tedious; (2.) unlike the previous part in which preference heterogeneity was ad hoc, the distribution of preferences here should reflect my baseline estimates. This requires *estimation* of the non-separable model allowing for preference heterogeneity.

To avoid these complications, I follow WK and simply calculate the welfare loss for the *average* household, i.e. the household with average preferences from the preferred specification in column 8 of table 4. I also compare the losses with those for a household that has preferences as in BPS (column 2, table 4). Given that average preferences in these two cases are separable, the Frisch elasticities become deep parameters. I can thus directly plug in the parameters from table 4 without estimating the quantitative model.[10]

The results are in the first row of table H.2. Households are willing to give up about 17% of their lifetime consumption in order to avoid wage risk. This is 2 percentage points larger than the cost in BPS, simply because households here are more exposed to risk given the limited insurance role of family labor supply. It is also closer by 20% to welfare losses with exogenous labor supply (the latter are about 25% of lifetime consumption - table 6 in WK). The welfare losses stem predominantly from consumption as wage risk translates into consumption risk; this is stronger in the preferred specification where the role of labor supply is limited. Details of this decomposition are in appendix B of WK.

[10]Given that average preferences are separable, I use WK's separable specification for utility, namely $U_t = \beta^t\big((1-\sigma)^{-1}C_{it}^{1-\sigma} - \psi_1(1-1/\eta_1)^{-1}H_{1it}^{1-1/\eta_1} - \psi_2(1-1/\eta_2)^{-1}H_{2it}^{1-1/\eta_2}\big)$, instead of specification (H.1). I allow for taxes precisely as in WK's welfare calculations.

LUXEMBOURG INSTITUTE OF
SOCIO-ECONOMIC RESEARCH

11, Porte des Sciences
Campus Belval
L-4366 Esch-sur-Alzette

T +352 58 58 55-1
F +352 58 58 55-700

www.liser.lu