I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14226

Accounting for Limited Commitment between Spouses When Estimating Labor-Supply Elasticities

Christian Bredemeier
Jan Gravert
Falko Juessen

MARCH 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14226

Accounting for Limited Commitment between Spouses When Estimating Labor-Supply Elasticities

Christian Bredemeier
University of Wuppertal and IZA

Jan Gravert
University of Wuppertal

Falko Juessen
University of Wuppertal and IZA

MARCH 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Accounting for Limited Commitment between Spouses When Estimating Labor-Supply Elasticities

The Frisch elasticity of labor supply can be estimated by regressing hours worked on the hourly wage rate, controlling for consumption of the individual worker. However, most household panel surveys contain consumption information only at the household level. We show that proxying individual consumption by household consumption biases estimated Frisch elasticities downward as limited commitment in the household induces individual consumption to behave differently from household consumption. We develop an improved estimation approach that eliminates this bias by exploiting information on the composition of household consumption to infer its distribution. Using PSID data, we estimate Frisch elasticities of about 0.7.

JEL Classification: D13, D15, J12, J22, E21, E24

Keywords: labor-supply elasticity, limited commitment, intra-household decision making, couple households, consumption

Corresponding author:
Christian Bredemeier
Schumpeter School of Business and Economics
University of Wuppertal
Gaußstraße 20
42119 Wuppertal
Germany

E-mail: bredemeier@uni-wuppertal.de

1 Introduction

The Frisch elasticity of labor supply is an important concept in both labor economics and macroeconomics. It measures how willingly individuals substitute hours worked intertemporally. The Frisch elasticity determines labor-supply responses to temporary changes in wage rates and to predictable life-cycle patterns in wage rates. It is also a decisive determinant of the cost of business cycles and the fiscal multiplier. Despite its importance, there is no consensus about a range of values for the Frisch elasticity with estimates from microeconometric studies diverging substantially from the results of macroeconomic studies, see, e.g., Keane and Rogerson (2015).

The Frisch elasticity is defined as the reaction of labor supply to changes in wage rates holding marginal utility of wealth constant. Hence, it can be estimated in a regression of hours worked on the wage rate when controlling for consumption, the latter being closely tied to the marginal utility of wealth, see, e.g., Altonji (1986). Building on recent advances in family economics emphasizing that commitment within the household is limited (Voena 2015, Chiappori and Mazzocco 2017, Ábrahám and Laczó 2018), we show that the relevant consumption variable would be consumption of the individual worker, which is, however, usually not observed in household panel surveys, and that using household consumption instead of individual consumption biases the estimated Frisch elasticity.[1] Due to limited commitment, individual household members' shares in total household consumption vary and reflect members' varying bargaining positions for which the respective wage rates are important determinants. Omitting information on how consumption is distributed between household members thus leads to a bias. Using simulated data from a quantitative model of household decision making with limited commitment between spouses, we find that a standard labor-supply regression yields an estimated Frisch elasticity that is about 20% below its true value when household consumption rather than individual consumption is controlled for.

The intuition behind the estimation bias is as follows. When commitment between spouses is limited, a temporary rise in an individual's hourly wage rate exerts three effects on the individuals' labor supply. First, labor supply increases due to a conventional substitution effect that is governed by the Frisch elasticity and can be used to recover this elasticity. Second, household consumption increases which induces labor supply to fall due to a wealth effect. Third, the individual's bargaining position in the household may increase, reflecting the improvement of the individual's outside option. This leads the household to grant more

[1] There is strong empirical evidence supporting the limited-commitment approach to family decision making, see Dercon and Krishnan (2000), Duflo and Udry (2004), Mazzocco (2007), Robinson (2012), Cesarini et al. (2017) and Lise and Yamada (2019).

leisure to the individual thereby reducing the individual's labor supply through a limited-commitment effect. Both the income effect and the limited-commitment effect could be accounted for with data on individual consumption, but data on household consumption can only account for the income effect but not the limited-commitment effect.

We use our model to develop an improved estimation approach that eliminates the bias and can be applied when data on individual consumption is not available. The key idea of our approach is that information on *who* consumes how much of total household consumption can be inferred from *what* the household purchases. Information on the latter is provided in household panel surveys such as the Panel Study of Income Dynamics (PSID) which covers a variety of different consumption items. The limited-commitment paradigm implies that relative bargaining weights determine the distribution of consumption between household members and, when preferences over different consumption goods differ across household members, also how much a household spends on the different goods. The (observable) composition of household consumption is thus informative about its (unobservable) distribution. We show analytically that in a labor-supply regression which also controls for the composition of household consumption, i.e., households' expenditure shares on different goods and services, and not just their total consumption expenditures, the coefficient on the hourly wage rate is an almost unbiased estimate of the Frisch elasticity.[2]

We perform Monte-Carlo experiments using synthetic data from a calibrated model version to show that our approach all but eliminates the bias, with estimates deviating less than 1% from the true value even for small differences in preferences between household members. We then apply our approach to household panel data from the PSID. The results are strongly supportive of our theoretical results and our improved estimation approach yields a significantly larger estimate for the Frisch elasticity. In our specifications that correct for the bias due to limited commitment, the estimated Frisch elasticity is around 0.7.

Our paper contributes to both, the literature on estimating labor-supply elasticities and the literature on limited commitment between spouses in marriages. The latter has been surveyed by Chiappori and Mazzocco (2017).[3] Keane (2011) provides a comprehensive survey of the literature on estimating labor-supply elasticities. The micro/macro puzzle on the elasticity of labor supply is a central question in this literature, see Keane and Rogerson

[2]Our approach does not require an a-priori assumption about which spouse prefers which type of consumption goods, it is only necessary that spouses' preferences over the included consumption items *differ*.

[3]A large part of this literature documents empirical evidence for limited commitment, i.e., rejects the competing full-commitment paradigm, see, e.g., Mazzocco (2007), Robinson (2012), and Lise and Yamada (2019). A second strand of the literature studies the consequences of limited commitment, see, e.g., Mazzocco et al. (2014), Voena (2015), and Ábrahám and Laczó (2018). While most papers in this second group focus on welfare or policy effects, we study the consequence of limited commitment for labor-supply estimations.

(2015) for an overview. Several studies have contributed to reconciling micro and macro estimates by pointing out a number of downward biases in microeconometric estimates, see, e.g., Blomquist (1985, 1988), Alogoskoufis (1987), Heckman (1993), Rupert et al. (2000), Imai and Keane (2004), and Domeij and Floden (2006). Our empirical analysis will take these biases into account and our paper contributes to this literature by showing that limited commitment between spouses – if not appropriately corrected for – also biases estimates of the Frisch elasticity downward.

While our econometric approach improves estimations of the labor-supply condition in *levels*, Altonji (1986) proposed an alternative approach that estimates the labor supply condition in *growth rates* and identifies the Frisch elasticity as the coefficient on expected wage growth. The latter is uncorrelated with changes in marginal utility when households have sufficient access to insurance or borrowing possibilities. The approach using growth rates is not affected by problems due to limited commitment in the household because bargaining mostly reacts to unexpected changes in wage rates while the Frisch elasticity is identified through changes in expected wage rates in this approach. However, this approach faces two substantial challenges in empirical applications. First, consumption insurance possibilities are limited in the real world and a substantial share of households has close to no wealth (see Kuhn and Ríos-Rull 2016) reducing their ability to self-insure. Domeij and Floden (2006) have shown that this results in a substantial downward bias in the estimated Frisch elasticity – which Bredemeier et al. (2019) have shown can be overcome by exploiting the double-earner structure of most households. The second challenge is that the Altonji (1986) approach requires information on expected wage growth which is usually obtained using instrumental variables. But wage growth – in contrast to wage levels which are used in this paper – is hard to predict and hence instruments tend to be weak, see Keane (2011) for a discussion. Our bias-corrected estimate for the Frisch elasticity is at the upper bound of estimates presented in previous studies and close to the values reported in previous studies that have used expected wage growth instead of consumption information in levels, see Pistaferri (2003), Domeij and Floden (2006), and Bredemeier et al. (2019).

The remainder of this paper is organized as follows. In Section 2, we develop a limited-commitment model of couple households, show analytically that conventional labor-supply regressions are biased, and derive an unbiased estimation approach. In Section 3, we perform Monte-Carlo experiments using synthetic data from a calibrated version of our model to quantify the bias in conventional estimations and to evaluate our estimation approach. In Section 4, we apply our estimation approach to PSID data. Section 5 concludes.

2 Theory

We consider couple households whose members take decisions cooperatively and act under a joint budget constraint. There are two goods which are consumed privately by individual household members and there is one non-state contingent asset in which they invest jointly but cannot go short. Members of the household face stochastic wage rates and each individual member has the option to leave the household unilaterally at any time. They would do so whenever the outside option of single life is preferable to life within the household. Spouses acknowledge this possibility in their bargaining process and understand that they have no possibility to enforce promises made by the partner. This gives rise to limited commitment, as in, among others, Mazzocco (2007), Voena (2015), and Ábrahám and Laczó (2018). Any bargained plan over savings, consumption, and labor supply has to ensure that partners never actually want to leave the household. This can be understood as a self-enforcing contract between spouses as, under such a plan, incentives are always as such that spouses want to stay in the household - or, more technically, that participation constraints hold at any time.

Marcet and Marimon (2019) have shown that a decision problem with limited commitment can be represented as a simple Pareto planning problem with time-varying weights on individual utility functions. Whenever one spouse has an incentive to leave the household, i.e., his or her participation constraint is binding, his or her Pareto weight increases such that his or her participation constraint holds with equality. This is achieved by adjusting the Pareto weight of the spouse by exactly the Kuhn-Tucker multiplier on the respective participation constraint, see Marcet and Marimon (2019).

2.1 A simple labor-supply model with limited commitment

The notation is as follows. An individual agent is identified by the index gj, where $g \in \{m, f\}$ denotes the individual's gender (m = male, f = female) and j denotes the household in which the individual lives. Individual gj forms a household with an individual of the opposite gender, denoted by $-gj$. Household wealth is denoted by a_j. w_{gj} is individual gj's wage rate, μ_{gj} her Pareto weight, $c_{gj,k}$ her consumption of good $k = A, B$, and n_{gj} her hours worked. $S_{gj}(w_{gj}, a_{gj})$ is the individual's expected lifetime utility if she were single and had wealth a_{gj} such that $S_{gj}(w_{gj}, a_j/2)$ is her expected lifetime utility in case of divorce, where the individual would obtain half the household's wealth, and hence the relevant outside option to the household bargaining problem (following Voena 2015).[4] V_{gj} is the individual's

[4] The maximization problem of the single household can be found in Appendix A.1.

expected lifetime utility within the couple household. β is the rate of time preference and r the exogenous interest rate. Throughout, variables with a prime denote next period values.

For given state variables $\Omega_j = (w_{mj}, w_{fj}, \mu_{mj}, \mu_{fj}, a_j)$, the couple chooses $\Gamma_j = (c_{mj,A}, c_{fj,A}, c_{mj,B}, c_{fj,B}, n_{mj}, n_{fj}, a'_j)$ and the planning problem reads

$$\begin{aligned} V_j(\Omega_j) = \max_{\Gamma_j} \; & v\left(c_{mj,A}, c_{fj,A}, c_{mj,B}, c_{fj,B}, n_{mj}, n_{fj}, \mu'_{mj}, \mu'_{fj}\right) \\ & - \phi_{mj} \cdot S_{mj}(w_{mj}, a_j/2) - \phi_{fj} \cdot S_{fj}(w_{fj}, a_j/2) + \beta \cdot \mathrm{E}\left[V_j\left(\Omega'_j\right)\right], \end{aligned}$$

with the household target function given by

$$v = \mu'_{mj} \cdot u_{mj}(c_{mj,A}, c_{mj,B}, n_{mj}) + \mu'_{fj} \cdot u_{fj}(c_{fj,A}, c_{fj,B}, n_{fj}),$$

subject to the joint budget constraint

$$c_{mj,A} + c_{mj,B} + c_{fj,A} + c_{fj,B} + a'_j \leq w_{mj} n_{mj} + w_{fj} n_{fj} + (1+r)\, a_j, \tag{1}$$

and the joint borrowing constraint

$$a'_j \geq 0, \tag{2}$$

with the Pareto weights being updated according to

$$\mu'_{gj} = \mu_{gj} + \phi_{gj},$$

where ϕ_{gj} is the Kuhn-Tucker multiplier on this period's participation constraint

$$V_{gj}(\Omega_j) \geq S_{gj}(w_{gj}, a_j/2)$$

for individual gj. Initial Pareto weights are determined by Nash (1950) bargaining.[5]

Individual preferences of agent gj are represented by

$$u_{gj}(c_{gj}, n_{gj}, I_j) = \frac{c_{gj}^{1-\sigma_g} - 1}{1-\sigma_g} - \alpha_{gj} \cdot \frac{n_{gj}^{1+1/\eta_g}}{1+1/\eta_g} + \Psi_j \cdot I_j,$$

where c_{gj} is total consumption of individual gj (an aggregate of two consumption goods, see below), σ_g and η_g are the rate of risk aversion and the Frisch elasticity of labor supply for individuals of gender g, respectively, I_j is an indicator variable for staying in the household,

[5]Details are provided in Appendix A.2.

and Ψ_j is a direct utility gain from being married, which may differ across households. The latter term captures all gains and losses from marriage not explicitly modeled such as love, companionship, intimacy, children, returns to scale in consumption, home production, etc., as well as avoided costs and foregone benefits of divorce. In our quantitative model, we will calibrate Ψ_j to match the extent of commitment issues to an empirical target.

Aggregate consumption c_{gj} consists of two goods A and B over which spouses have different preferences,

$$c_{gj} = \gamma_g^{\gamma_g} (1 - \gamma_g)^{1-\gamma_g} \cdot c_{gj,A}^{\gamma_g} c_{gj,B}^{1-\gamma_g},$$

where $\gamma_g \in (0, 1)$ with $\gamma_g \neq \gamma_{-g}$ are preference weights. Differences in preferences over the two goods will enable us to exploit information on household expenditures over the two goods to infer how household consumption is distributed over the individual members. We choose the units in which the two goods are measured such that their prices are the same and normalized to $P_A = P_B = 1$. This implies that the price of the consumption bundle c is the same for both genders, $P_g = P_A^{\gamma_g} P_B^{1-\gamma_g} = P_{-g} = P = 1$.[6]

Wage rates evolve according to a stochastic process with probability function $f(w'_{mj}, w'_{fj} | w_{mj}, w_{fj})$. Our analytical results do not depend on a particular specification of the wage process. In our calibrated model, we will assume that wage rates follow stationary first-order autoregressive processes with heterogenous fixed components.

For our analytical results, it is important that the value of an individual's outside option depends on her wage rate. This is responsible for the bias in conventional labor-supply estimations as wage raises improve bargaining positions which tends to increase leisure time and reduce labor supply. Otherwise, our derivations do not depend on the details of the outside option. For our quantitative model, we will assume that divorcees have access to the same saving technology as have married individuals and the evolution of their wage rates is not affected by the change in their marital status. Following Mazzocco (2007) and Ábrahám and Laczó (2018), we assume in the quantitative analysis that the outside option of a spouse is being a single for the rest of her life. This assumption is for computational simplicity and a more elaborated modelling of the outside option would yield the same implications in our context. Specifically, giving divorcees the possibility to remarry (as in Voena 2015) would raise the value of the outside option and lead us to recalibrate the direct utility value

[6] While the model abstracts from price shifts between different consumption goods, such shifts occur empirically. Theoretically, this raises the possibility that a shift in the household consumption bundle toward one good does not reflect rising bargaining power of the spouse who prefers this good but results from this good becoming relatively cheaper. In our empirical analysis, price shifts between consumption goods are accounted for by time fixed effects.

from marriage, Ψ_j, accordingly. Since we model being divorced as an absorbing state, the calibrated utility gain Ψ_j also includes the negative value of remarriage possibilities as a divorcee.

The first-order conditions of couple j's decision problem are

$$\lambda_j = \mu'_{mj} \cdot c_{mj}^{-\sigma_m} = \mu'_{fj} \cdot c_{fj}^{-\sigma_f}, \tag{3}$$

$$c_{gj,A} = \gamma_g \cdot c_{gj}, \tag{4}$$

$$c_{gj,B} = (1-\gamma_g) \cdot c_{gj}, \tag{5}$$

$$w_{mj} = \mu'_{mj} \cdot \frac{n_{mj}^{1/\eta_m}}{\lambda_j}, \tag{6}$$

$$w_{fj} = \mu'_{fj} \cdot \frac{n_{fj}^{1/\eta_f}}{\lambda_j}, \tag{7}$$

$$\lambda_j - \xi_j = \beta \cdot \mathrm{E}\left[(1+r) \cdot \lambda'_j - \frac{1}{2} \cdot \phi'_{mj} \cdot \frac{\partial S_m}{\partial a'_{mj}} - \frac{1}{2} \cdot \phi'_{fj} \cdot \frac{\partial S_f}{\partial a'_{fj}}\right], \tag{8}$$

the budget constraint (1) and the borrowing constraint (2) with its corresponding Kuhn-Tucker conditions, given the wage rate realizations of the spouses, w_{mj} and w_{fj}, current asset holdings a_j, and the previous-period Pareto weights μ_{mj} and μ_{fj}. λ_j denotes the Lagrange multiplier on the budget constraint and ξ_j denotes the Kuhn-Tucker multiplier on the borrowing constraint. The ratio of marginal utilities of consumption of the spouses may vary over time due to potential time variation in the Pareto weights which can be seen from the risk-sharing condition (3). Thus, intra-household risk-sharing is imperfect since commitment is limited. The composition of the aggregated consumption bundle, c_{gj}, is described by (4) and (5). The labor-supply conditions are given by (6) and (7). These conditions also reflect spouses' varying bargaining powers μ'_{gj}. The right-hand side of the Euler equation (8) includes two additional terms compared to a standard Euler equation in the presence of borrowing constraints. These terms capture the impact of savings on the likelihood of binding participation constraints in the future and thereby the impact on changes in future Pareto weights.

Under full commitment, the static first-order conditions take the same form, but the Pareto weights are determined once and for all such that participation constraints at the time of household formation are fulfilled. In this case, the Euler equation simplifies to the standard form $\lambda_j - \xi_j = \beta \cdot \mathrm{E}((1+r) \cdot \lambda'_j)$ as Pareto weights are constant, $\phi'_{mj} = \phi'_{fj} = 0$.

2.2 The estimation problem

We now use the first-order conditions of the model to derive conditions which can be used to estimate the Frisch elasticity of labor supply. We introduce time indices to clarify the panel dimension of the estimation. Rearranging (6) or (7), respectively, the labor-supply condition of agent gj in period t is given by

$$n_{gjt} = \left(\frac{w_{gjt} \cdot \lambda_{jt}}{\alpha_{gj} \cdot \mu_{gjt}} \right)^{\eta_g},$$

where we define $\mu_{gjt} = \mu'_{gj}$ for notational convenience. Using the first-order conditions for individual consumption, $\lambda_{jt} = \mu_{gjt} \cdot c_{gjt}^{-\sigma_g}$, and taking logs, the labor-supply condition becomes

$$\log n_{gjt} = -\eta_g \cdot \log \alpha_{gj} + \eta_g \cdot \log w_{gjt} - \eta_g \cdot \sigma_g \cdot \log c_{gjt}. \tag{9}$$

Hence, if one had access to data on individual consumption, one could simply estimate

$$\log n_{gjt} = \delta_{0,gj} + \delta_{1,g} \cdot \log w_{gjt} + \delta_{2,g} \cdot \log c_{gjt} + \varepsilon_{gjt}, \tag{10}$$

where $\delta_{1,g}$ and $\delta_{2,g}$ are regression coefficients, $\delta_{0,gj}$ is an individual fixed effect, and ε_{gjt} is a residual stemming from measurement errors. Estimating equation (10) would yield an unbiased estimate, $\hat{\delta}_{1,g}$, for the Frisch labor-supply elasticity, $\mathrm{E}\,\hat{\delta}_{1,g} = \eta_g$.

However, usually only household consumption is observed in household panel data and, hence, equation (10) can only be estimated using household consumption,

$$\log n_{gjt} = \kappa_{0,gj} + \kappa_{1,g} \cdot \log w_{gjt} + \kappa_{2,g} \cdot \log c_{jt} + \tilde{\varepsilon}_{gjt}, \tag{11}$$

Estimating (11) instead of (10) results in an estimation bias when household consumption and individual consumption are not perfectly correlated.

We now derive this bias analytically and discuss how it depends on the degree of commitment in the household. The risk-sharing condition (3) can be used to link household consumption to consumption of one of the spouses. Specifically, household consumption, $c_{jt} = c_{mjt} + c_{fjt}$, can be written as

$$c_{jt} = c_{gjt} + x_{gjt}^{-1/\sigma_{-g}} \cdot c_{gjt}^{\sigma_g/\sigma_{-g}}, \tag{12}$$

where $x_{gjt} = \mu_{gjt}/\mu_{-gjt}$ denotes the relative Pareto weight of individual gj and the second term on the right-hand side is c_{-gjt} according to (3). A first-order Taylor approximation of

(12) yields[7]

$$\log c_{jt} \approx \left(\frac{\overline{c}_g}{\overline{c}} + \frac{\sigma_g}{\sigma_{-g}} \cdot \frac{\overline{c}_{-g}}{\overline{c}}\right) \cdot \log c_{gjt} - \frac{1}{\sigma_{-g}} \cdot \frac{\overline{c}_{-g}}{\overline{c}} \cdot \log x_{gjt}, \tag{13}$$

where $\overline{c}_g$, $\overline{c}_{-g}$, $\overline{c}$, and $\overline{x}_g$ are the sample averages of c_{gjt}, c_{-gjt}, c_{jt}, and x_{gjt}. Using this condition in the labor-supply condition (9) gives

$$\begin{aligned}\log n_{gjt} \approx &- \eta_g \cdot \log \alpha_{gj} + \eta_g \cdot \log w_{gjt} - \eta_g \cdot \sigma_g \cdot \frac{\overline{c}}{\overline{c}_g + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{c}_{-g}} \cdot \log c_{jt} \\ &- \eta_g \cdot \frac{\sigma_g}{\sigma_{-g}} \cdot \frac{\overline{c}_{-g}}{\overline{c}_g + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{c}_{-g}} \cdot \log x_{gjt}.\end{aligned} \tag{14}$$

Due to the omission of $\log x_{gjt}$ in (11) the estimated coefficient on the log wage rate is

$$\mathrm{E}\,\widehat{\kappa}_{g,1} \approx \eta_g - \eta_g \cdot \frac{\sigma_g}{\sigma_{-g}} \cdot \frac{\overline{c}_{-g}}{\overline{c}_g + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{c}_{-g}} \cdot \frac{\operatorname{cov}(\log w_{gj}, \log x_{gj})}{\operatorname{var}(\log w_{gj})}. \tag{15}$$

Hence, a labor-supply regression which controls for household consumption would still yield an unbiased estimate of the Frisch elasticity if the omitted relative Pareto weight were a constant, i.e., if there were full commitment. Then, the final term in (14) would simply become part of the individual fixed effect, $\kappa_{0,gj}$. However, under limited commitment, the final term in (14), which is omitted from (11), varies over time. The deviation from its individual-specific mean is then part of the combined residual $\tilde{\varepsilon}_{gjt}$. This is problematic because the relative Pareto weight and hence the combined residual is correlated with the wage rate. Wage raises improve outside options and hence bargaining positions and can therefore lead to an increase in the individual's Pareto weight. Thus, $\log w_{gjt}$ and $\tilde{\varepsilon}_{gjt}$ (which includes a term decreasing in x_{gjt}) are negatively correlated and, as a consequence, an omitted-variable bias occurs and the estimate of the Frisch elasticity, $\hat{\kappa}_{1,g}$, is downward biased,

$$\mathrm{E}\,\widehat{\kappa}_{g,1} < \eta_g.$$

2.3 Solving the estimation problem

While individual consumption is usually not reported in real-world household panel data, household expenditures on different types of goods and services often are. This information can be used for an unbiased estimation of the Frisch elasticity.

Using the first-order conditions for the composition of the two individual consumption

[7] Details can be found in Appendix A.3.

bundles c_{mj} and c_{fj}, (4) and (5), household expenditures for the two goods are given by

$$c_{jt,A} = \gamma_m c_{mjt} + \gamma_f c_{fjt}, \text{ and} \tag{16}$$

$$c_{jt,B} = (1-\gamma_m)\, c_{mjt} + (1-\gamma_f)\, c_{fjt}. \tag{17}$$

Hence, the necessary but unobservable *individual*-specific consumption variables c_{mjt} and c_{fjt} can be expressed as linear functions of the observable *good*-specific consumption variables $c_{jt,A}$ and $c_{jt,B}$. Specifically, we have

$$c_{mjt} = \frac{1-\gamma_f}{\gamma_m - \gamma_f} c_{jt,A} - \frac{\gamma_f}{\gamma_m - \gamma_f} c_{jt,B}, \text{ and}$$

$$c_{fjt} = \frac{1-\gamma_m}{\gamma_f - \gamma_m} c_{jt,A} - \frac{\gamma_m}{\gamma_f - \gamma_m} c_{jt,B}.$$

To use this information in the log-linear labor-supply condition (10), we subtract (17) from (16), divide by c_{jt}, and rearrange terms to obtain

$$\frac{c_{jt,A}}{c_{jt}} = \gamma_{-g} + (\gamma_g - \gamma_{-g}) \cdot \frac{c_{gjt}}{c_{jt}}. \tag{18}$$

Applying a first-order Taylor approximation to this equation and rearranging terms gives[8]

$$\log c_{gjt} \approx \log c_{jt} + \frac{1}{\gamma_g - \gamma_{-g}} \cdot \frac{\bar{c}_A}{\bar{c}_g} \cdot \log(c_{jt,A}/c_{jt}). \tag{19}$$

We can use (19) in (10) to obtain

$$\log n_{gjt} \approx \eta_g \cdot \underbrace{\log w_{gjt}}_{\text{observable}} - \eta_g \sigma_g \cdot \underbrace{\log c_{jt}}_{\text{observable}} + \frac{\eta_g \sigma_g}{\gamma_{-g} - \gamma_g} \cdot \frac{\bar{c}_A}{\bar{c}_g} \cdot \underbrace{\log(c_{jt,A}/c_{jt})}_{\text{observable}} - \underbrace{\eta_g \log \alpha_{gj}}_{\text{fixed effect}}. \tag{20}$$

Note that the variables $\log w_{gjt}$, $\log c_{jt}$, and $\log(c_{jt,A}/c_{jt})$ are all observable in real-world panel data such as the PSID. The final term can be captured by individual fixed effects.

Hence, in a regression where one accounts for the composition of household consumption, measured by the (log) share spent on good A, $\log(c_{gjt,A}/c_{gjt})$, next to the level of household consumption,

$$\log n_{gjt} = \zeta_{0,gj} + \zeta_{1,g} \cdot \log w_{gjt} + \zeta_{2,g} \cdot \log c_{jt} + \zeta_{3,g} \cdot \log(c_{jt,A}/c_{jt}) + \varepsilon_{gjt}, \tag{21}$$

[8] Details can be found in Appendix A.3.

the estimated coefficient on the log wage rate is an unbiased estimate of the Frisch elasticity,

$$\mathrm{E}\,\widehat{\zeta}_{1,g} \approx \eta_g.$$

Note that our estimation approach does not require an a-priori assumption about which spouse prefers which type of consumption goods. It is only necessary that spouses' preferences over the two types of goods *differ*. In our empirical application in Section 4, we will consider alternative aggregations of consumption items to consumption bundles and we will show that different consumption aggregations lead to similar estimates for the Frisch elasticity.

3 Monte-Carlo study

In this section, we calibrate our model and solve it numerically. We then simulate synthetic household panel data and use them to estimate labor-supply regressions. This procedure has two purposes. First, we want to quantify the estimation bias due to limited commitment. Second, we want to evaluate the performance of our improved estimation approach, i.e., determine how strong an approximation error is induced by the log-linear approximation of the relation between the composition of consumption and its distribution, (18).

To evaluate our model quantitatively, we introduce specific wage processes and estimate their parameters from PSID data (see Section 4 for details on our PSID sample). Our calibration strategy focuses on targets in the labor market, the wealth distribution, and the extent of commitment issues since these aspects are the most relevant for our analysis. We also perform sensitivity analyses in terms of key model parameters.

3.1 Calibration

The parametrization is a combination of setting some parameters to values taken from the literature, estimating some parameters, and calibrating others. Table 1 summarizes the baseline parameter values.

Wages. We assume that individual wage rates consist of two components, a fixed component ψ_{gj} and a time-varying component z_{gjt},

$$\log w_{gjt} = \psi_{gj} + z_{gjt}. \tag{22}$$

The time-varying component follows a gender-specific first-order autoregressive process

$$z_{gjt} = \rho_g \cdot z_{gjt-1} + \epsilon_{gjt},$$

with autocorrelation ρ_g and innovations $\epsilon \sim N(0, s^2_{\epsilon_g})$. We allow for a non-zero covariance between the wage-rate shocks of husband and wife, with the covariance being $\omega \cdot s_{\epsilon_m} \cdot s_{\epsilon_f}$, where κ denotes the correlation. We estimate the parameters of the couple wage process using the Generalized Method of Moments (GMM) and the PSID sample which we use for the empirical analysis. Details on the moment conditions can be found in Appendix A.4.

The fixed wage components ψ induce heterogeneity in mean wages and allow us to account for the gender wage gap and inequality within gender. We interpret the fixed wage components as reflecting returns to education and distinguish between two education levels per gender (low and high).[9] We take the relative frequencies of the four combinations of education levels in a household, $\tau_{high,high}$, $\tau_{high,low}$, $\tau_{low,high}$, and $\tau_{low,low}$, directly from our PSID sample and choose the fixed wage components, $\psi_{m,high}$, $\psi_{m,low}$, $\psi_{f,high}$, and $\psi_{f,low}$, to match mean wage rates in each education-gender cell. We normalize the average wage rate in the economy to one.

Preference parameters. We set the discount factor to a standard value for an annual model frequency, $\beta = 0.95$. For the true Frisch labor-supply elasticities, we assume $\eta_m = 0.7$ for men and $\eta_f = 1.05$ for women, in line with the results in Bredemeier et al. (2019). We set the average relative risk aversion in the economy to $(\sigma_m + \sigma_f)/2 = 1.5$, which is a standard value in the literature. Croson and Gneezy (2009) survey the literature on gender differences in risk aversion and conclude that women are on average more risk averse than men. However, little is known about the size of these gender differences. We therefore assume moderate differences in our baseline calibration and later perform sensitivity analyses where we vary the extent of gender differences. Our baseline calibration uses $\sigma_m = 1.3$ and $\sigma_f = 1.7$.

Our estimation approach for the Frisch elasticity relies on the assumption that preference weights for consumption goods, γ_m and γ_f, differ. For the calibration of these parameters, we use our PSID sample and construct two consumption bundles, where one bundle contains goods and services more strongly preferred by men and the other bundle consists of goods and services more strongly preferred by women. To assign the individual consumption items covered in the PSID to either bundle, we use a sample of bachelor households, whose

[9] Low education is defined as 12 grades and below, i.e., high school with and without nonacademic training as well as no schooling. High education is defined as college dropout, college degree, or college and advanced/professional degree.

Table 1: Parameter values, baseline model

Description	Parameter	Value	Comment
Wages			
Autocorrelation, men	ρ_m	0.86	estimated
Autocorrelation, women	ρ_f	0.88	
Shock variance, men	$s^2_{\varepsilon,m}$	0.06	
Shock variance, women	$s^2_{\varepsilon,f}$	0.04	
Correlation wage shocks	ω	0.20	
Mean log wage, high education, men	$\psi_{m,high}$	0.29	observed
Mean log wage, low education, men	$\psi_{m,low}$	-0.15	
Mean log wage, high education, women	$\psi_{f,high}$	-0.09	
Mean log wage, low education, women	$\psi_{f,low}$	-0.46	
Share of hh's with high/high education	$\tau_{high,high}$	0.58	
Share of hh's with high/low education	$\tau_{high,low}$	0.11	
Share of hh's with low/high education	$\tau_{low,high}$	0.15	
Share of hh's with low/low education	$\tau_{low,low}$	0.16	
Preferences			
Discount factor	β	0.95	standard
Risk aversion men	σ_m	1.3	$\overline{\sigma} = 1.5$
Risk aversion women	σ_f	1.7	$\sigma_f = 1.3 \cdot \sigma_m$
Preference for good A, men	γ_m	0.56	$\overline{c}_B/\overline{c} = 0.42$
Preference for good A, women	γ_f	0.61	
Frisch elasticity, men	η_m	0.7	Bredemeier et al. (2019)
Frisch elasticity, women	η_f	1.05	
Disutility weight labor ...			
... highly educated men	$\alpha_{m,\text{high}}$	15.39	$\overline{h}_m^{high} = 0.3413$
... less educated men	$\alpha_{m,\text{low}}$	17.23	$\overline{h}_m^{low} = 0.3411$
... highly educated women	$\alpha_{f,\text{high}}$	16.56	$\overline{h}_f^{high} = 0.3234$
... less educated women	$\alpha_{f,\text{low}}$	20.46	$\overline{h}_f^{low} = 0.3248$
High direct utility gain from marriage	Ψ_{high}	5	$\text{var}(\log x) = 0$ in this group
Low direct utility gain from marriage	Ψ_{low}	0.05	$\partial \log x / \partial (z_m - z_f) = 0.338$
Share households with Ψ_{high}	ν_{high}	0.07	Del Boca and Flinn (2012)
Share households with Ψ_{low}	ν_{low}	0.93	
Prices			
Interest rate	r	1.76%	wealth share(bottom 40%) = 0.6%
Price of good A	P_A	1	normalization
Price of good B	P_B	1	

consumption decisions are not affected by intra-household bargaining. We identify a consumption item as more strongly preferred by men, if, conditional on age and education, the item makes up a larger share of consumption expenditures for male singles than it does for female singles. Comparing expenditures on the two consumption bundles, male singles spend 4.62 percentage points more on the male bundle than female singles spend on this bundle. In our model, this number identifies the difference between the two preference parameters, $\gamma_m - \gamma_f$. To determine their level, we target the average consumption composition of couple households who in our sample spent 42 percent on the male consumption bundle. Using this procedure and understanding good A as the bundle of items more strongly preferred by women, we obtain $\gamma_m = 0.56$ and $\gamma_f = 0.61$. In robustness checks, we considered alternative values for γ_m and γ_f and found that even smaller differences in preference parameters than those used in our baseline calibration deliver almost unbiased estimates. Hence, for our approach, it is not necessary to classify goods a priori. Any categorization of goods and services into bundles over which men and women have different preferences suffices, even if those preference differences are small.

We set group-specific disutility weights on labor to match average hours worked by gender and education. We normalize these targets such that their average is 0.33.

Commitment issues. We calibrate the model to be in line with an empirically reasonable strength of commitment issues. This is of particular importance since we want to assess the quantitative importance of the estimation bias that originates from limited commitment and to evaluate our improved estimation approach quantitatively.

We use the direct utility gain from marriage, Ψ_j, to target the strength of commitment problems and distinguish between two values for Ψ_j.[10] Del Boca and Flinn (2012) estimate a household decision model with participation constraints using PSID data and their results indicate that for 7% of households participation constraints are never binding. Accordingly, for $\nu_{high} = 7\%$ of households in our model, we set the direct utility gain from marriage to a value Ψ_{high} that implies that Pareto weights never need to be adjusted. For the remaining $\nu_{low} = 93\%$ of households, we calibrate direct utility gains from marriage to match empirical variation in bargaining positions within households. Within-household variation in bargaining positions can be identified from variation in the consumption shares of different individuals within households, which is however not available in the PSID. Using data for Japan, Lise and Yamada (2019) estimate a household decision model with time-varying Pareto weights

[10] Together with the four household types in the wage fixed-effect dimension, there are eight household types in total. We assume that utility gains Ψ_j are independent of the fixed wage component ψ_{gj}.

and find an elasticity of relative Pareto weights to relative wages in the household of 0.338. We calibrate Ψ^{low} to match this target and consider alternative targets in sensitivity analyses.

Wealth distribution. The savings behavior of couples has an impact on the extent of commitment issues, see Ábrahám and Laczó (2018). Their analysis implies that households at the bottom of the wealth distribution tend to have stronger commitment issues. In our calibration, we set the interest rate r to match the wealth share of the bottom 40% of the wealth distribution which is 0.6% in the U.S. economy (see Kuhn and Rios-Rull, 2016).

3.2 Simulation results

For the Monte-Carlo experiments, we simulate a large household panel and discard the first 500 periods to avoid dependence on initial conditions. We then repeatedly draw samples of 4,000 households and 5 years (roughly the size of our PSID sample), taking into account the relative frequencies of the different household types, $\tau_j \cdot \nu_j$. With each sample, we run different labor-supply regresions for men.[11] We report average point estimates and average standard errors across 10,000 samples.

For the sake of illustration, we start with a purposely misspecified regression where we regress log hours worked on log wage rates, taking into account individual fixed effects. As no variable is included that attempts to control for the marginal utility of wealth and hence income effects, the estimate for the Frisch elasticity is strongly biased downward, see column (1) of Table 2.

Next, we estimate condition (11), i.e., a regression of log hours worked on an individual fixed effect, the log wage rate, and log household consumption. If there were full commitment, controlling for household consumption (instead of individual consumption) would be sufficient to obtain an unbiased estimate of the Frisch elasticity, as, with constant Pareto weights, individual consumption can be expressed as a function of household consumption alone and the former can hence be proxied perfectly by the latter, see (13). However, with limited commitment between spouses, controlling for household consumption instead of individual consumption leads to a bias in the estimated Frisch elasticity of about 20%, see column (2) of Table 2.

We now turn to our improved estimation approach (21). Column (3) of Table 2 shows that a labor-supply regression where one accounts for the composition of household consumption,

[11] We focus on men because the model abstracts from a number of aspects which are arguably relevant for female labor supply, such as fertility decisions and the availability of child care. Results for women are similar to those for men and are available on request.

Table 2: Labor-supply regressions for men, from synthetic household panel data.

	(1)	(2)	(3)
log wage rate, $\log w_{mjt}$	0.182	0.572	0.699
	(0.002)	(0.001)	(0.000)
log household	–	-0.855	-0.905
consumption, $\log c_{jt}$	–	(0.002)	(0.000)
log share of consumption goods	–	–	20.218
preferred by women, $\log c_{Ajt}/c_{jt}$	–	–	(0.020)
log individual	–	–	–
consumption, $\log c_{mjt}$	–	–	–
relative bias, $(\widehat{\eta}_m - \eta_m)/\eta_m$	-0.740	-0.183	-0.002
fixed effects	yes	yes	yes
number of observations	20,000	20,000	20,000

Notes: Dependent variable is log hours worked $\log n_{mjt}$. The true Frisch elasticity in the model is $\eta_m = 0.7$. The table shows average estimates from 10,000 Monte-Carlo draws, with average standard errors in parentheses.

in this case measured by the (log) share spent on good A, next to the level of household consumption, yields an all but unbiased estimate of the Frisch elasticity. For our baseline parameter values of γ_m and γ_f, the estimation bias decreases to less than one percent.

Sensitivity. We have performed a number of sensitivity checks to corroborate our results. These checks are described in detail in Appendix A.6. Figure 1 and Table 8 in the Appendix show the results, which we summarize here. As discussed before, our improved estimation approach exploits differences between spouses in preferences over the different consumption items. We considered many combinations of γ_m and γ_f and found that even very small differences in preference weights are sufficient for virtually unbiased estimates. The bias in the conventional approach does not depend on these preference weights.

Equation (15) shows that the relative estimation bias $(\widehat{\eta} - \eta)/\eta$ of the conventional approach depends on spouses' risk aversion σ_g. In sensitivity checks, we find that the bias of the conventional approach does not fall below 14% for reasonable values of risk aversion. Our improved estimation approach is successful in effectively removing the bias also in these alternative parameterizations.

Equation (15) also shows that the relative estimation bias of the conventional approach depends on the covariance term between bargaining weights and relative wages. In our baseline calibration, we target the relation between spouses' relative wages and the bargaining weight taken from Lise and Yamada (2019). To explore the sensitivity of our results to this statistic, we run specifications where we vary this target and thus the direct utility gain from marriage. We find that the bias in the conventional approach remains strong even for substantial variations of the calibration target.

Finally, we vary the correlation between the stochastic wage components of husband and wife. This correlation can be important for the estimation bias because we target the correlation between spouses' relative wages and bargaining weights but the estimation results depend on the relation between the husband's absolute wage rate and the bargaining weight, see equation (15). We find that varying the correlation affects the bias in the conventional estimation only little and our improved estimation approach continues to deliver a virtually unbiased estimate.

For completeness, we have also considered a full-commitment version of our model where Pareto weights are constant. Using simulated data from this model version, we confirm that our improved estimation approach yields an unbiased estimate of the Frisch elasticity also under full commitment. Thus, applying our approach, one does not have to make a-priori assumptions about the degree of commitment between spouses.

4 Empirical analysis

4.1 Sample selection and variable definitions

Data and sample selection. We use waves 1999-2017 of the Panel Study of Income Dynamics (PSID). Since 1999, the PSID data has been collected biennially. A key advantage of recent waves of the PSID is the newly included detailed information on subitems of household consumption. The new design of the PSID covers over 70% of all consumption items covered in the Consumer Expenditure Survey (CEX) and 23 subitems are covered in a consistent way over time.[12] Our sample selection closely follows Blundell et al. (2016) who use PSID waves 1999–2009 to investigate family labor supply as an insurance mechanism.

We consider married households where both spouses are 25-60 years old and we drop the Survey of Economic Opportunity sample and the immigrant sample. In our baseline analysis, we drop couples in the period where they dissolve but include these household heads when

[12] See Andreski et al. (2014) and Blundell et al. (2016) for a comparison of consumption data from the PSID with the CEX and NIPA.

they marry again. We drop observations with wages below half the hourly minimum wage, observations where couples report very high asset values ($20 million and more), couples who receive transfers higher than twice total household earnings and we do not use data displaying extreme jumps from one PSID wave to the next.[13] As in the Monte-Carlo analysis, our regressions consider the husbands' hours worked in this sample. Throughout, we use PSID sampling weights.

Hours and wages. The hours variable is annual hours worked, calculated as weeks worked times usual weekly hours plus overtime hours. As is standard in the labor-supply literature, the hourly wage rate is constructed by dividing annual earnings by annual hours worked. Our regressions will correct for the division bias that would otherwise result from the constructed wage rate variable. Earnings are measured in 2000 dollars and consist of labor earnings, the labor part of business income, and the labor part of farm income. Blomquist (1985, (1988)) has emphasized the importance of using net rather than gross wage rates in labor-supply regressions. To transform gross wages into net wages, we calculate taxes and eligible amounts of EITC and food stamps benefits using program information for the various years. We take into account that benefits vary by demographic characteristics, e.g., the number and age of children, etc. We determine marginal tax rates by considering the change in income after taxes and transfers induced by a $500 increase in gross earnings. The net wage rate is then given by the gross wage rate times one minus the marginal tax rate. In the following, we denote the net wage rate simply as wage rate in line with our model where we only consider net wage rates. We also consider a robustness check using gross wage rates.

Consumption. Table 7 in Appendix A.5 shows the various consumption items that are covered in the PSID in a consistent way over time and it also shows different aggregations of consumption items to consumption bundles that we use in our analysis. For the calibration of our model, we have already introduced one possible dichotomous classification of consumption items, in terms of goods and services more strongly preferred by men and women, respectively. Yet, an advantage of our approach is that is does not require a-priori information about which spouse prefers which type of consumption goods. We therefore additionally consider a more agnostic dichotomous classification where we follow Blundell et al. (2016) and distinguish between nondurable goods on the one hand and services on the other hand.[14]

[13] Such a jump is defined as an extremely positive (negative) wage change from $t-2$ to t, followed by an extremely negative (positive) change from t to $t+2$, see Blundell et al. (2016) for details.

[14] Nondurable goods consist of food at home, food stamps, and gasoline consumption and services consist of food out of home, health insurance, health services, utilities, transportation, education, child care, home insurance, and rent (or rent equivalent). For homeowners 6% of the reported value of the house is used as a rent equivalent, following Poterba and Sinai (2008).

In our quantitative model, small differences in preferences of household members over two types of goods are sufficient for consistent estimation of the Frisch elasticity, see Section 3. In an empirical application, estimation can be complicated, e.g., by measurement error, such that using more information can be helpful. We therefore consider also finer breakdowns of household consumption. Specifically, we assign the PSID consumption items into the personal consumption expenditure (PCE) major product types used in the national product and income accounts (NIPA), which yields eight non-empty product types.[15] As another alternative, we use the expenditure shares on all 23 consumption items as individual control variables in our regressions. In all specifications, we omit one consumption share as a base category.[16]

Additional variables. Our labor-supply regressions include individual and time fixed effects. Individual fixed effects account for heterogeneity in the taste for work, as indicated by the model, see (9) together with (10) and (11) as well as (20) together with (21). By including time effects in the labor-supply regression we filter out effects related to the economy-wide business cycle, i.e., demand-determined fluctuations in hours worked from which we abstract in the model. While an individual's preference for non-working time is constant in the model, it may vary empirically. To account for taste shifters that display time variation, we include a third-order polynomial in age and the number of young and old children in the household.[17] In robustness checks, we consider additional potential taste shifters. Variables affecting the taste for work which are mostly constant over time, such as education, are accounted for by the individual fixed effect.

Wage regression. As is well known in the labor-supply literature (see, e.g., Altonji 1986; Borjas 1980; Pencavel 1986 and Keane 2011), labor-supply regressions are subject to a division bias when wage rates have to be computed as earnings divided by hours worked. This generates a spurious negative correlation of the calculated wage rate with hours worked

[15] The categories are food and beverages purchased for off-premises consumption, gasoline and other energy goods, housing and utilities, health care, transportation services, food services and accommodations, financial services and insurance, and other services. The PSID does not contain consistent information on the other PCE major product types: motor vehicle parts, furnishings and durable household equipment, recreational goods and vehicles, other durable goods (all belonging to the durable goods category), clothing and footwear, other nondurable goods (nondurable), and recreation services (services).

[16] When aggregating the various consumption items to consumption bundles, missing values in the items are treated as zeros. The more detailed the consumption bundles are defined (e.g., using eight instead of two bundles), the more often a household may not consume the bundle at all. To account for the resulting zeros, we add a one to the different consumption quantities before including them as log shares in the labor-supply regression. We ensure that we leave the sample unchanged in the various specifications we consider. We achieve this by first checking the availability of all potentially required variables for each household and only using those households in the analysis for which all variables are available.

[17] Young children are below age 7. Old children are age 7-17.

because measurement error in hours worked occurs on both sides of the regression equation. As first suggested by Borjas (1980), we therefore run an initial wage regression and use it to determine predicted wage rates which are uncorrelated with the measurement error in hours worked. We then use predicted log (net) wage rates $\log \widetilde{w}_{gjt}$ in the labor-supply regression.

In the wage regression, we include a third-order polynomial in age and interactions of these terms with education, race (white, African American, other), firm tenure, firm tenure squared, state dummies, year dummies, and following Altonji (1986), the other variables from the labor-supply regression.[18] By taking into account interaction terms in age and education, predicted wages are identified through education-specific life-cycle wage profiles, taking into account a set of control variables as well as time effects that exploit trend and cyclicality of wages for prediction. This procedure is standard in the labor-supply literature, where the Frisch elasticity is inferred from the reaction of hours to wage changes over the life cycle, see, e.g., Rupert et al. (2000). In our sample, the wage regression has an R^2 of about 35%.

4.2 Empirical results

Baseline labor-supply regressions. Column 1 of Table 3 refers to a simple regression where we regress (log) hours worked on (log) predicted wage rates, individual and time fixed effects, as well as taste shifters. This specification is included for illustration only. As discussed before, such a regression is unable to identify the Frisch elasticity as no variable is included that attempts to control for the marginal utility of wealth and hence income effects. As expected, we estimate the wage sensitivity of hours worked to be very small in this specification. In this (purposely misspecified) regression, we estimate an elasticity of only 0.18. The second column shows that the coefficient on the wage rate increases substantially when total household consumption is included and hence the income effect is accounted for. Such a regression is in the spirit of Altonji (1986) and yields an estimated elasticity of 0.59. In line with the presence of wealth effects, we estimate a statistically significant negative coefficient on household consumption. These results confirm that Altonji's (1986) fundamental point of taking into account consumption information at all is of great importance when estimating labor-supply elasticities.

[18] Education is measured in three categories: Category 1 is 0-11 grades (includes those with no schooling); category 2 is 12 grades, i.e., high school with and without nonacademic training; category 3 is college dropout, college degree, or college and advanced/professional degree.

Table 3: Labor-supply regressions for men, from PSID data.

	(1)	(2)	(3)
log wage rate, $\log \tilde{w}_{mjt}$	0.181	0.593	0.677
	(0.015)	(0.030)	(0.032)
log household	–	-0.297	-0.351
consumption $\log c_{jt}$	–	(0.019)	(0.020)
log share of consumption items	–	–	0.021
preferred by women, $\log(c_{jt,A}/c_{jt})$	–	–	(0.003)
fixed effects	yes	yes	yes
time effects	yes	yes	yes
cubic in age	yes	yes	yes
# young children	yes	yes	yes
# old children	yes	yes	yes
observations	19,797	19,797	19,797

Notes: Dependent variable is log hours worked $\log n_{mjt}$. Standard errors in parentheses. See Table 7 in Appendix A.5 for details on the aggregations of consumption items. Young children are below age 7. Old children are age 7-17.

Yet, our analysis has shown that total household consumption is an imperfect proxy for individual consumption since commitment between spouses is limited. As discussed before, a regression using total household consumption leads to an underestimation of the Frisch elasticity. Our Monte-Carlo estimations have shown that time variation in bargaining weights can be captured in a regression where one accounts for the composition of household consumption, measured by the (log) shares spent on different consumption bundles, next to the level of household consumption. Our empirical results for the PSID are in line with these theoretical predictions. When using the classification into consumption items more strongly preferred by men or women, respectively, the estimate for the Frisch elasticity increases to a value of 0.68, see column (3) of Table 3. This indicates that labor-supply regressions which do not control for the composition of household consumption suffer from a non-negligible underestimation of the Frisch elasticity.

The coefficient on the consumption bundle more strongly preferred by women is positive. This is in line with the theory of limited commitment, as a high consumption share of the wife is indicative for the wife having strong bargaining power in the household which translates into more leisure for her and less for her husband, who accordingly has to work more.

Table 4: Estimation results for alternative disaggregations of consumption, from PSID data.

	(1)	(2)	(3)
log wage rate, log $\tilde{w}_{mjt}$	0.670 (0.032)	0.671 (0.032)	0.711 (0.034)
consumption categorization	Nondur./Serv.	NIPA	all items
observations	19,797	19,797	19,797

Notes: Dependent variable is log hours worked $\log n_{mjt}$. Standard errors in parentheses. All regressions control for individual and time fixed effects, total household consumption, the log shares of consumption categories, a cubic in age, as well as the number of young and old children, respectively. See Table 7 in Appendix A.5 for details on the aggregations of consumption items. In all regressions, one consumption share is omitted as a base category. Young children are below age 7. Old children are age 7-17.

Table 4 reports results for specifications where we use alternative disaggregations of consumption. Column (1) shows that the agnostic classification of consumption items into mondurables and services yields a very similar estimate for the Frisch elasticity. This also holds when we use the finer NIPA categorization in terms of eight consumption bundles, see column (2). When we account for all consumption items covered in the PSID separately, the estimated Frisch elasticity increases to 0.71, see column (3). Our finding that the estimates for the Frisch elasticity do not change much relative to our baseline specification reported in Table 3 corroborates that it is sufficient to account for the composition of household consumption in an agnostic way. In particular, it is not necessary to assign consumption items to either husband or wife a priori.

Robustness checks. We have corroborated our estimates for the Frisch elasticity in several robustness checks. The results of the most important checks are summarized in Table 5. For these regressions, we use the specification with all consumption items covered in the PSID.[19]

Rupert et al. (2000) have shown that ignoring changes in the productivity in non-market activities can bias estimates for the Frisch elasticity downward. The reason is that peak earnings years coincide with the period in which individuals have the greatest productivity, or the greatest demands on their time, in home production. Rupert et al. (2000) et al. mention having children and buying a house as events which call for more time spent in household activities and mostly occur in the phase of the life cycle in which wages are high. Our baseline estimations already control for age and the presence of children. In column (1)

[19] We obtain similar results when using the alternative breakdowns of consumption.

Table 5: Additional estimation results, from PSID data.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log wage rate, $\log \tilde{w}_{mjt}$	0.709 (0.034)	0.732 (0.041)	0.676 (0.039)	0.715 (0.034)	0.717 (0.034)	0.722 (0.036)	0.729 (0.038)
homeownership	yes	no	no	no	no	no	no
wage rate	net	net	net	gross	net	net	net
sample	baseline	w/o prof. occ.	w/o similar earnings couples	baseline	cohabiting couples included	only 1st marriages	w/o eventual divorcees
observations	19,797	15,376	14,975	19,797	21,603	17,395	14,914

Notes: Dependent variable is log hours worked $\log n_{mjt}$. Standard errors in parentheses. All regressions control for individual and time fixed effects, total household consumption, 22 log shares of consumption items covered in the PSID (see Appendix A.5), a cubic in age, as well as the number of young and old children, respectively. In all regressions, one consumption share is omitted as a base category. Young children are below age 7. Old children are age 7-17. Professional occupations: managers, business specialists, financial specialists, lawyers, judges, and physicians. Similar earnings = wife's share in household earnings between 45% and 55%.

of Table 5, we show that additionally controlling for homeownership affects the estimated Frisch elasticity only very little in our application.

Imai and Keane (2004) have shown that estimates for the Frisch elasticity can be biased downward because young workers with low wages invest in their human capital by working much which renders the life-cycle profile of hours worked flat. To address this point, we run a regression where we omit individuals working in occupations with the highest returns to working long hours. Cortés and Tessada (2011) emphasize that professional occupations such as physicians and lawyers are characterized by people having to work long hours to have successful careers. We therefore choose managers, business specialists, financial specialists, lawyers, judges, and physicians as the occupation groups to be excluded from the sample. Omitting individuals in professional occupations slightly increases the estimate of the Frisch elasticity, in line with the Imai and Keane (2004) argument, see column (2) of Table 5.

Evidence documented by Bertrand et al. (2015) suggests an aversion to a situation where the wife earns more than her husband and that households make costly choices to avoid this situation. This aversion is absent from our model and we check the importance of this assumption by omitting households where the average share of the wife in household earnings lies between 45% and 55%. Without these households, our results remain similar and the estimate for men's Frisch elasticity is 0.68, see column (3) of Table 5.

As discussed before, Blomquist (1985, 1988) has stressed the importance of using net rather than gross wages in labor-supply regressions because changes in gross wages overstate changes in decision-relevant net wages when the tax system is progressive. In our baseline regressions, we therefore used net wage rates as regressors. In order to assess how important the construction of net wage rates is for our results, we also run a specification where we use the gross wage rate instead. The results, shown in column (4) of Table 5, are only little affected by this change. This is consistent with the findings of Guner et al. (2014) who show that actual tax liabilities of U.S. households are substantially less progressive than a simple look on tax rates by bracket suggests.

Next, we include cohabiting (unmarried) couples in the analysis.[20] Column (5) shows that the estimated Frisch elasticity is similar in this enlarged sample.

Finally, we address the point that our model has no divorces in equilibrium and all limited-commitment effects are driven by the *threat* of divorce while our empirical sample contains households that do divorce later. We consider two additional sample restrictions. First, different from our baseline sample, we not only drop couples in the period where they dissolve but also do not include these household heads when they marry again. Second, we omit all households that eventually separate while included in the PSID. We find that this has little effect on the estimated Frisch elasticity, see columns (6) and (7) of Table 5.

Across specifications, our improved estimation approach delivers estimates for the Frisch elasticity ranging from 0.67 to 0.73. These numbers suggest that the simple labor-supply regression that does not control for the composition of household consumption (column (2) of Table 3) underestimates the true value of the Frisch elasticity by up to 20%, in line with the results of our Monte-Carlo estimations using the model-generated data.

5 Conclusion

We have analyzed the consequence of limited commitment in the family for the estimation of labor-supply elasticities. In principle, the Frisch elasticity can be estimated in a regression of hours worked on the hourly wage rate where one controls for consumption as a measure of the marginal utility of wealth. We have shown that such a regression would yield an unbiased estimate only if information on individual consumption were available or if there were full commitment in the household. We have developed and calibrated a dual-earner household model to analyze the quantitative importance of the bias that results when commitment between household members is limited and household consumption is used as a proxy for individual consumption. Our results show that standard labor-supply regressions yield estimated Frisch elasticities that are about 20% below their true values.

[20] In this specification, we include a marital status indicator in the wage regression.

We have developed an improved estimation approach that all but eliminates this bias. Our approach uses data on the composition of household consumption which is available in many household panel surveys and is informative about how consumption of a household is distributed among its members. In Monte-Carlo experiments, we have found that our estimation approach reduces the bias in the estimated Frisch elasticity to below 1%. We have applied this approach to U.S. panel data from the PSID and found Frisch elasticities for men of about 0.7. In line with the predictions of our model, these estimates are substantially larger than those obtained from labor-supply regressions that fail to account for the effects of limited commitment. Hence, individuals seem more willing to substitute hours worked intertemporally than conventional estimation approaches suggest.

References

Ábrahám, Á. and S. Laczó (2018). Efficient risk sharing with limited commitment and storage. *The Review of Economic Studies 85*(3), 1389–1424.

Alogoskoufis, G. S. (1987). On intertemporal substitution and aggregate labor supply. *Journal of Political Economy 95*(5), 938–960.

Altonji, J. G. (1986). Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy 94*(3, Part 2), S176–S215.

Andreski, P., G. Li, M. Z. Samancioglu, and R. Schoeni (2014). Estimates of annual consumption expenditures and its major components in the PSID in comparison to the CE. *American Economic Review 104*(5), 132–35.

Bertrand, M., E. Kamenica, and J. Pan (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics 130*(2), 571–614.

Blomquist, N. S. (1985). Labour supply in a two-period model: The effect of a nonlinear progressive income tax. *The Review of Economic Studies 52*(3), 515–524.

Blomquist, N. S. (1988). Nonlinear taxes and labor supply. *European Economic Review 32*(6), 1213–1226.

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

Borjas, G. J. (1980). The relationship between wages and weekly hours of work: The role of division bias. *The Journal of Human Resources 15*(3), 409–423.

Bredemeier, C., J. Gravert, and F. Juessen (2019). Estimating labor supply elasticities with joint borrowing constraints of couples. *Journal of Labor Economics 37*(4), 1215–1265.

Cesarini, D., E. Lindqvist, M. J. Notowidigdo, and R. Östling (2017). The effect of wealth on individual and household labor supply: Evidence from Swedish lotteries. *American Economic Review 107*(12), 3917–3946.

Chiappori, P.-A., M. C. Dias, and C. Meghir (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy 126*(S1), S26–S72.

Chiappori, P.-A., J. I. Gimenez-Nadal, J. A. Molina, A. Theloudis, and J. Velilla (2020). Intrahousehold commitment and intertemporal labor supply. IZA Discussion Paper 13545.

Chiappori, P.-A. and M. Mazzocco (2017). Static and intertemporal household decisions. *Journal of Economic Literature 55*(3), 985–1045.

Cortés, P. and J. Tessada (2011). Low-skilled immigration and the labor supply of highly skilled women. *American Economic Journal: Applied Economics 3*(3), 88–123.

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–74.

Del Boca, D. and C. Flinn (2012). Endogenous household interaction. *Journal of Econometrics 166*(1), 49–65.

Dercon, S. and P. Krishnan (2000). In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of Political Economy 108*(4), 688–727.

Domeij, D. and M. Floden (2006). The labor-supply elasticity and borrowing constraints: Why estimates are biased. *Review of Economic Dynamics 9*(2), 242–262.

Duflo, E. and C. Udry (2004). Intrahousehold resource allocation in Cote d'Ivoire: Social norms, separate accounts and consumption choices. NBER Discussion Paper 10498.

Floden, M. and J. Lindé (2001). Idiosyncratic risk in the United States and Sweden: Is there a role for government insurance? *Review of Economic Dynamics 4*(2), 406–437.

Garner, T. I., G. Janini, W. Passero, L. Paszkiewicz, and M. Vendemia (2006). The CE and the PCE: A comparison. *Monthly Labor Review, September 2006*, 20–46.

Groneck, M. and J. Wallenius (2021). It Sucks to Be Single! Marital Status and Redistribution of Social Security. *The Economic Journal 131*(633), 327–371.

Guner, N., R. Kaygusuz, and G. Ventura (2014). Income taxation of U.S. households: Facts and parametric estimates. *Review of Economic Dynamics 17*(4), 559–581.

Heathcote, J., F. Perri, and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics 13*(1), 15–51.

Heckman, J. J. (1993). What has been learned about labor supply in the past twenty years? *American Economic Review 83*(2), 116–121.

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

Imai, S. and M. P. Keane (2004). Intertemporal labor supply and human capital accumulation. *International Economic Review 45*(2), 601–641.

Keane, M. and R. Rogerson (2015). Reconciling micro and macro labor supply elasticities: A structural perspective. *Annual Review of Economics 7*(1), 89–117.

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

Kuhn, M. and J.-V. Ríos-Rull (2016). 2013 Update on the US earnings, income, and wealth distributional facts: A view from macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review 37*(1), 2–73.

Lise, J. and K. Yamada (2019). Household sharing and commitment: Evidence from panel data on individual expenditures and time use. *The Review of Economic Studies 86*(5), 2184–2219.

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.

Marcet, A. and R. Marimon (2019). Recursive contracts. *Econometrica 87*(5), 1589–1631.

Mazzocco, M. (2007). Household intertemporal behaviour: A collective characterization and a test of commitment. *The Review of Economic Studies 74*(3), 857–895.

Mazzocco, M., C. Ruiz, and S. Yamaguchi (2014). Labor supply and household dynamics. *American Economic Review 104*(5), 354–59.

Nash, J. F. (1950). The bargaining problem. *Econometrica 18*(2), 155–162.

Pencavel, J. (1986). Labor supply of men: A survey. In *Handbook of Labor Economics*, Volume 1, pp. 3–102. Elsevier.

Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics 21*(3), 729–754.

Poterba, J. and T. Sinai (2008). Tax expenditures for owner-occupied housing: Deductions for property taxes and mortgage interest and the exclusion of imputed rental income. *American Economic Review 98*(2), 84–89.

Robinson, J. (2012). Limited insurance within the household: Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics 4*(4), 140–164.

Rupert, P., R. Rogerson, and R. Wright (2000). Homework in labor economics: Household production and intertemporal substitution. *Journal of Monetary Economics 46*(3), 557–579.

Storesletten, K., C. I. Telmer, and A. Yaron (2004). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy 112*(3), 695–717.

Voena, A. (2015). Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples? *American Economic Review 105*(8), 2295–2332.

A Appendix

A.1 Maximization problem of bachelor households (outside option)

The outside option of agent gj in household j in a particular period is given by the expected lifetime utility of being single. When agent gj lives as a single, the maximization problem is

$$S_{gj}\left(w_{gj}, a_{gj}\right) = \max_{c_{gj}, c_{gj,A}, c_{gj,B}, n_{gj}, a'_{gj}} u_{gj}\left(c_{gj}, n_{gj}\right) + \beta \cdot \mathrm{E}\left[S_{gj}\left(w'_{gj}, a'_{gj}\right)\right],$$

with

$$c_{gj} = \gamma_g^{\gamma_g}\left(1-\gamma_g\right)^{1-\gamma_g} \cdot c_{gj,A}^{\gamma_g} c_{gj,B}^{1-\gamma_g},$$

subject to the period budget constraint

$$c_{gj} + a'_{gj} \leq w_{gj} n_{gj} + (1+r)\, a_{gj}$$

and the borrowing constraint

$$a'_{gj} > 0.$$

A.2 Nash bargaining

Initial Pareto weights are determined using Nash (1950) bargaining before shocks in the first period realize. To highlight this timing, we use the time index 0 in the following. Spouses enter the couple household only when there is a gain from marriage such that their expected lifetime utility of being married is higher than expected lifetime utility as a single,

$$\begin{aligned} V_{mj}\left(\Omega_{j0}\right) &\geq S_{mj}\left(w_{mj0}, a_{mj0}\right), \\ V_{fj}\left(\Omega_{j0}\right) &\geq S_{fj}\left(w_{fj0}, a_{fj0}\right), \end{aligned}$$

where w_{gj0} and a_{gj0} denote wage rates and asset holdings at household formation. Ω_{j0} summarizes the state variables at household formation, $\Omega_{j0} = (w_{mj0}, w_{fj0}, \mu_{mj0}, \mu_{fj0}, a_{mj0} + a_{fj0})$. The resulting Nash program is

$$N = \max_{\mu_{mj0}, \mu_{fj0}} \left(V_{mj}\left(\Omega_{j0}\right) - S_{mj}\left(w_{mj0}, a_{mj0}\right)\right)^{1/2} \cdot \left(V_{fj}\left(\Omega_{j0}\right) - S_{fj}\left(w_{fj0}, a_{fj0}\right)\right)^{1/2},$$

To simplify notation and without loss of generality, we assume that initial Pareto weights sum up to one such that $\mu_{-gj0} = 1 - \mu_{gj0}$. Then, the first-order condition of the Nash program is

$$\frac{V_{mj}(\Omega_{j0}) - S_{mj}(w_{mj0}, a_{mj0})}{V_{fj}(\Omega_{j0}) - S_{fj}(w_{fj0}, a_{fj0})} = -\frac{\partial V_{mj}(\Omega_{j0}) / \partial \mu_{gj0}}{\partial V_{mj}(\Omega_{j0}) / \partial \mu_{gj0}}.$$

A.3 Taylor approximations

Obtaining (13) from (12) involves the following steps:

$$\begin{aligned}
(12) \Leftrightarrow c_{jt} - \overline{c}_j &\approx c_{gjt} - \overline{c}_g - \frac{1}{\sigma_{-g}} \overline{x}_g^{-1/\sigma_{-g}-1} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} \cdot (x_{gjt} - \overline{x}_g) \\
&\quad + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}-1} \cdot (c_{gjt} - \overline{c}_g) \\
\Leftrightarrow \overline{c}_j \frac{c_{jt} - \overline{c}_j}{\overline{c}_j} &\approx \overline{c}_g \frac{c_{gjt} - \overline{c}_g}{\overline{c}_g} - \frac{1}{\sigma_{-g}} \overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} \cdot \frac{x_{gjt} - \overline{x}_g}{\overline{x}_g} \\
&\quad + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} \cdot \frac{c_{gjt} - \overline{c}_g}{\overline{c}_g} \\
\Leftrightarrow \overline{c}_j \log c_{jt} &\approx \overline{c}_g \log c_{gjt} - \frac{1}{\sigma_{-g}} \overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} \cdot \log x_{gjt} \\
&\quad + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} \cdot \log c_{gjt} \\
&= \overline{c}_g \log c_{gjt} - \frac{1}{\sigma_{-g}} \overline{c}_{-gj} \cdot \log x_{gjt} + \frac{\sigma_g}{\sigma_{-g}} \cdot \overline{c}_{-g} \cdot \log c_{gjt} \Leftrightarrow (13)
\end{aligned}$$

where the penultimate step uses $\overline{x}_g^{-1/\sigma_{-g}} \cdot \overline{c}_g^{\sigma_g/\sigma_{-g}} = \overline{c}_{-g}$ which follows from (3).

Obtaining (19) from (18) involves the following steps, where $\tilde{c}jt, A = c_{jt,A}/c_{jt}$ and $\tilde{c}gjt = c_{gjt}/c_{jt}$ are defined for notational convenience:

$$\begin{aligned}
(18) \Leftrightarrow \tilde{c}jt, A - \overline{s}_A &= (\gamma_g - \gamma_{-g}) (\tilde{c}gjt - \overline{s}_g) \Leftrightarrow \frac{\tilde{c}jt, A - \overline{s}_A}{\overline{s}_A} \overline{s}_A = (\gamma_g - \gamma_{-g}) \frac{\tilde{c}gjt - \overline{s}_g}{\overline{s}_g} \overline{s}_g \\
\Leftrightarrow \overline{s}_A \log \tilde{c}jt, A &\approx (\gamma_g - \gamma_{-g}) \overline{s}_g \log \tilde{c}gjt \Leftrightarrow \log \tilde{c}gjt \approx \frac{1}{(\gamma_g - \gamma_{-g})} \frac{\overline{s}_A}{\overline{s}_g} \log \tilde{c}jt, A \Leftrightarrow (19).
\end{aligned}$$

A.4 Stochastic wage process

We estimate the parameters of the gender-specific AR(1) processes for both genders simultaneously to allow for arbitrary covariance structures. We use the Generalized Method of Moment (GMM) and our baseline PSID sample. We first estimate a filter regression to eliminate predictable components of observed log wage rates, $\log w_{gjt}$,

$$\log w_{gjt} = f\left(q_{gjt}\right) + \hat{w}_{gjt},$$

where q_{gjt} denotes characteristics of an individual with gender g in household j in period t and $f\left(q_{gjt}\right)$ consists of a cubic polynomial in job market experience, dummies for education, time dummies, as well as interactions of all variables. Following Chiappori et al. (2020), job market experience is constructed as age minus six and the number of years in school. Education is measured in three categories: Category 1 is 0-11 grades (includes those with no schooling); category 2 is 12 grades, i.e. high school with and without nonacademic training; category 3 is college dropout, college degree, or college and advanced/professional degree. The wage process is then estimated for residual log wage rates $\hat{w}_{gjt}$.

We follow the applied literature, see, e.g., Heathcote et al. (2010), and account for individual fixed effects, measurement errors, and time-varying factor loadings. While the variance terms are time invariant in our theoretical model, the applied literature has shown the importance of allowing for such flexibility in the estimated processes to correctly identify persistence and idiosyncratic risk. Residual log wages are decomposed into

$$\hat{w}_{gjt} = \pi_{gt} \cdot \chi_{gj} + \theta_{gt} \cdot z_{gjt} + \varphi_{gt} \cdot u_{gjt},$$

where π_{gt}, θ_{gt}, and φ_{gt} are gender-specific time-varying factor loadings, χ_{gj} is an individual fixed effect (which in our model is, together with $f\left(q_{gj}\right)$, part of the fixed wage component ψ_{gj}), z_{gjt} is the autocorrelated component, and u_{gjt} denotes measurement error. The fixed effect has variance $s^2_{\chi_g}$. The measurement error has variance $s^2_{u_g}$. z_{gjt} is assumed to follow a gender-specific first-order autoregressive process,

$$z_{gjt} = \rho_g \cdot z_{gjt-1} + \epsilon_{gjt}.$$

with innovation variance $s^2_{\epsilon_g}$ and an arbitrary correlation between the innovations of husband and wife, ϵ_{mjt} and ϵ_{fjt}, denoted by $s_{\epsilon_m \epsilon_f}$. We follow the approach by MaCurdy (1982) and treat the variances and the covariance of the auto-regressive components at the start of the sample period, $s^2_{z_g 1}$ $(g = m, f)$ and $s_{z_m z_f 1}$, as additional parameters to be estimated.

In the moment conditions of the GMM estimation, we use the variances and covariances of $\hat{w}_{mjt+s}$ and $\hat{w}_{fjt+s}$. The wage variances are

$$\begin{aligned}
\mathrm{var}(\hat{w}_{gj1}) =& \pi^2_{g1} \cdot s^2_{\chi_g} + \theta^2_{g1} \cdot s^2_{z_g 1} + \varphi^2_{g1} \cdot s^2_{u_g},\ t = 1, \\
\mathrm{var}(\hat{w}_{gjt}) =& \pi^2_{gt} \cdot s^2_{\chi_g} + \theta^2_{gt} \cdot \left(\rho_g^{2t-2} \cdot s^2_{z_g 1} + s^2_{\epsilon_g} \cdot \sum_{s=0}^{t-2} \rho_g^{2s} \right) + \varphi^2_{gt} \cdot s^2_{u_g},\ t > 1.
\end{aligned}$$

The auto-covariances are

$$\begin{aligned}
\mathrm{cov}\left(\hat{w}_{gj1}, \hat{w}_{gjt}\right) =& \pi_{g1} \cdot \pi_{gt} \cdot s^2_{\chi_g} + \theta_{g1} \cdot \theta_{gt} \cdot s^2_{z_g 1},\ t > 1, \\
\mathrm{cov}\left(\hat{w}_{gjt}, \hat{w}_{gjt+s}\right) =& \pi_{gt} \cdot \pi_{gt+s} \cdot s^2_{\chi_g} + \theta_{gt} \cdot \theta_{gt+s} \cdot \left(\rho_g^{2t+s-2} \cdot s^2_{z_g 1} + \rho_g^s \cdot s^2_{\epsilon_g} \cdot \sum_{k=0}^{t-2} \rho_g^{2k} \right),\ t > 1,\ s > 0.
\end{aligned}$$

The covariances between $\hat{w}_{mjt}$ and $\hat{w}_{fjt}$ are

$$\begin{aligned}
\mathrm{cov}\left(\hat{w}_{mj1}, \hat{w}_{fj1}\right) =& \theta_{m1} \cdot \theta_{f1} \cdot s_{z_m z_f 1}, \\
\mathrm{cov}\left(\hat{w}_{mjt}, \hat{w}_{fjt}\right) =& \theta_{mt} \cdot \theta_{ft} \cdot \left(\rho_m^{t-1} \cdot \rho_f^{t-1} \cdot s_{z_m z_f 1} + s_{\epsilon_m \epsilon_f} \cdot \sum_{s=0}^{t-2} \rho_m^s \cdot \rho_f^s \right),\ t > 1.
\end{aligned}$$

The covariances between male and female wages at different points in time, $\mathrm{cov}\left(\hat{w}_{mjt}, \hat{w}_{fjt+s}\right)$, would be redundant as the information is included in $\mathrm{cov}\left(\hat{w}_{mjt}, \hat{w}_{fjt}\right)$ together with $\mathrm{cov}\left(\hat{w}_{gjt}, \hat{w}_{gjt+s}\right)$. GMM estimation is carried out by replacing population moments by their sample analogs. We estimate

$$\Theta_g = \{s^2_{\chi_g}, \rho_g, s^2_{z_g 1}, s^2_{\epsilon_g}, s^2_{u_g}, \pi_{g2}, ..., \pi_{gT}, \theta_{g2}, ..., \theta_{gT}, \varphi_{g2}, ..., \varphi_{gT}\}, \tag{23}$$

for both genders $g = m, f$ as well as the covariances $s_{z_m z_f 1}$ and $s_{\epsilon_m \epsilon_f}$. For identification, the first-period factor loadings π_{g1}, θ_{g1}, and φ_{g1} are normalized to 1.

Column (1) of Table 6 shows the point estimates (point estimates for the time-varying factor loadings are not shown) and column (2) shows standard errors. Since the PSID data are biennial, we annualize the estimates by dividing estimated variances and covariances by two and taking the square root of the estimated autocorrelation, see column (3). The implied annualized autocorrelations, while closer to the lower bound of values typically used in the incomplete markets literature, see, e.g., Floden and Lindé (2001) and Storesletten et al. (2004), are in line with other studies which estimate jointly the wage processes of

Table 6: Estimated process for residual wage rates, men and women

	(1) Point estimate	(2) Standard error	(3) Annualized parameter
Men			
$s^2_{\chi m}$	0.057	0.030	0.029
ρ_m	0.736	0.112	0.858
$s^2_{w_m 1}$	0.185	0.023	0.093
$s^2_{\epsilon_m}$	0.126	0.072	0.063
$s^2_{u_m}$	0.056	0.022	0.028
Women			
$s^2_{\chi f}$	0.001	0.068	0.001
ρ_f	0.779	0.171	0.882
$s^2_{w_f 1}$	0.209	0.058	0.105
$s^2_{\epsilon_f}$	0.084	0.053	0.042
$s^2_{u_f}$	0.069	0.025	0.034
Covariances			
$s_{w_m w_f 1}$	0.045	0.007	0.045
$s_{\epsilon_m \epsilon_f}$	0.020	0.010	0.010

Notes: GMM estimation results for the covariance structure of residual wage rates.

spouses, see Bredemeier et al. (2019) and Groneck and Wallenius (2021) for PSID estimates, and Chiappori et al. (2018) for estimates using the British Household Panel Survey.[21] The estimated variances are similar to typical values used in the literature. The implied correlation of income shocks $\omega = s_{\epsilon_m \epsilon_f}/(s_{\epsilon_m} \cdot s_{\epsilon_f})$ equals 0.196 and is similar to the estimate reported in Hyslop (2001).

Next to the stochastic component, the wage process in our model includes constant terms to account for long-run wage differences, including the gender wage gap. We quantify these constant terms by calibration as described in the main text.

[21] Bredemeier et al. (2019) estimate $\rho_m = 0.84$ and $\rho_f = 0.81$ and Groneck and Wallenius (2021) find values between 0.80 and 0.87. Chiappori et al. (2018) report estimates between 0.83 and 0.91.

A.5 Categorization of consumption items

Table 7 summarizes how we map the different consumption items covered in the PSID into the categories used in our analysis. Total consumption is the sum of expenditures on all items. As discussed in the main text, our baseline classification consists of two consumption bundles, where one bundle contains goods and services more strongly preferred by men and the other bundle consists of goods and services more strongly preferred by women. To assign the individual consumption items to either bundle, we use a sample of bachelor households and run, for each item, a regression with the expenditure share for this item on the left-hand side and age, education, and gender dummies on the right-hand side. We identify an item as more strongly preferred by women if the estimated coefficient on the female dummy is positive.

Table 7: Categorization of PSID consumption items into consumption bundles

		baseline		ND/S		NIPA							
	Total	A (female)	B (male)	Nondurable	Services	Food off-premise	Gasoline/energy	Housing/utilities	Health care	Transportation	Food/accommodations	Finance/insurance	Other services
Food at home and food delivered	×		×	×		×							
Food out of home	×		×		×						×		
Food stamps	×	×		×		×							
Gasoline cost	×	×		×			×						
Health insurance	×	×			×							×	
Out of pocket for nursing home, hospital	×		×		×				×				
OOP for doctor, outpatient surge, dental	×	×			×				×				
OOP prescription, in-home mc	×				×		×		×				
Home owner insurance	×	×			×							×	
Electricity expenditure	×	×			×			×					
Heating expenditure	×	×			×			×					
Water expenditure	×	×			×			×					
Other utilities (e.g., phone and cable)	×	×			×			×					
Car insurance	×	×			×					×			
Car repair	×	×			×					×			
Parking and car pool	×	×			×					×			
Bus and train fares	×				×					×			
Taxi fares	×		×		×					×			
Other transportation expenditure	×		×		×					×			
Tuition room and board (not day care)	×	×			×								×
Other school related expenses	×		×		×								×
Child care expenses	×	×			×								×
Monthly rent (or rent equivalent)	×	×			×			×					

Notes: ND/S = non-durables/services, OOP = out of pocket. Mapping of PSID items into NIPA categories based on Andreski et al. (2014) and Garner et al. (2006).

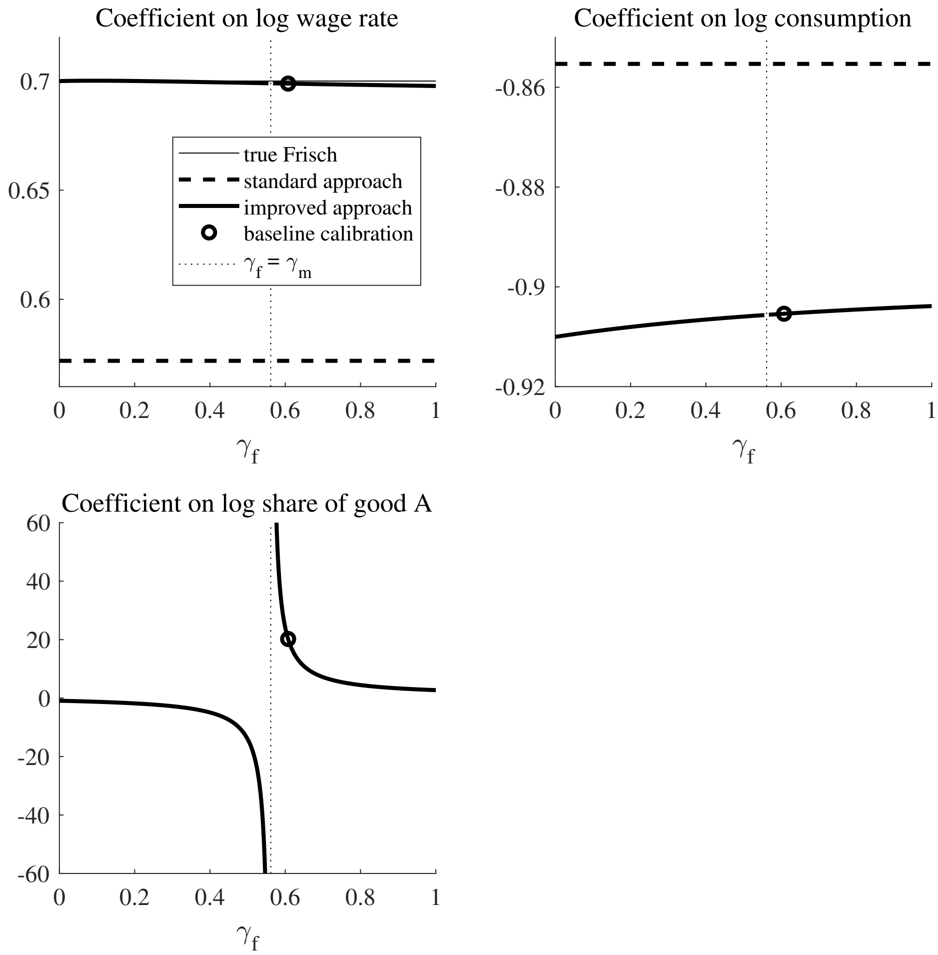
A.6 Sensitivity analysis for Monte-Carlo analysis

In a first series of sensitivity checks, we vary gender differences in the preferences over the two consumption goods. Specifically, we keep the male preference parameter γ_m at its baseline value of 0.56 and vary the female preference parameter γ_f over the entire range from zero to one (the baseline value being 0.61). Figure 1 shows the mean estimates for the three coefficients in regressions of condition (21), i.e., a regression of log hours on the log wage rate, log household consumption, and the log share of consumption expenditures for good A (which here is not necessarily the good more preferred by women since we vary exactly these preferences). The estimates in the conventional approach (11) are not affected by the changing preferences over different consumption items, which is reflected by the dashed lines.

Our main interest lies on the coefficient on the log wage rate which is shown in the upper-right panel of Figure 1 and corresponds to the estimate of the Frisch elasticity. The dotted horizontal line highlights the case where preferences over the two goods are identical across spouses. If this is the case, the expenditure share on good A is a constant and our estimation approach is not feasible. The circle marks the baseline calibration and the corresponding estimate of 0.699. Reducing the weight on good A in wives' preferences increases the estimated coefficient and hence the precision of the Frisch-elasticity estimate. This is because the closer γ_f is to zero, the closer equation (18) gets to a proportional relation between the expenditure share on good A and the consumption share of, e.g., the husband. When (18) is a proportional relation, the log-linear approximation (19) holds with strict equality. As a consequence, there is also no approximation error in the labor-supply condition (20). More importantly, the figure shows that even smallest differences in preferences, as long as they are non-zero, are sufficient for our improved estimation approach to deliver very good estimates of the Frisch elasticity. Even for the least favorable calibration, where γ_f is close to one and the approximation error in (19) is maximally large, the approach yields a bias of less than one percent.

Consistent with the above, the coefficient on log consumption (upper-right panel) takes the theoretically predicted value of $\sigma_m \eta_m = 0.91$ (see (20)) when $\gamma_f \to 0$. For larger values of γ_f, the coefficient on log consumption increases with the approximation error in (20). The lower-left panel of Figure 1 reflects that the expenditure share on good A enters the labor-supply regression as a proxy for bargaining positions. When $\gamma_f > \gamma_m$, a high share of expenditures on good A indicates strong bargaining power of the wife and these are situations where the husband works long hours – accordingly, the coefficient on the proxy is positive. The reverse holds for $\gamma_f < \gamma_m$. The more similar the two preference parameters are, the less the expenditure shares on certain categories of goods react to bargaining weights and the

Figure 1: Monte-Carlo results for men for different values of γ_f.



Notes: Solid lines show average coefficients on $\log w_{mjt}$ (top panel), $\log c_{jt}$ (bottom left panel), and $\log(c_{jt,A}/c_{jt})$ (bottom right panel) from 10,000 estimations of (21) for $g = m$ with samples obtained from simulated model data. Dashed lines show average coefficients from regressions of log hours worked on the log wage rate, log household consumption, and fixed effects. Simulations are performed for different values of γ_f.

proxy relation is thus relatively flat, which makes the estimated coefficient large in absolute value.

We now turn to parameters that affect both the conventional and our improved estimation approach. Regarding risk aversion, there is some consensus in the empirical literature that women tend to be more risk averse than men, but the size of this difference is unknown. In our baseline calibration, women are 30% more risk averse than men. In the first two sensitivity checks, we shut off and enlarge to 100%, respectively, this gender difference, each time pertaining the average value (1.5) of risk aversion from the baseline calibration. Columns (2) and (3) of Table 8 show that the bias in the conventional approach is largest when risk aversion is identical for women and men but, also with the strong differences, the bias remains substantial. The estimate from our improved estimation approach is almost unaffected from these two recalibrations. Next to gender differences, also the *average* level of risk aversion in the economy is debated in the literature. In columns (4) and (5) of Table 8, we set *average* risk aversion to a rather low value of one and a rather large value of two, respectively, while keeping fix the baseline difference between men and women. The bias in the conventional estimation is not substantially affected by these recalibrations. The impact on the estimate from our improved estimation approach is again minimal.

To explore the sensitivity of our results to the relation between bargaining weights and relative wages, we run two specifications where we vary this calibration target. In column (5) of Table 8, we add two standard deviations to the Lise and Yamada (2019) estimate, which reduces direct utility gains from marriage and hence amplifies commitment problems. In column (6), we subtract two standard deviations. As one would expect, the bias in the conventional approach is larger for the specification with more severe commitment problems (column 5), but differences between specifications are moderate. Again, the recalibrations have little impact on the results of our improved estimation approach.

Next, we vary the correlation between the stochastic wage components of husband and wife. We run specifications where we add and subtract, respectively, two standard deviations to our point estimate for the covariance between spouses' wage shocks, see Table 6. Columns (7) and (8) of Table 8 indicate that the bias in the conventional approach is somewhat less pronounced when wages are more strongly correlated. Our improved estimation approach continues to deliver a virtually unbiased estimate, independent of the wage correlation.

Finally, we estimate the two estimation approaches with simulated data from a model that has full commitment between spouses. This case is achieved by setting $\chi_{high} = 1$. The results are shown in column (9) and corroborate that the conventional approach *would* yield an unbiased estimate of the Frisch elasticity *if* there were full commitment. Importantly, the lower block of the Table shows that our extended estimation approach works well also in a setting with full commitment. In such a setting, the consumption composition is not constant as, with rising consumption, the household consumes less of the good more preferred by the more risk averse spouse. In the regression, the expenditure-share variable picks up the correlation between hours and the composition of consumption driven by the wage shocks of the individual's partner but does not influence the estimated coefficient on the wage rate. Hence, applying our approach to estimate the Frisch elasticity does not impose an a-priori assumption on commitment in the household.

Table 8: Sensitivity analysis for Monte-Carlo analysis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Conventional labor-supply estimations									
log wage rate, log w_{mjt}	0.553	0.600	0.570	0.572	0.560	0.583	0.576	0.568	0.701
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
log household consumption, log c_{jt}	-0.889	-0.760	-0.538	-1.172	-0.842	-0.869	-0.856	-0.855	-1.012
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.000)
bias	-0.210	-0.142	-0.186	-0.183	-0.200	-0.167	-0.176	-0.188	0.001
Improved estimation approach									
log wage rate, log w_{mjt}	0.697	0.702	0.698	0.699	0.698	0.699	0.699	0.698	0.700
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log household consumption, log c_{jt}	-1.043	-0.699	-0.601	-1.209	-0.905	-0.906	-0.907	-0.903	-0.701
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log share of goods preferred by women, $\log(c_{Ajt}/c_{jt})$	23.159	15.932	13.480	27.120	20.164	20.277	20.212	20.244	61.648
	(0.020)	(0.023)	(0.016)	(0.024)	(0.020)	(0.019)	(0.020)	(0.020)	(0.172)
relative bias $(\widehat{\eta} - \eta)/\eta$	-0.004	0.003	-0.003	-0.001	-0.002	-0.001	-0.001	-0.003	0.000
change to baseline	$\sigma_m = 1.5$ $\sigma_f = 1.5$	$\sigma_m = 1$ $\sigma_f = 2$	$\sigma_m = 0.867$ $\sigma_f = 1.133$	$\sigma_m = 1.733$ $\sigma_f = 2.267$	$\partial \log x / \partial \Delta z = 0.368$	$\partial \log x / \partial \Delta z = 0.308$	$\omega = 0.390$	$\omega = 0.004$	$\chi_{high} = 1$
fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
number of observations	20,000	20,000	20,000	20,000	20,000	20,000	20,000	20,000	20,000

Notes: Results in columns (1) through (8) and (10) are for men. Results in column (9) are for women. The true Frisch elasticities in the model are $\eta_m = 0.7$ and $\eta_f = 1.05$. The dependent variable is (logged) hours of the male in household j in period t. The table shows average estimates from 10,000 Monte-Carlo draws, with average standard errors in parentheses. $\Delta z = z_m - z_f$.