Heterogeneous Preferences, Spousal Interaction, and Couples' Time-Use

Almut Balleer*, Monika Merz†, Tamás K. Papp‡

Abstract

We study the role of heterogeneous preferences at the spousal level and that of spousal mutual insurance against wage shocks for couples' labor supply in a time-use model. We estimate the model for couples in the German Time-Use Survey with Bayesian techniques and generate gender-specific wage-elasticities of market hours in the cross-section and for all couples. Missing mutual insurance ignores males' behavior which we find to be substantial. Together with preference heterogeneity it shapes the elasticities' size and distribution, especially for high and low hours worked and wage groups. Our setting is suitable to analyze nonlinear and distributional economic policy.

JEL-Classification: D13, E24, J22

Keywords: preference heterogeneity, spousal labor supply, Bayesian estimation, Marshallian wage elasticity

*RWI - Leibniz Institute for Economic Research, TU Dortmund, IIES, and CEPR; Email: almut.balleer@rwi-essen.de

†University of Vienna, CEPR, and IZA; Email: monika.merz@univie.ac.at

‡Institute for Advanced Studies, Vienna; Email: tpapp@ihs.ac.at

1 Introduction

The total supply of hours worked in the market is central to the production of goods and services. At a given point in time it depends on the underlying distribution of preferences, wage rates and reservation wage rates across the workforce, since they determine the actual time-allocation. When an individual's labor supply interacts with that of her partner, e.g., because partners mutually insure against adverse wage shocks, her reservation wage rate depends not only on her own tastes and non-labor income, but also on her partner's wage rate. In fact, the bulk of market hours typically is supplied by individuals living in couples. Therefore, we need to understand how the *joint* distribution of preferences, wage rates and reservation wage rates affects spousal time-allocation and the associated own- and cross-wage elasticities of market hours within and across couples.

In this paper, we depart from the observed time-allocation of spouses in actual couples according to the German Time-Use Survey of 2001/02. In Germany, about two-thirds of individuals of prime working-age, i.e., between 25 and 54 years old are married or cohabiting. They are well represented in this survey. We measure the time they allocate across market work, homework, and leisure.[1] Individual hours devoted to home production matter for the total amount of goods available for consumption and for measuring leisure. Since the data report each activity in a broadly or a narrowly defined sense, we can distinguish, e.g., core homework from homework including childcare. We also have wage information for each employed person and non-labor income for each household.

The couples in our sample display a wide variety of time-allocation choices. There co-exist dual-career couples, more traditional ones where the male works in the market and the female stays at home, less traditional couples with roles switched, as well as those where neither partner works.[2] Figure 1 illustrates a striking fact which motivates our analysis: Time-allocation clearly differs across couples of different labor market states. However, even within couples with at least one spouse employed, the observed time-allocation is very diverse.

Treating spousal time-allocation as mutual insurance device against adverse individual wage changes, we ask whether preference heterogeneity — in addition to wage heterogeneity — among spouses helps replicate the observed diversity in couples' time-allocation. We also study whether and how preference heterogeneity affects spouses' own- and cross-wage elasticities of labor supply measured along the intensive and the extensive margin both in the cross-section and for any couple type.

We make three important contributions. First, taking spousal time-allocation as a mutual insurance device, we model preference and wage heterogeneity at the spousal level. Second,

[1] All couples in our sample consist of a male and a female. To simplify language, we use partners and spouses interchangeably irrespective of their marital status. We commonly refer to the female partner as wife and to the male partner as husband. For similar reasons, we use the term *preferences* to capture actual preferences as well as other economic determinants of individual time-allocation not otherwise formulated in our model.

[2] 61 percent of all couples are dual-career, 27 percent are traditional ones. Couples with only the female or no one employed each constitute six percent of the total.

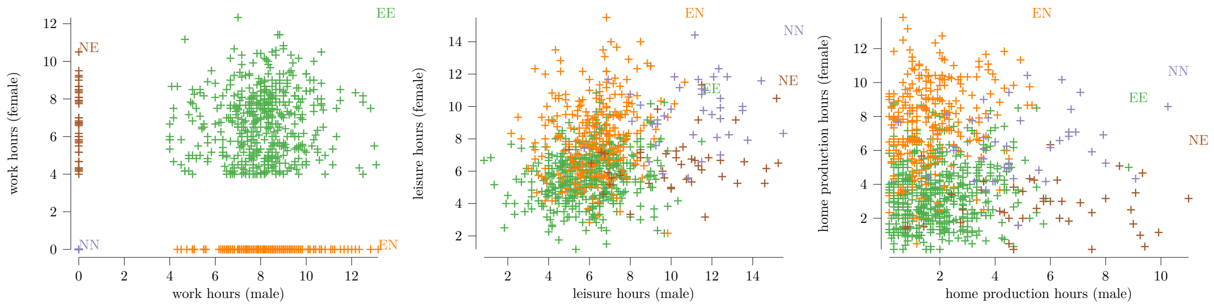

Figure 1: Scatterplots of average daily work, leisure, and home production hours by each partner's labor market status (EE: both employed, EN: male employed, female non-employed, etc.)

we structurally estimate these heterogeneous parameters for all spouses, using time-use data on actual couples and a Bayesian multilevel model. Our model replicates the observed heterogeneity in time-use across and within couples of a given labor market status. We show that wage heterogeneity alone cannot reproduce the patterns we observe in the data. Third, we use the estimated model for generating own- and cross-wage elasticities of market hours by gender in the cross-section and for all couples.

We subsequently perform two types of counterfactual experiments to further explore the role of preference heterogeneity and of spousal mutual insurance via time-use adjustment. In one experiment, we re-estimate the model using gender-specific homogeneous preferences. This allows us to contrast the implied elasticities to those from Blundell, Pistaferri, and Saporta-Eksten (2016) who assume heterogeneous wages, spousal interaction, but homogeneous preferences. To elicit the importance of spousal mutual insurance, we further restrict our model by eliminating the possibility for males to react to their spouse experiencing a wage shock. This setting permits us to compare our results to those from Attanasio et al. (2018) who study married women's labor supply behavior while taking their husband's income as given.

We address our research question using a non-cooperative model of spouses' time-allocation decisions.[3] Agents endogenously sort into market work, or homework and leisure, where market work is subject to a positive lower bound.[4] The implied equilibrium outcome features dual-career couples, those with only one spouse employed, and couples where neither partner works in the market. The equilibrium typically is unique, but inefficient. For the sake of our positive analysis efficiency is not essential. The model is static, because we focus on and exploit the rich heterogeneity in couples' time-use. We report the relative change in hours for continuously employed spouses and also the transition probabilities of those who change their labor market status when wages change. The implied wage-elasticities are long-run Marshallian elasticities. Throughout this paper, we take couples and their members' individual

[3]Our model builds upon the analytical framework of Del Boca and Flinn (2012) who launched the idea of spouses interacting in their time-allocation.

[4]This lower bound is a simple way of modelling fixed costs of taking up market work. It helps replicate the lack of very low market hours worked in our data.

characteristics as given. In this sense, all of our results are conditional on the observed status quo processes of family formation, fertility, or education.

We estimate the parameters of the structural model using multilevel Bayesian methods. The approach has several advantages. It allows us to directly address the cross-sectional parameter heterogeneity (in wages and preferences). Moreover, we can model hierarchical data, i.e., we can handle observations on the individual and on the couple level as well as day-to-day variation of time use observations as natural part of the empirical model. Furthermore, the posterior sample allows us to quantify the parameter uncertainty even with small sample size. Finally, the structural model is not approximated, but evaluated exactly. The posterior distribution then propagates the nonlinear mapping from parameters to hours. Methods that provide point estimates or use linear approximations do not fully account for non-linearities. Hence, counterfactual calculations also reflect the non-linearity which then also propagates to estimated elasticities.

We briefly summarize our main findings. Preference heterogeneity at the spousal level in addition to wage heterogeneity is essential for replicating the observed dispersion in couples' market hours. Together with spousal mutual insurance it also matters for generating gender-specific wage-elasticities in the cross-section and for all couples. Without preference heterogeneity, own-wage elasticities are predicted to be different especially at very high or very low wages. Similarly, lack of mutual insurance underpredicts females' own-wage elasticities at all wage levels. As for adjustments along the intensive margin, males' and females' responses are highly symmetric, but lower in absolute values for females and also for spouses of a given sex in single-earner couples compared to dual earners. The adjustment along the extensive margin is much more asymmetric across gender and also across couple types.

Our results show why modelling heterogeneity at the spousal level matters for macroeconomics and especially for macroeconomic policy analysis that involves the labor market. Mean or median wage-elasticities for men and women are similar and of plausible size across different model specifications, implying that the predicted average impact of a particular policy may be similar. Yet, their incidence obviously differs for spouses depending on their wage rates or hours worked, and also on the couple's labor market status. Our framework contains crucial ingredients for studying the implications of particular non-linear policies at the spousal level and their impact on labor supply. For example, by introducing labor income taxes or transfers, we could assess their incidence on different types of couples. This would be suitable for identifying the couple types which are most likely to change their market hours worked in reaction to such policies, thereby addressing the ongoing labor shortage in the German labor market.

The paper proceeds as follows. Section 2 relates to the literature. Section 3 introduces details of the German Time-Use Survey. Section 4 presents the model setup, while Section 5 lays out the estimation strategy. Section 6 discusses the results. Finally, Section 7 concludes.

2 Related Literature and Contributions

The main contribution of our work relates to the growing literature on family labor supply that studies its determinants and quantitative implications in the cross-section, or for the aggregate economy. Existing papers consider spousal insurance against adverse labor market outcomes. They differ with respect to the underlying model of family labor supply, i.e., the type and extent of heterogeneity considered, the details of the spousal decision-making, their relevant time horizon, and sometimes also regarding closely related decisions such as education, couple formation, or fertility. None of them allows for individual preference heterogeneity. Studies of family labor supply typically emphasize the importance of each spouse's wage-elasticity of labor supply for the implications of public policies or welfare. This paper does not address policy, but focuses on estimating a structural model that replicates patterns in the data and then allows counterfactual experiments, including the calculation of wage elasticities. It is the first to study spousal time-allocation as a mutual insurance device against idiosyncratic wage-shocks while allowing for heterogeneous preferences — in addition to heterogeneous wages — at the spousal level and treating males and females in couples as equals, allowing them to endogenously sort into employment, or non-employment.

One strand of this literature uses a partial equilibrium life-cycle model of two-earner (male and female) families of the unitary type. It assumes that all family members have identical preferences and share the same objective and constraints, but their wage or earnings processes may differ.[5] Husbands are typically considered as primary earners, whereas wives are secondary earners. They can divide their time between leisure and market work. In this setting, Guner, Kaygusuz, and Ventura (2012) evaluate the effects of income tax reforms in the U.S. on the labor supply of married men and women; Blundell, Pistaferri, and Saporta-Eksten (2016) use data from the PSID and the Consumer Expenditure Survey (CEX) of the U.S. to quantify the importance of family labor supply — in addition to savings and governmental transfers — as consumption insurance device against idiosyncratic wage shocks to males or females. Their sample consists of stable married couples with continuously employed males; Attanasio et al. (2018) take husbands' earnings as given when studying married females' labor supply in connection with couples' savings and consumption decisions. They aggregate across couples to illustrate how the aggregate wage-elasticity of labor supply varies with the underlying type and degree of heterogeneity, including the distribution of reservation wages in the cross-section; Wu and Krueger (2021) investigate whether a life-cycle two-earner household model with endogeneous labor supply can replicate the consumption and labor supply responses to transitory and permanent wage shocks estimated by Blundell, Pistaferri, and Saporta-Eksten (2016); Birinci (2019) allows spousal labor supply to depend on governmental transfers that vary over the business cycle. He finds that in case of a male's job displacement public insurance partially crowds out private insurance in form of the wife increasing her labor supply; Golosov et al. (2021) study the effect of unexpected wealth and

[5]As Blundell, Pistaferri, and Saporta-Eksten (2016) point out, assuming identical preferences in a dynamic model helps avoid the difficulty of identifying heterogeneous preferences in a dynamic setting.

unearned income on family consumption and labor supply. We share with these papers the focus on family labor supply as insurance device against idiosyncratic labor market shocks. Our paper differs in that we model heterogeneous preferences and heterogeneous wages — potential or actual — at the spousal level, but abstract from time and age. Heterogeneous preferences within couples are essential for replicating the heterogeneity in time-allocation observed even among couples of similar labor market states. To insure the couple's consumption against idiosyncratic wage shocks, spouses can adjust their time-allocation across market work, homework, and leisure. We treat spouses equally and do not *ex ante* distinguish between prime earners and secondary earners, thereby allowing both of them to sort across employment and non-employment.

A second vast strand of this literature explicitly considers individual spouses with their respective objectives and constraints and allows them to interact in their decision-making.[6] Spousal interaction can be cooperative, or non-cooperative. Although a bargaining process typically is not specified, the implied allocations can be interpreted as if bargaining under given outside options had occurred. Cooperative models, also known as collective models, consider marriage as a cooperative game where partners settle on Pareto optimal outcomes.[7] Following this line of research, Goussé, Jacquemet, and Robin (2017) use a collective model of household consumption and individual time-allocation to study how family values affect the mating and time-allocation decisions. Preference heterogeneity relates to groups of males and females that are characterized by their market wage, education, and family value. The authors endogenize the consumption sharing-rule by explicitly modelling a marriage market for singles, allowing marriages to dissolve. They structurally estimate the model using data from the British Household Panel Survey from 1991 to 2008 and retrieve Marshallian type own- and cross-wage elasticities for male and female market hours; Obermeier (2023) models unobserved preference heterogeneity across spouses in a finite-horizon collective model of spousal market consumption and time-use. He adds a dynamic marriage market to an otherwise static decision problem which helps identify the joint distribution of partners' preferences. He estimates the model using indirect inference and UK data on individual time-use, consumption expenditures, and income. He uses his setting for welfare comparisons at the individual and household level. Our work differs from his in that we use Bayesian techniques to structurally estimate our model and that we elicit a broad spectrum of wage-elasticities in the cross-section and in the aggregate. Non-cooperative models assume partners play best responses as outlined, for example, by Lundberg and Pollak (1994).[8] Following John Nash's logic who argued that any cooperative model should be preceded by a non-cooperative

[6]We limit our survey below to static model versions, since they are closest to our work.

[7]Prominent early examples of divorce-threat cooperative models that take divorce or remaining single as outside option include McElroy and Horney (1981), and Manser and Brown (1980). Collective models were pioneered by Pierre-Andre Chiappori (1988) and Pierre-André Chiappori (1992), Apps and Rees (1988), and Browning and Pierre-Andre Chiappori (1998). They extend their predecessors by treating variations in the internal distribution of power as exogenous, or stemming from a search model of the marriage market. All these models are static, and they feature preference heterogeneity across males and females.

[8]Other early examples of non-cooperative models at work are Del Boca and Flinn (1995) and Del Boca and Flinn (2012).

one in order to establish outside options for the parties involved, non-cooperation while maintaining the relationship is a legitimate alternative. In fact, Gobbi (2018) and Doepke and Tertilt (2019) independently show that such models perform well in accounting for empirical patterns related to childcare and expenses on children. We share with these papers that spouses interact in their time-allocation decisions and that their reservation wages affect the outcomes. Our paper extends the existing work by modelling preference heterogeneity at the spousal level und estimating the full distribution of individual parameters with the help of a multi-level Bayesian procedure. The estimated model not only accounts for the empirical patterns of couples' time-use in the cross-section, it also serves as a lab to quantitatively assess the role of preference and wage heterogeneity for labor supply elasticities in general and the extent of mutual spousal insurance in particular.

3 The German Time-Use Survey

The German TUS is a quota sample survey of all private households in Germany that is designed and carried out by the Federal Statistical Office (Destatis)[9]. The quotation is based on the German census. Excluded are homeless people and individuals living in group quarters or similar living institutions. Participating households enter voluntarily. Time-use surveys exist for three independent waves, namely 1991/92, 2001/02 and 2012/13. The first wave cannot be used for our purposes, since it does not contain information on usual hours worked or on income which is necessary for estimating our model. The latter two waves comply with Eurostat's recommendations regarding the harmonization of time-budget surveys, and therefore are comparable with the content of the MTUS. For each wave, the reference period ranges from April of the earlier year to the end of March of the subsequent year in order to avoid seasonal distortions. The original data consist of three survey documents which we merge into our baseline dataset: information at the household level, each household member who is at least 10 years old provides socio-economic information about herself, and the same individual also keeps a diary over 24 hours on each of up to three days including both weekdays and weekends. These diaries contain activities in intervals of ten minutes. We use the 2001/02 wave for our baseline analysis. We compare our main results to the respective figures from the 2012/2013 wave in a robustness check. We aggregate the individual records journalized in the diaries to daily measures of activities we need for our model estimation. Via the household dimension, we can identify couples and have detailed information about each spouse's time-use.

Our sample contains couples with partners each of whom is between 25 and 54 years old, i.e., in their prime working-age. We exclude from our sample couples with children below 6 years. Our model abstracts from children and also from time spent on childcare. Young children are known to impose a large tax on a couple's time-use and significantly affect partners' time-allocation.[10]

[9]Seehttps://www.destatis.de/EN/FactsFigures/SocietyState/IncomeConsumptionLivingConditions/TimeUse/TimeUse.html for a detailed description of the data.

[10]We plan to explore this topic in future research. There exist household models in which young children are captured as a public good that both partners can enjoy and to which they have to contribute goods or available time in order to foster them. See, e.g. Blundell, Pierre-Andre Chiappori, and Meghir (2005), or Doepke and Tertilt (2019).

Table 5 in Appendix A reports the relative frequency of different types of couples in their prime working-age in the German TUS (using the appropriate representative weights) and contrasts them against the respective figures from the German microcensus.[11] The entries show that over 60 percent of all individuals in the indicated age-range live in couples, and that the vast majority of them are couples without children younger than six years. This group, which is the object of our study, is representative in the TUS compared to the microcensus. Couples with children younger than six years are over-represented in the TUS, while persons living in other conditions than single or couples are under-represented by design.

We define and compute three categories of time use: market work, home production and leisure. In doing so, we follow Aguiar and Hurst (2007) as closely as possible and distinguish between a core activity and a more broadly defined activity. Core market work comprises time spent in the main or secondary job as well as training on the job. Total market work adds related activities such as searching for another job, taking breaks and commuting. We will use core market work in our estimation below. We further exclude persons with core market work of less than 4 hours a day, since this ensures a reasonable distinction between market work and non-market work in our model estimation below. Core home production encompasses preparing meals and maintenance activities in the home. Total home production adds shopping, gardening, construction and childcare. These categories refer to the primary task that is carried out during the assigned time interval. Since we cannot separately measure care for elderly or handicapped in home production, we deviate from Aguiar and Hurst (2007) by including these activities in home production. We will use total home production in our estimation below. We compute daily leisure as a residual by subtracting six hours for sleep and personal care, core market work and total home production from 24 hours. We consider only regular working days in our sample.

We categorize the couples in our sample by each partner's labor market status: both partners work in the market, only the man works, only the woman works, and no partner works. Not working encompasses both the formal definitions of being unemployed and out-of-the-labor force. We discard unreasonable work hours per day, i.e., more than 14 hours of core market work and more than 16 hours of total market work, or less than two hours.

Apart from spouses' time-use, the German TUS provides information on each spouse's individual characteristics. It also allows us to infer individual hourly wage rates as well as the household's non-labor income. These variables are crucial for estimating our model. In order to obtain individual earnings, we construct the wage income from the main job. When only bracketed information is available, we use the mid-point of the bracket as an approximation for the earnings. We then compute the hourly wage rate by dividing wage income from the main job by usual hours worked. We discard unreasonably high hourly wages, i.e., wage rates above 200 Euros. We take total household income from the survey and compute the household's non-wage income as the difference between total household income and the sum of the individual wage incomes. All wages and income are net of taxes.

[11]See https://www.forschungsdatenzentrum.de/en/household/microcensus

couple status	gender	market work		home production			leisure
		core	total	core	total	child	
EE	male	7.958	9.262	0.590	1.924	0.162	8.118
	female	6.612	7.557	1.760	3.363	0.334	8.025
EN	male	7.950	9.226	0.477	1.715	0.250	8.335
	female	0	0.112	4.183	7.457	0.902	10.54
NE	male	0	0.0808	2.020	5.419	0.473	12.58
	female	6.517	7.388	1.471	2.950	0.232	8.533
NN	male	0	0.182	1.460	4.977	0.109	13.02
	female	0	0.118	3.445	6.187	0.337	11.81

Table 1: Average daily time use 2001/2002. Figures show daily averages of time-use aggregates in hours. In home production, the sum of core and child is less than total home production, where child denotes childcare. Couple status refers to employment, E, or non-employment, N, of male and female partners. All numbers are averages using representative weights at the household level.

Our baseline sample encompasses 1,870 spouses in 935 couples and a total of 3,142 observations. Table 1 shows the average daily time-use of couples according to their labor market status. When both partners work, women work less in the market and more at home compared to their partners, while both enjoy a similar amount of leisure. When only one partner works, the other works more at home and enjoys more leisure. However, if women are the sole wage-earner, they provide fewer market hours and more home production than men in case they are the sole earner. As is to be expected, childcare is negligible as a primary component of home production if kids are older than 6 years.

Table 6 in Appendix A exhibits unweighted means, standard deviations, percentiles and min and max values of daily core market work, total home production and leisure by gender for the full sample and by couple type. Table 7 further documents correlations between these key time use variables. Market work is generally negatively related to home production and leisure for individuals. Own home production is weakly positively related to the partners market work, own leisure is weakly negatively related to the partners market work. Home production of partners in a couple are only very weakly positively related, but leisure of partners in a couple is positively related. Hourly wage rates of spouses are barely correlated (−0.0077, not shown in Table) indicating little assortative mating by their respective productivity in our sample.

Table 8 in Appendix A documents labor and non-labor income as well as age and educational degree of the couples in our sample sorted by their respective labor market status. Even for couples of the same labor market status the variation in wages and income is high. Women tend to earn substantially lower market wages than men. Also, couples with no partner working in the market or only the woman working have substantially higher non-labor income than others. Couples with no partner working tend to be somewhat older than other couples. Women are on average a few years younger than their male partner. Table 8 also reports the educational achievement of men and women by couples' labor market status. Men tend to be more highly educated than their female partners. Education is highest among dual-career couples.

Table 9 in Appendix A shows the main source of income for couples according to their labor market status. For dual-career and traditional couples, EE and EN, respectively, the main source of income is wage income. The main source of non-wage income are pensions and unemployment benefits. In addition to the variables already mentioned we use information on whether or not a couple is married and whether they reside in the east or the west of Germany for our empirical work.

4 The Model

We model each couple as a pair of a male m and a female f, who interact in the allocation of their available time and also in their goods consumption. The model is static. We take couples as given and consider neither their mating or marriage decisions, nor their decisions to maintain the relationship or break up. Members of a couple gain from a partnership, because they can at least partially specialize in the type of goods production in which they have a comparative advantage and subsequently consume more goods than if they remained single.[12]

First, we describe the economic environment. Then we describe the solution under the *non-cooperative Nash equilibrium*, in which members of couples optimize taking the strategy of the other party as given.[13]

4.1 The economic environment

The economy consists of couples, comprised of two individuals, which we label *male* and *female* for notational convenience. We index couples with $j \in C$, but suppress this in this section and we focus on the decision problem of a given couple. Each individual $i \in \{m,f\}$ in a couple can allocate his or her available time T between market work, n_i, home work h_i, and leisure ℓ_i, thus facing the time constraint:

$$\ell_i + h_i + n_i \leq T. \tag{1}$$

Both individuals can choose between *non-employment*, $n_i = 0$, and *employment*, $n \in [n_0, T]$, where n_0 is the minimum number of work hours that are allowed by the model. We introduce the latter constraint to account for the fact that very low hours of employment (e.g. one hour of work per day) are atypical in the data.

Individual consumption comprises goods that are either purchased in the market, c, or domestically produced, z, using home work as sole input. Due to the lack of available data on consumption expenditures and home-produced goods, we assume both types of consumption to be public goods. Each partner can voluntarily contribute to the production of these goods. Market consumption goods are purchased using total non-labor income M, plus total earnings $w_m n_m + w_f n_f$, plus unemployment benefits $\rho w_i T$ for each member when $n_i = 0$, where w_i denotes the net hourly real wage rate of individual i and ρ parametrizes the unemployment

[12]They may also gain from economizing on household maintenance costs, but we do not explicitly model them.

[13]A table summarizing notation is available in Appendix B.

benefit that is proportional to the individual's wage.[14] Hence, we assume partners in a household to pool their income, since we have information on individual earnings if employed, but not on the individual share of non-labor income. The household faces the budget constraint

$$c \le M + w_m n_m + w_f n_f + \mathbf{1}_{n_m=0} \cdot \rho w_m T + \mathbf{1}_{n_f=0} \cdot \rho w_f T \tag{2}$$

where $w_i n_i$ denotes the wage income of each individual, and the last two terms account for the unemployment benefits. Given that our approach is static, we model M, w_m, and w_f as exogenous.

Without loss of generality, we normalize the price of the market good to unity. The domestic good z is nontradable, and its production is captured by a Cobb-Douglas home production function:

$$z(h_m, h_f) = h_m^{\gamma_m} h_f^{\gamma_f}, \tag{3}$$

where

$$\gamma_m + \gamma_f = 1 \qquad \text{and} \qquad 0 \le \gamma_m, \gamma_f \le 1$$

characterize the home production, for symmetry of the formulas it is convenient to use both γ_m and $\gamma_f = 1 - \gamma_m$. This particular function treats male and female time in home production as partially substitutable. Consistent with the empirical evidence on actual time use of couples, it ensures that in equilibrium, each spouse contributes some positive amount of home work.

Individual preferences are defined over a market consumption good, and a non-market consumption good plus leisure. They are captured by a Cobb-Douglas utility function that is continuous, linear homogeneous and strictly concave conditional on employment status. The parameter α_i denotes individual i's utility weight on market consumption, and $1-\alpha_i$ captures the weight on non-market consumption and leisure, which are aggregated using a Cobb-Douglas form with weights β_i and $1-\beta_i$ on the non-market good and leisure, respectively.

We assume that a stochastic term ζ_i that is added to the utility of each *non-employed member*, with $E[\zeta_i] =$, where ζ_m and ζ_f are identically and independently distributed. The purpose of this is to account for the fact that seemingly similar couples have different employment patterns.

Consequently, we model each individual's utility as

$$U_i(c, z, \ell_i) = \alpha_i \log(c) + (1-\alpha_i)(\beta_i z + (1-\beta_i)\ell_i) + \zeta_i \mathbf{1}_{n_i=0} \qquad \text{for } i = m, f \tag{4}$$

To make the model well-behaved, we also introduce the condition

$$\rho T \le n_0. \tag{5}$$

This ensures that choices are monotonic in w and M. Intuitively, without this condition, non-employment can be "too attractive", so in a region of low wages an individual could switch from employment to non-employment then employment again as wages are increasing.[15]

4.2 Non-cooperative equilibrium

Assume that the partners forming a household interact non-cooperatively, in that each of them individually maximizes utility while taking their partner's decisions as given. Hence, each

[14] We use the term "non-labor income" for M for brevity, while technically it is income other than wages and unemployment benefits.

[15] See the proof of Lemma 1 in the Appendix for the exact role of this condition.

member $i \in \{m,f\}$ of a couple solves the following decision problem:

$$\max_{n_i,h_i,l_i} U(c,z,\ell_i)$$

subject to her individual time constraint (1), the budget constraint (2), the home production function (3), and several non-negativity constraints:

$$c,\ z,\ \ell_i,\ h_i > 0,\ n_i \in \{0\} \cup [n_0,T].$$

Thus, each member i of the household takes the leisure, home production, and market hours choices ℓ_k, h_k, n_k of the other member k as given. Reaction functions then provide two mappings

$$(\ell_m,h_m,n_m) \mapsto (\ell_f,h_f,n_f) \tag{6}$$

$$(\ell_f,h_f,n_f) \mapsto (\ell_m,h_m,n_m), \tag{7}$$

a fixed point of which is an equilibrium. Since the utility function (4) is separable in market hours n_i and the joint leisure-home production choice (ℓ_i,h_i), we can solve our problem in three steps:

1. Holding n_m and n_f fixed, we derive the optimal choices of (ℓ_i,h_i), $i=m,f$, and the indirect utility functions $\hat{U}_i(n_m,n_f)$, $i=m,f$.
2. Conditional on the employment status for each member, we obtain the utility of being employed and unemployed.
3. Using the indirect utility functions $\hat{U}_i$, we derive the reaction functions

$$n_m \mapsto n_f$$

$$n_f \mapsto n_m$$

   and find a fixed point, which yields an equilibrium.

The details of the solution procedure can be found in Appendix C.

5 Estimation

We use Bayesian methods to estimate the model, which we find advantageous for several reasons. Bayesian methods allow us to quantify parameter uncertainty, estimate the structural model *as is* without resorting to approximations, directly address the cross-sectional parameter heterogeneity (in wages and preferences) and incorporate covariate information at the couple and individual level, and perform consistency checks on the results. We explain these advantages in detail below.

First, the posterior sample we obtain allows us to *quantify the parameter uncertainty*, which is important given the sample size. Parameter uncertainty is reflected in all the results we report, and is especially relevant for the hours response and elasticity calculations in Section 6.4, since couple-level outcomes are determined by a nonlinear mapping described in Section 4.2 from the posterior distribution of parameters to hours, and the elasticity calculations are themselves nonlinear. Some parameters are unobserved, which means that posterior uncertainty should be propagated to outcomes. Propagating the uncertainty through nonlinear calculations is important because it can affect the moments of the distribution — anticipating results, in Section 6.4 we find that various elasticities derived from the model have a skewed distribution.

Methods that provide point estimates would not be able to deal with couple-level uncertainty in a consistent way. In order to illustrate this, consider a hypothetical couple where the preference parameters α_m, α_f have a posterior distribution that would make 20% of the unemployed females switch from non-employment to employment in response to a 10% wage increase, which would translate to a $\geq 0.2 \cdot n_0$ expected increase in hours.[16] In contrast, a method that would replace these parameters with point estimates might miss this effect entirely, depending on where they fall in the parameter range. Likelihood-based methods which use approximations around a mode are not sufficient for our purposes either, as the approximation used is usually normal-linear, which would not work well in a setting where the posterior distribution of couple-level parameters consistent with observed outcomes is not well-approximated by a normal distribution, which is the case for our model.[17]

Second, using Bayesian methods allows us to estimate the structural model exactly as described in Section 4 in a consistent manner, without resorting to approximations. Our data generating process may be nonlinear, complicated, and somewhat computationally intensive compared to simple reduced form models, but from a statistical point of view it is just a function which can be computed, and thus incorporated in the likelihood. This allows to capture interactions within the couple and also allows a form of unemployment benefits.

Third, Bayesian methods allow us to use a *multilevel model* with weakly informative priors for the cross-sectional parameters which is especially important because our model has a high number of preference and wage parameters in the cross-section. Multilevel methods automatically pool cross-sectional information and are a natural fit for the hierarchical nature of our data. They also allow us to estimate a conditional distribution of unobserved wages (for the non-employed) relying only on *a priori* exchangeability of individual-specific error terms.[18] Specifically, our setup has the following hierarchical layers: individual-specific parameters are assumed to be drawn *ex ante* from a common distribution that also depends on the couple- and individual-specific information we have in our data (such as education, age, marital status, etc), then, given the parameters for each couple, we allow the allocation implied by the model to be observed with a noise, allowing for the daily variation in time use patterns we observe in the data.

Fourth, our estimation method provides posterior estimates and predictions at the couple level, allowing us to check model predictions against data, which is important to ensure that the model specification is reasonable.[19] While we specify a multilevel model for the cross-section, this only serves to sharpen the couple-level estimates, but does not constrain them. Because of this, our estimation procedure retains the flexibility to provide an estimate of cross-sectional heterogeneity at the couple level based on the data. Contrast this, for example, with a simulated method of moments methodology which minimizes a vector of moments generated from cross-sectional

[16] Recall that n_0 is the minimum amount of work hours allowed in the model.

[17] Also, most of these methods have no theoretical guarantees for predictive accuracy when variable transformations are not continuously differentiable. This of course does not mean that other methods, e.g., bootstrap, cannot account for uncertainty. It is just that working with posteriors is very convenient, as we have already implemented the model solution for the estimation, and get to reuse it in the exact same form for calculating elasticities.

[18] For recent introductions to multilevel models, we recommend Gelman and Hill (2007), Snijders and Bosker (2011), and Hox, Moerbeek, and Van de Schoot (2017). The importance of exchangeability in hierarchical models is discussed in Bernardo (1996).

[19] We do this in Section 6.2.

parameters, which would lack couple-level predictions that can be tied to a particular couple.

However, it should be mentioned that the trade-off for the above-mentioned advantages is that Bayesian estimation for our model is computationally intensive. We have more than 900 couples in the data, and depending on employment status they have 7 to 9 couple-specific parameters each. With the multilevel parameters discussed in Section 5.1, this means we have approximately 7,000 scalar parameters, which we reduce somewhat with marginalizing the ζ's, but given that we solve for structural outcomes, this is still computationally expensive.[20]

We explain how we map the prediction of the model to observed data in Section 5.1. After discussion of the fixed parameters in Section 5.2, Section 5.3 describes our choices for multilevel and noise distributions, while Section 5.4 briefly summarizes the priors. Section 5.5 describes the couple-level likelihood, while Section 5.6 briefly touches on some technical details of the estimation.

5.1 Mapping the model to observations

The model introduced in Section 4 maps couple-level preference parameters and non-labor income $\alpha_{m,j}, \alpha_{f,j}, \beta_{m,j}, \beta_{f,j}, w_{m,j}, w_{f,j}, M_j$ to observed work hours $n_{m,j}, n_{f,j}$ (may be 0 for the unemployed), $h_{m,j}, h_{f,j}$ (home production) and $\ell_{m,j}, \ell_{f,j}$ (leisure), where j is the couple index. Of these, hours are directly observable for every couple, wages are observable for employed members of couples, and we only observe the *sum* of M_j, wage income, and unemployment benefits for each couple. However, this sum is only available in brackets, so we model the survey response as

$$D_j \sim \mathrm{N}\left(M_j + \underbrace{\mathbf{1}_{n_{m,j}>0} w_{m,j} n_{m,j} + \mathbf{1}_{n_{f,j}>0} w_{f,j} n_{f,j}}_{\text{wage income}} + \underbrace{\mathbf{1}_{n_{m,j}=0} w_{m,j} \rho T + \mathbf{1}_{n_{f,j}=0} \rho T}_{\text{unemployment benefits}},\ \sigma_\eta^2 \right) \tag{8}$$

where σ_η is the standard deviation of the IID noise for this process.

Recall that time use information is collected in 10-minute blocks, while the model delivers a non-negative real number below the total time endowment T for each member of the couple. Since we observe multiple days for each couple, which are not necessarily equal, we assume that daily hours are observed with an IID (within- and between couple) $\mathrm{N}(0, \sigma_\varepsilon^2)$ noise:

$$n_{j,i,\text{observed}} - n_{j,i,\text{model}} \sim \mathrm{N}(0, \sigma_\varepsilon^2) \tag{9}$$

$$h_{j,i,\text{observed}} - h_{j,i,\text{model}} \sim \mathrm{N}(0, \sigma_\varepsilon^2), \tag{10}$$

which are IID across couples $j \in C$ and for $i = m, f$. This ensures that the expected values are preserved. We estimate σ_ε^2. We assume a similar noise structure to the noise $\zeta_{i,j} \sim \mathrm{N}(0, \sigma_\zeta^2)$ for $i = m, f$, IID.

Finally, while we do not directly observe the share of unemployment benefits from non-wage income, the survey provides a question that asks the main source of this income. We make

[20] We have developed optimized Julia code that allows estimation on a PC within less than an hour (depending on chain length), which is available from the authors on request.

$u_j = 1$ iff the couple answers "unemployment benefits", and assume that

$$\Pr(u_j = 1) = \left(\frac{\text{unemployment benefits}}{\text{total non-wage income}}\right)^{\kappa} \tag{11}$$

where κ is another parameter we estimate. The purpose of this formulation is to make the likelihood smoothly differentiable and allow for the possibility that survey respondent error while still assuming a connection between the answer and the actual fact. It makes no perceptible difference to the results but improves convergence by regularizing the posterior a bit.

5.2 Fixed parameters

Here we explain our choices for parameters which are *not* estimated. Time is measured in hours, and we fix the total length of working hours at $T = 16$, which is consistent with our data. Implicitly, this assumes that individuals need at least 8 hours for sleep and personal care that do not qualify as market work, home production, or leisure. We also fix $n_0 = 4$, assuming that part-time work has to take at least 4 hours per day, and $\rho = 4.0/16.0 = 0.25$ which implies that public transfers are equivalent to 50% of 8-hour income. This approximates an average of the German replacement rate of 65% for recently unemployed persons and a lower bound to social security payments that is not tied to the previous wage.

The parameters γ_m, γ_m that describe home production are not separately identifiable from α_m,β_m and α_f,β_f at the level of the couple.[21] That is, for every allocation we observe at the couple level, there is a manifold of (α_m, β_m, γ_m, α_f, β_f, γ_f) values that can generate it. Statistically, once we assume particular cross-sectional and noise distributions, the γ_m,γ_f are *weakly* identified, but as usually happens in cases like this, this leads to very bad convergence of the MCMC estimator so we do not pursue this. Instead, we perform robustness checks on all of the fixed parameters in Section 6.5.

We set σ_η to match the standard deviation of a uniform distribution with the median bracket.[22]

5.3 Cross-sectional distributions

We also need to assume a functional form for the *ex ante* cross-sectional distribution of wages, and preference parameters. Since we would like to avoid overfitting the model, it is important to choose a simple functional form, but at the same time we would like to avoid ruling out possible *correlations* between preferences and wages, either for the same individual (e.g. between α_i, β_i, and w_i), or between spouses. In order to strike a reasonable balance between these two

[21]See (14) and (15) in Appendix C.

[22]The median bracket width is 500 euros a month, so $\sigma_\eta = 500/20/\sqrt{12}$ assuming 20 workdays in a month, since the standard deviation of a Uniform(0,A) distribution is $A/\sqrt{12}$.

requirements, we use distributions of the form

$$\begin{bmatrix} \log(M_j) \\ \text{logit}^{-1}(\alpha_{j,m}) \\ \text{logit}^{-1}(\beta_{j,m}) \\ \log(w_{j,m}) \\ \text{logit}^{-1}(\alpha_{j,f}) \\ \text{logit}^{-1}(\beta_{j,f}) \\ \log(w_{j,f}) \end{bmatrix} \sim \text{N}\Big(\mu_\omega + B_m X_{m,j} + B_f X_{f,j} + B_c X_{c,j},\, L_\omega \text{Diag}(\sigma_\omega)^2 L'_\omega\Big), \tag{12}$$

IID$\forall j$, where X_m, X_f, and X_c are matrices that contains individual-specific and couple-specific covariates (such as gender and age, discussed in detail in Section 6.1) for members of the couple, augmented by a constant to capture the level, and Bs is coefficient matrices. The variance is parametrized with the standard deviations σ_ω and the Cholesky decomposition of the correlation matrix L_ω.

As is common in multilevel models, (12) sharpens our estimates to the extent that it happens to fit the data, but does not constrain it. For example, if it is not a good fit, that is reflected in a large standard deviation σ_ω and couple-level observations are affect only to a small extent by (12). Also, the actual estimates may end up being non-normal, the multivariate normal is just a common starting point that can be adjusted if the model provides a poor fit to the data.[23]

This transformed distribution family is flexible, yet at the same time simple to parameterize and has parameters which are easy to interpret intuitively. For example, if $L_\omega L'_\omega$ is close to being diagonal, then there would be no correlation between the model parameters and wages, while a block-diagonal structure would demonstrate correlation for individuals (e.g. between α_i and w_i), but no correlation between spouses. Deviations from this allow us to model assortative matching between couples. It is important to emphasize that (12) is IID *ex ante*, but conditional on the actual realizations of hours, individuals and couples will of course be different *ex post* — for example, a couple where both members are working will probably have higher wages or α's compared to a couple where both members are non-employed. This is especially important for wages, which we observe directly only for the employed individuals. When analyzing the results, we are careful about distinguishing *ex ante* wages, which are realizations from the distribution (12) and may or may not be observable, and *observed* wages, which are wages for the employed individuals.

5.4 Priors

Following standard Bayesian practice, we use weakly informative prior distributions,[24] which we describe briefly. We choose an IID N(0,2.5) prior for the elements of the Bs, μ_ω. For L_ω, we use the a prior defined in Lewandowski, Kurowicka, and Joe (2009), equivalent to

$$p(\Omega \,|\, s_\Omega) \propto \det(\Omega)^{s_\Omega - 1}$$

where $\Omega = L_\omega L'_\omega$, with $s_\Omega = 2.0$, which ensures a vague but unimodal prior. For the elements of σ_ω, and for σ_ζ, σ_ε, we follow Polson, Scott, et al. (2012) and use the half-Normal prior with

[23]We check model predictions against the data in Section 6.2.

[24]See Gelman (2004).

location zero and scale 2.5, which is also vague but sufficient to make the posterior proper. For κ, we use the uniform prior on [0,2].

5.5 Some details of the likelihood calculation

The sections above completely specify the likelihood and prior as a function of

1. couple-level parameters $M_j, \alpha_{j,m}, \alpha_{j,f}, \beta_{j,m}, \beta_{j,f}, (w_{j,m}), (w_{j,f}), \zeta_{j,m}, \zeta_{j,f}$ [with the understanding that for employed persons we know the wages, hence the ()s],
2. cross-sectional parameters μ_ω, σ_ω, L_ω, B_m, B_f, B_c,
3. noise parameters κ, σ_ζ, and σ_ε.

We discuss some technical details of the estimation below.[25]

The equilibrium that we have discussed in Section 4 provides a mapping from the other income M, wages w_i, preference parameters α_i, β_i for $i = m, f$ to choices of market, leisure, and home production hours given ζ_m, ζ_f:

$$(M, w_m, w_f, \alpha_m, \alpha_f, \beta_m, \beta_f) \mapsto (n_m, n_f, \ell_m, \ell_f, h_m, h_f) \tag{13}$$

We marginalize out ζs in the estimation for more efficient for traversal of the posterior parameter space.[26] The ζs are then redrawn in a manner consistent with thresholds and σ_ζ for the posterior estimates that have been obtained.

From the αs, βs, wages, and the couple status, we calculate *expected* hours using the mapping in (13). We assume that these are observed with a noise as described in (10) and (9).

When calculating hours, we also obtain the *thresholds* for $\zeta_{j,m}$ and $\zeta_{j,f}$ that are consistent with a given employment status that we observe in the data. At these thresholds $\bar{\zeta}_{j,m}$ and $\bar{\zeta}_{j,f}$ individuals are just indifferent between employment and non-employment.

For employed individuals, we know that $\zeta_{j,i} < \bar{\zeta}_{j,i}$, and thus add

$$\log\left(\int_{-\infty}^{\bar{\zeta}_{j,i}} p(\zeta \mid \sigma_\zeta)\right)$$

to the log likelihood, and similarly the integral of the right tail for the non-employed.

5.6 Estimation details and diagnostics

We use a variant of the NUTS sampler Hoffman and Gelman (2014), as described by Betancourt (2017), implemented in Julia.[27] Whenever log Jacobian determinants are necessary because of transformations, we account for them in the code but do not list them specifically in the main text as they are a technical detail. The estimation methodology is tested on simulated data; we verify that it recovers parameters that were used in the simulation. Bayesian convergence diagnostics are available in Appendix D. Here, we also show plots of prior and posterior distributions. It is fair to describe our priors as very uninformative (loose).

[25]We do not include a formula for the posterior density as it is unwieldy and does not add anything but requires even more notation.

[26]For readers unfamiliar with Bayesian terminology: marginalization is a technique of sampling from marginal distributions $p_Y(y) = \int p_{XY}(x,y)\mathrm{d}x$ when the integral can be calculated or approximated analytically.

[27]See https://github.com/tpapp/DynamicHMC.jl.

6 Results

We discuss estimated parameters in subsection 6.1, then check predicted hours against the data in Section 6.2 where we also discuss the importance of modeling heterogeneous preferences. We illustrate identification in subsection 6.3. In Section 6.4, we discuss the implied elasticities.

6.1 Estimated parameters and distributions

The posterior estimates of the parameters α and β and also of the wage rates in the B matrix are depicted in Figure 2. Each panel reports how each of these parameters varies with selected demographic variables. Blue dots correspond to the point estimates for males, and red dots to the ones for females. The horizontal bars represent posterior quantiles with the thick bars corresponding to the 25%–75% quantile and the thin ones to the 5%–95% quantile. Each panel depicts for males and females the direction and extent by which any of the explanatory variables affects the particular parameter considered. For example, the upper left panel indicates that α does not react to age, but tends to decline for men and women in all explanatory variables except for females who are married or live in the west, and females who hold a university degree. Remarkably, these estimates are by and large mirrored in the lower panel which reports the reaction of male and female wage rates to the explanatory variables. Wages do not react to age, but they rise for men and women in all other variables except for when women are married, live in the west, or have children who are at least 6 years old. To check the plausibility of these results, we use OLS and estimate Mincer-type wage regressions for men and women with the same explanatory variables as in the Bayesian estimation. The point estimates are represented by a small cross. When contrasting the wage estimates from these two procedures, we conclude that the posterior wage estimates are consistent with the OLS estimates and plausible.[28]

6.2 Hours predictions and the importance of modeling preferences

Throughout this paper we have argued that couples matter for replicating total hours worked. Moreover, within couples of the same labor market status, spousal time allocated across market work, home work, and leisure varies widely (see Figure 1). We have also argued that preference heterogeneity at the spousal level — in addition to wage heterogeneity — is essential for replicating this observed heterogeneity in the data. Below we corroborate this view by invoking some plausibility checks based on parameter estimates followed by a more rigorous illustration of the fact that wage heterogeneity alone cannot replicate the observed heterogeneity in couples' time-use data.

For the sake of illustration, we focus on the observed heterogeneity of market hours worked by couples' labor market state which is determined primarily by the heterogeneity in wages and in α's. Figure 3 consists of three panels each of which depicts joint parameter densities for actual wage rates for employed or potential wage rates for non-employed and α's by couple type. The top panel depicts joint densities for w_m and w_f, and the panels at the bottom separately illustrate

[28] Posterior quantiles for noise parameters are in Table 10 of Appendix E.

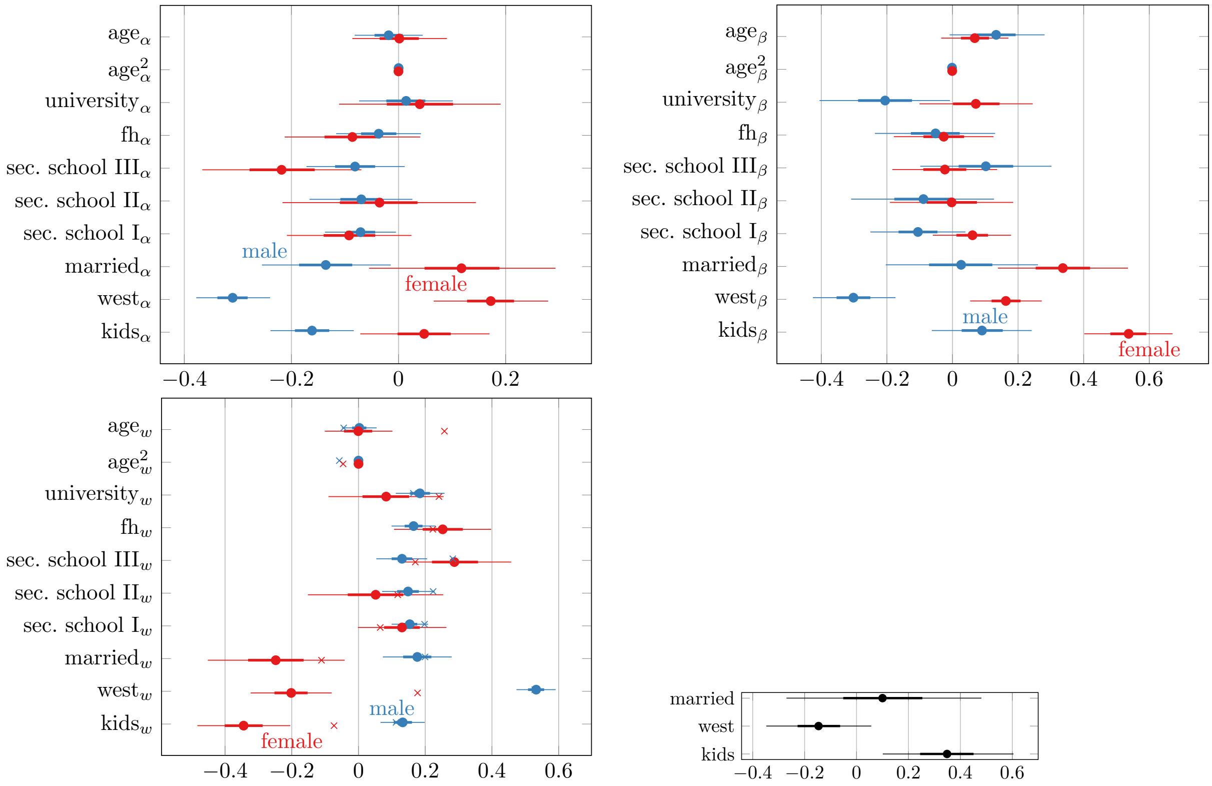

Figure 2: Posterior estimates for the B matrix in (12), ▬ male, ▬ female. Top left: coefficients α, top right: coefficients for β, bottom left: coefficients wages (× show OLS estimates on the subset of employed, by gender), bottom right: coefficients M (couple-level).

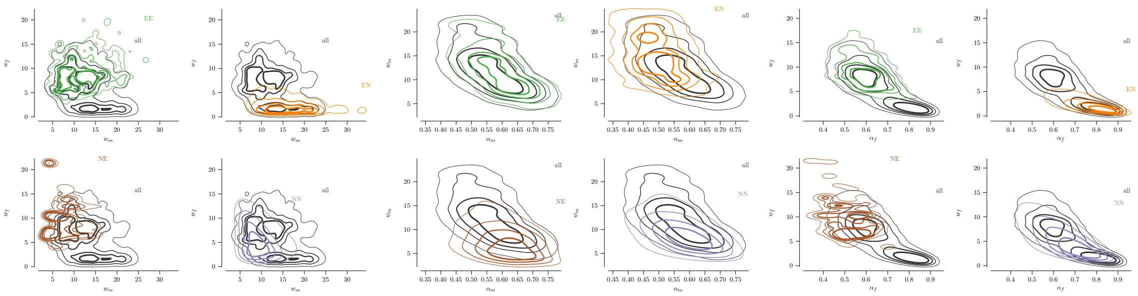

Figure 3: Selected cross-sectional marginal posterior densities for w_m,w_f,α_m,α_f by couple type. Overall cross-section in grey.

the joint densities between the wage rate and the corresponding α by gender and couple type. To make our point, we first consider the joint wage densities for single-earner couples, i.e., EN and NE. Although many females in traditional couples have a positive potential wage, they do not work in the market. These facts are compatible, because w_f and α_f are negatively correlated. That is, those females don't work in the market, because they assign a rather low value to market consumption. A similar argument holds for males in NE couples. Note that for the working spouses, such a negative correlation is not observed. What about dual-career couples? The wage rate and α are only weakly negatively correlated for females, but strongly negatively correlated for men. We take this to indicate that what keeps these males with a rather

	w_m	n_m	ℓ_m	h_m	ℓ_f	h_f
day 1	10.18	11.0	4.67	0.33	7.83	8.17
day 2		11.17	4.5	0.33	8.33	7.67

Table 2: Average daily time-use for a selected EN couple used to illustrate identification.

high wage from working even more hours is their rather low value of market consumption.

Figure 4 depicts for males and females, respectively, by their labor market status the weekly market hours from the data on the horizontal axis and the predicted posterior average on the vertical axis. The top panels correspond to our benchmark model with preference heterogeneity, and the bottom panels correspond to the model with homogeneous preferences. Visual inspection of these graphs suggests that our benchmark model with preference heterogeneity reproduces the patterns in the data for market hours compared to assuming homogeneous preferences. Specifically, without assuming heterogeneity in preferences, wage heterogeneity alone is unable to replicate any connection between the model and the data within EE couples, and predicts mostly high (around 10 hours/day) market work hours for males in EN couples. In contrast, for the model with wage- and preference heterogeneity, predicted vs observed hours are mostly around the diagonal, suggesting a reasonably good model fit.

6.3 Illustration of identification

In this section we illustrate how identification works in our multilevel model. Recall that for each couple in the data, we model couple-specific parameters *as if* they were drawn from a common distribution (12). However, once *couple-specific* observations are considered, they of course sharpen the estimates of parameters that characterize each particular couple.

While the cross-sectional distribution is estimated jointly with couple-specific parameters, for the purposes of illustration, *in this section we pretend that it was estimated using observations from all couples except a specific one*, and then build up the posterior step-by-step for a specific EN couple to illustrate identification.[29] The data for the couple we consider is shown in Table 2. Note that for this particular EN couple, we have two weekdays of observations, the variation of which is fairly typical for our data.

As is standard in multilevel models, the distribution parameters α_i,β_i for each individual, and for the non-employed also w_i are *jointly* identified conditional on the data and the hyperparameters. That is to say, conditional on noise magnitudes σ_ε, σ_η, hyperparameters σ_ω, L_ω, and Bs, the model assigns a posterior to the couple-level parameters. For the purposes of illustration and to avoid visual clutter, we fix all parameters at their posterior means *except* for the male preference parameter α_m and the female wage w_f, and plot components of the posterior in Figure 5, which we discuss below.

[29]We chose an EN couple because it demonstrates identification for both an employed and a non-employed member, EE and NN couples work analogously, *mutatis mutandis*. All other parts of the model are standard, with straightforward identification. See Appendix D for posterior-prior comparisons of common parameters.

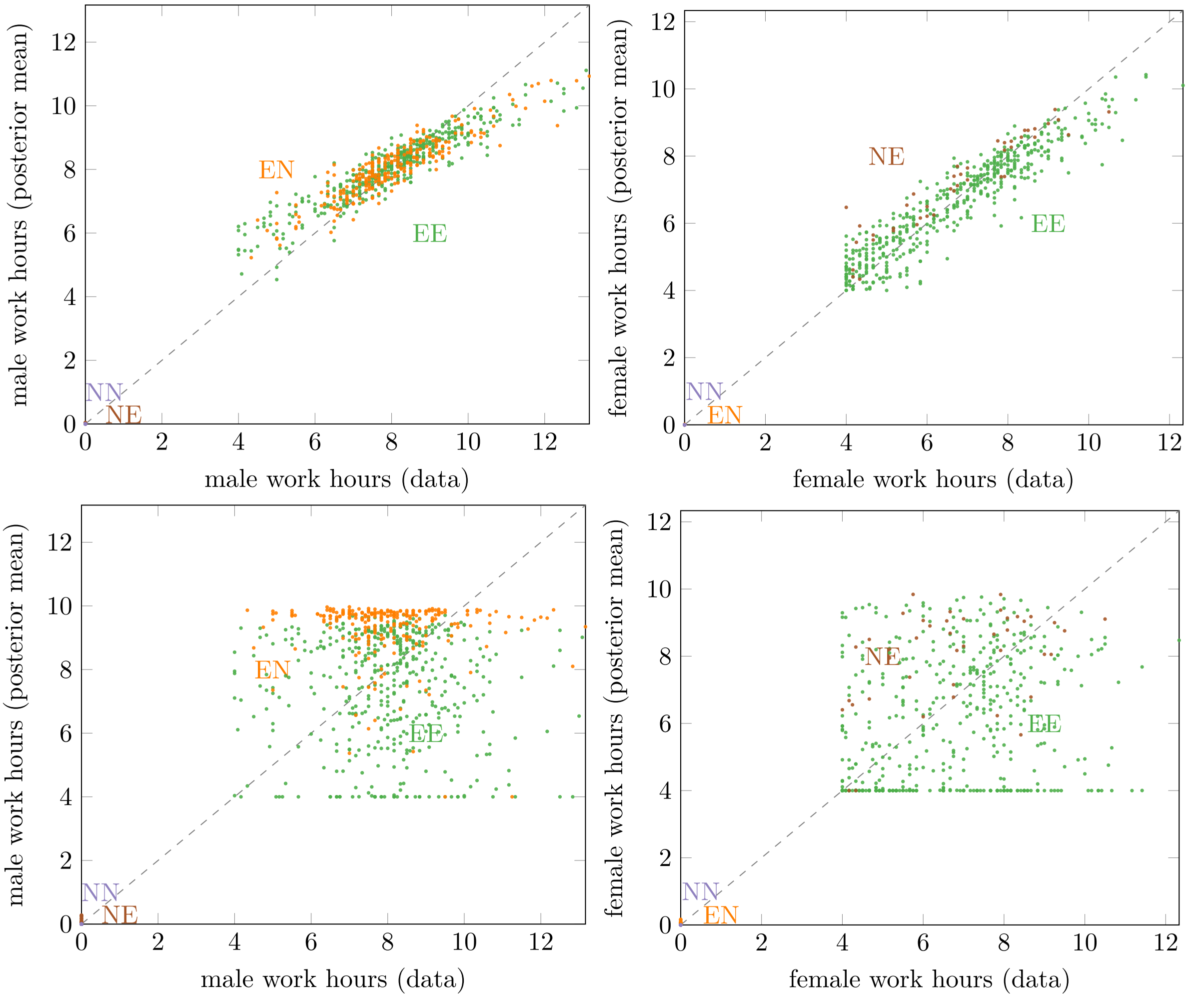

Figure 4: Market hours in the data and predicted from the model (posterior means), for males (left) and females (right). Top: model with heterogeneous preferences, bottom: model with homogeneous preferences. The diagonal can be used to assess model fit, observe how predicted vs data hours are clustered around the diagonal for the model with wage- and preference heterogeneity (top panels), while for the model without preference heterogeneity the fit is poor (bottom panels).

The top panel of Figure 5 shows a slice of (12) along $\hat{\alpha}_m$ and $\hat{w}_f$, as determined by the covariates for this specific couple. Without considering the particular time-use observations for this couple, we start with a strong correlation between $\hat{\alpha}_m = \text{logit}(\alpha_m)$ and $\hat{w}_f = \log(w_f)$.[30] That is to say, without knowing anything about a particular couple, but having seen all other couples, the patterns in our data (that is to say, for the majority of couples where both members are employed) are compatible with male preferences for work counterbalancing female wages. If either of these were off, only one member of the couple would work. The posterior density is concentrated around $\alpha_m \approx \text{logit}^{-1}(0.4) \approx 0.6$ and hourly $w_f \approx \exp(1.0) \approx 2.7$.

Second, in the middle panel of Figure 5 shows the likelihood that we would obtain from nothing but the time use observations of this couple, with noise parameters at their posterior

[30] We are plotting transformed variables because they reveal the highest density region shapes better.

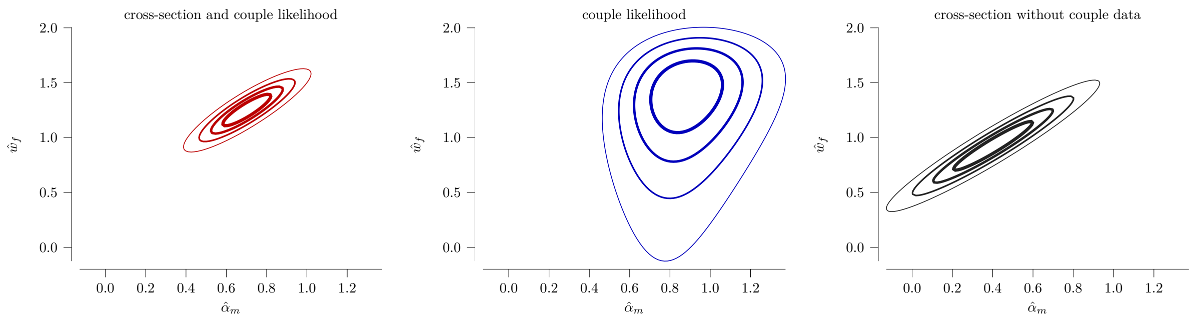

Figure 5: Illustration of identification in a multilevel model. All plots show 20%, 40%, . . . , 80% highest posterior density regions for $(\hat{\alpha}_m, \hat{w}_f) = (\text{logit}(\alpha_m), \log(w_f))$ for a selected EN couple, with all other parameters fixed at their posterior means. Left: cross-sectional hyperprior, without couple-specific data. Middle: likelihood of the observations for the selected couple, based solely on the data. Right: posterior from the two other plots combined.

means. *For this particular couple*, we know that the above-average male hours suggest a strong preference for consumption from the male member (high α_m), and the fact that the female is unemployed is only compatible with a low female wage. Thus, in the middle panel (blue), we see density regions that that are concentrated around $\alpha_m = \text{logit}^{-1}(0.8) \approx 0.7$, while at the same time allowing a wide range of wages around $w_f \approx \exp(1.4) \approx 4$. The couple-level uncertainty is high for wages since for the non-employed member, we do not have wage information, and the wage is only identified from the structural implications of observing the outcome EN. Note that there is considerable uncertainty about all parameters, which is a natural consequence of having noise terms in the hours observation and the employment margin, which is in turn required for modeling the actual variation in day-to-day allocations.

Finally, when we *combine* the hyperprior and the couple-level data, we get the posterior density regions in the bottom panel of Figure 5 (red). Note how this much sharper than either of the panels above: multilevel models combine information from couple-specific and cross-sectional patterns. For this specific couple, the information pooled from other couples can be considered a prior, which is then combined with the data from actual time use observations. The significant influence of the hyperparameters here reflects the fact that the regression in (12) turns out to be a good predictor of couples' wages and parameters, so in the combination of the hyperparameters and the couple-specific observation the former have a large weight.[31] Compared to similar couples in the whole sample, the observations of a particular couple may then be considered large or small, which explains this deviation. These deviations of course balance on average for the whole sample.

[31] If we had a simple linear model with normal errors and no correlations for parameters, posterior modes for the couple-specific parameters would be weighted averages of the mean "deterministic" values and the linear prediction from (12), with the weights proportional to diagonal of Σ^{-1} and σ_ε^{-2} or σ_η^{-2}, as described in Gelman, Carlin, et al. (2013, p 116). Our setup is nonlinear and more complicated, but has the same intuition.

6.4 Cross-sectional hours changes and elasticities

In this section we use the estimated model as a laboratory to perform counterfactual experiments. Specifically, for each couple j we take a posterior draw k of parameters $M_{j,k}$, $\alpha_{m;j,k}$, $\alpha_{f;j,k}$, $\beta_{m;j,k}$, $\beta_{f;j,k}$, $w_{m;j,k}$, $w_{f;j,k}$, $\zeta_{m;j,k}$, $\zeta_{f;j,k}$, and calculate the corresponding market hours $n_{m;j,k}$ and $n_{f;j,k}$. Then we perform the following experiments:

1. increase male wage by 10%,
2. increase female wage by 10%,
3. increase female wage by 10%, but do not allow the male to change market hours.

Using the parameters from the same posterior draw, especially the ζs, ensures that the calculation is consistent. As our wage data are net of taxes, one should interpret these elasticities as net wage elasticities. This means that they do not take into account how specifics of the tax system affect net wages and, hence, labor supply decisions.

For some of the plots below, we bin couples by some particular parameter, such as hour- or wage deciles of some member. Hours changes and elasticities are then calculated using the within the bins, as explained in Appendix F. For each bin, we calculate the 5%, 25%, 50%, 75% and 95% quantiles, and display them with vertical lines and a dot for the median. We also calculate the means, and connect them as a line plot. These plots thus display information about the dispersion of hours changes and elasticities, and the shape of the distributions, in particular skewness. We also report results for subgroups, e.g., for the couple's employment status before the counterfactual experiment.

It is important to emphasize that for each member in each couple, we get a *posterior sample* of hours responses and elasticities. For example, consider an NE couple where the male is non-employed, while the female is employed. For some combinations of posterior parameters (e.g. male wage close to the reservation wage, which depends on the αs and βs, and the female wage), a 10% increase in male wages can result in employment (EE), or even in the female withdrawing from employment (EN), while for wages further from the reservation wage, the male would remain non-employed regardless of the wage increase (NE). Since we have draws from the posterior distribution, the adjustments in hours are automatically weighted with probabilities of all of these events in hour calculations, and we report expected values that take the nonlinear hours responses and the posterior uncertainty into account.

6.4.1 Cross-sectional hours changes with heterogeneous preferences

We first consider males' hours responses to a rise in their own wage and how these responses vary with males' hours worked.[32] According to Figure 6, the reaction of mean hours worked of men in dual career couples decline from half an hour to ca. 20 minutes as daily hours worked increase from slightly less than 6 to around 10. For men in traditional couples, mean hours changes are negligible, but slightly rising in hours worked. Those adjustment patterns for

[32] Technically, the plots include extensive margin changes, i.e., members switching between employment and non-employment, but these are so small that omitting them would be visually indistinguishable.

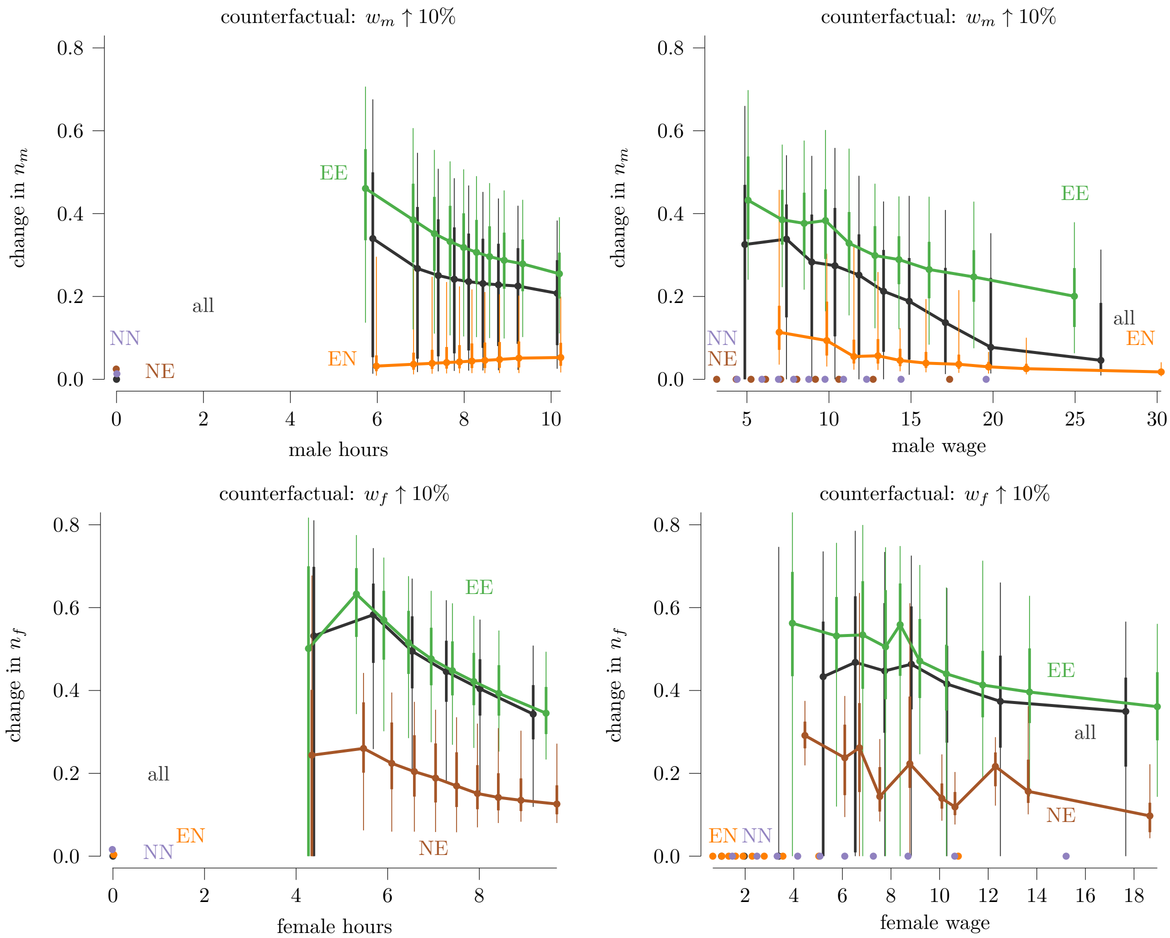

Figure 6: Male and female hours response to own wage change. Top row: male hours response to own wage change versus male hours (left) and male wages (right). Bottom row: female hours change to own wage change vs female hours (left) and female wages (right).

men in dual-career couples are very similar, if we plot their hours' reactions to a wage rise against their own wage. The mean hours reactions of men in EN couples decline from seven minutes at very low wages to a negligible amount at higher wages. Remarkably, the average of the mean hours reaction across all men regardless of their couple type strongly declines over the range of possible male wages. It is closer to the line for men in dual-career couples at lower wages, but quickly approaches that for men in traditional couples at higher hourly wage rates.

Females' mean reactions of hours worked when their own wages rise tend to exceed that of males in comparable couple types. When plotted against their own hours worked, we detect a hump-shaped pattern with an initial rise in additional hours followed by a steady decline. The mean reactions range between two-thirds of an hour and 25 minutes for females in dual-career couples, and between a quarter of an hour and ca. five minutes for females who are the sole breadwinners. When plotted against females' own wages, the mean hours reaction decreases for women in all couple types, and they are significantly lower for female single earners than for double earners at all wage levels.

Summing up, the mean reaction of market hours worked when one's own wage rises is larger for females than males in comparable couple types, it tends to decline in the own wage level, and it is lower for single earners than double earners.

6.4.2 Cross-sectional hours changes with spousal interaction

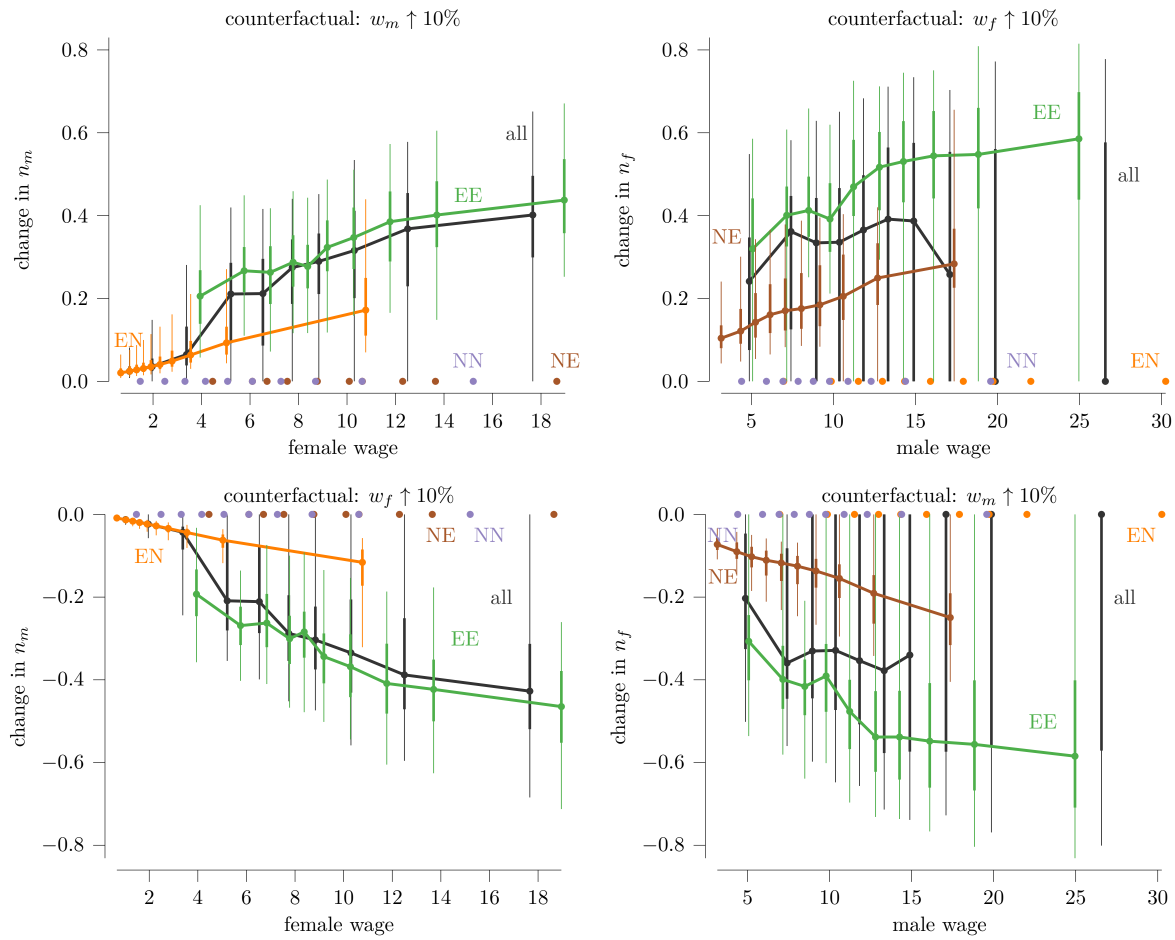

Figure 7: Male and female hours response to own and cross wage change. Top row: male and female hours response to own wage change vs spouse's wage. Middle row: male and female hours response to spouse's wage change vs spouse's wage.

We deviate from much of the literature on family labor supply by treating spouses as equals in terms of time-allocation decisions. In particular, we allow not only females, but also males to adjust their respective time-allocation in reaction to their spouse receiving a wage shock. Males' time-use reactions to wage changes are of interest in and of themselves, especially vis-a-vis their female partners. Cross-wage changes in market hours worked are one way to measure the extent to which spouses react to each other's wage changes. Another way is to depict males and females' own-wage hours changes against spouses' wage rates or hours worked and ask whether or not those changes are flat in these variables. Figure 7 illustrates our findings.

When plotted against females' wage rates, males' mean own-wage hours reactions almost steadily increase. That is to say males with a high-wage partner tend to have a larger own-wage hours change even if their partner is not employed. Similarly, when plotted against males' wage rates, females' own-wage hours reactions rise in their partner's wage rate. For females who are sole earners, those changes in hours worked are largest if their partner is assigned the highest possible wage.

Next, we explore spouses' mean cross-wage hours changes and how they vary with their partner's wage rate. If males experience a positive wage shock which induces them to work more, their female partners reduce their market hours with the relative reduction being strongest among females with an active high-wage partner. Similarly, if females receive a wage rise causing them to work more, their male partners work less, and this relative hours' reduction is strongest for men with a working spouse who has a high wage rate. These patterns clearly show that spouses react to each other's wage change, and that this reaction is relatively strong if a given wage change induces the affected partner to adjust her hours worked. This holds for females as well as for males.

Again we observe that the mean cross-wage hours reduction is larger in absolute values for females than males in similar couple types, and it is lower in absolute values for single earners than double earners. Figure 7 conveys another important message regarding relative wage levels by couple type. In traditional couples, female wage rates are significantly lower than in dual-career or progressive couples. On the other hand the highest male wage rates are earned by men in traditional couples, followed by those in dual career couples.

6.4.3 Own- and cross-wage elasticities in the cross-section and in the aggregate

Table 3 reports own- and cross-wage elasticities for males and females along the intensive margin and also along the extensive margin that are implied by our full-heterogeneity model. These elasticities are reported in total, and also by wage quartile and by couples' labor market status. We observe much symmetry in males' and females' adjustment along the intensive margin, but stark differences in their adjustment along the extensive margin.

Females' own- and cross-wage elasticities along the intensive margin exceed those of males in absolute terms. For each gender, these elasticities decrease in absolute value in the quartile of the own-wage. Regarding the adjustment along the extensive margin, females react slightly more strongly to an own-wage rise than males, but they reduce employment to a lesser degree when male wages rise than males do in reaction to rising female wages. There is another remarkable asymmetry. In reaction to rising male wages, females' withdrawal from employment rises in their own wage, but when females' wages rise, males' withdrawal from employment declines in their own wage. Lastly, when looking at the extensive margin adjustment from the perspective of couple types, we find that males in dual-career couples react much more strongly to a rise in their partner's wage than corresponding females. Similarly, the reaction of males in progressive couples to a rise in male wages is five times as big as that of females in traditional couples when female wages rise. When combining adjustment along both margins to total wage-elasticities, we find that the size and patterns of the intensive margin elasticities dominate the total.

	increase male wages						increase female wages					
group	extensive (%)		intensive		total		extensive (%)		intensive		total	
	m	f	m	f	m	f	m	f	m	f	m	f
all	0.24	−0.75	0.29	−0.63	0.30	−0.64	−1.46	0.27	−0.29	0.66	−0.30	0.66
			0.29,0.30	−0.64,−0.61	0.29,0.31	−0.67,−0.61			−0.29,−0.28	0.64,0.68	−0.33,−0.29	0.64,0.69
w Q1	0.49	−0.03	0.44	−0.71	0.44	−0.72	−1.55	0.20	−0.42	0.76	−0.44	0.82
			0.42,0.45	−0.82,−0.60	0.42,0.46	−0.83,−0.60			−0.43,−0.41	0.64,0.90	−0.49,−0.41	0.64,0.94
w Q2	0.27	−0.62	0.34	−0.69	0.35	−0.70	−1.49	0.36	−0.33	0.76	−0.35	0.77
			0.33,0.35	−0.72,−0.66	0.33,0.36	−0.75,−0.66			−0.34,−0.32	0.72,0.80	−0.39,−0.32	0.72,0.81
w Q3	0.15	−1.02	0.25	−0.67	0.25	−0.68	−1.57	0.28	−0.26	0.69	−0.27	0.70
			0.24,0.26	−0.69,−0.64	0.24,0.26	−0.73,−0.65			−0.26,−0.25	0.67,0.72	−0.31,−0.25	0.67,0.73
w Q4	0.06	−1.31	0.17	−0.54	0.17	−0.56	−1.22	0.25	−0.17	0.55	−0.18	0.55
			0.16,0.17	−0.56,−0.52	0.16,0.17	−0.60,−0.53			−0.18,−0.16	0.53,0.57	−0.22,−0.17	0.53,0.58
EE	0.00	−1.19	0.40	−0.67	0.40	−0.68	−2.21	0.00	−0.41	0.70	−0.43	0.70
			0.39,0.41	−0.68,−0.65	0.39,0.41	−0.71,−0.65			−0.42,−0.40	0.68,0.71	−0.47,−0.41	0.68,0.71
EN	0.00	0.00	0.09	—	0.09	—	−0.48	0.65	−0.05	—	−0.06	—
			0.08,0.09		0.08,0.09				−0.06,−0.05		−0.09,−0.05	
NE	3.25	−0.77	—	−0.20	—	−0.21	0.00	0.00	—	0.28	—	0.28
				−0.23,−0.18		−0.24,−0.18				0.25,0.30		0.25,0.30
NN	1.55	0.00	—	—	—	—	0.00	1.55	—	—	—	—

Table 3: Aggregate responses (as elasticities) to increasing male and female wages. Smaller numbers below cells are 10%–90% quantiles (omitted for extensive margin because they are hard to interpret for discrete changes). 2002 data, γ=0.5.

6.4.4 Own-wage elasticities in the cross-section for alternative model specifications

In this section we compare our model predictions with results from the literature and highlight the quantitative implications of preference heterogeneity at the spousal level as well as spousal interaction. To do so, we report own-wage elasticities for males and females who are originally employed and remain employed following a counterfactual wage rise. Hence, the resulting elasticities refer to adjustments along the intensive margin only. In Figure 8 we depict own-wage elasticities for all men and women by their wage respective hours worked decile from our full-heterogeneity benchmark model, and also from the two restrained model versions: first we assign gender-specific homogeneous preferences ("homog. pref.") and then we add constant males' income ("male fixed"). No clear ranking of these elasticities across the alternative model specifications emerges, but some facts stand out.

The benchmark model with spousal preference heterogeneity ("full") generates own-wage elasticities that tend to steadily decline in spouses own wage and own hours worked except for females with low hours. At the median wage, the own-wage elasticities for each gender are very similiar across the model with preference heterogeneity and the one with gender-specific homogeneous preferences. For males, the values equal 0.29 and 0.25, respectively, and for females, the values are 0.78 compared to 0.72.[33] However, much larger discrepancies between these two models' predictions emerge towards the margins of the respective wage distributions. That's because the homogeneous preference model assigns the average preference parameters that strongly differ from the one of the heterogeneous preference model. To illustrate this, consider working males. We know from Figure 3 that their wages tend to be negatively correlated with α. The model with homogeneous preferences assigns high-wage males a rather low α and low-wage males too high an α, thereby generating too high wage-elasticities towards the bottom and too low elasticities towards the top of males' wage distribution. The story for females is similar for high-wage earners who are predominantly spouses in dual-earner couples. Low-wage females often are the sole earner, and their individual α tends to exceed the average value that the homogeneous preference model assigns. Hence, this model predicts lower own-wage elasticities for females towards the bottom of their wage distribution than does our "full" heterogeneity model. Lastly, when we add to homogeneous preferences fixed male income, we observe a further decline in females' own-wage elasticities across the entire wage distribution. This is because if males cannot adjust their market hours in reaction to a rise in their spouse's wage rate, females have less of an incentive to increase their own market hours. At the median female wage, an elasticity results that is consistent with what Attanasio et al. (2018) report in their Table XI, i.e., 0.48.

In sum, we interpret the evidence from Figure 8 as support for our model, because it not only helps to replicate the observed heterogeneity across couples' hours worked, its restrained versions imply values for gender-specific wage-elasticities that are by and large consistent with what we know from the literature.

[33]This latter value exceeds 0.40 – the corresponding elasticity that Blundell, Pistaferri, and Saporta-Eksten (2016) report in their Table 5. When comparing these values, one needs to remember that their sample differs from ours not only in the country and time period covered, but their couples are stably married and include males who are continuously employed. Moreover, they use identical preferences for everyone whereas we can only use gender-specific homogeneous preferences.

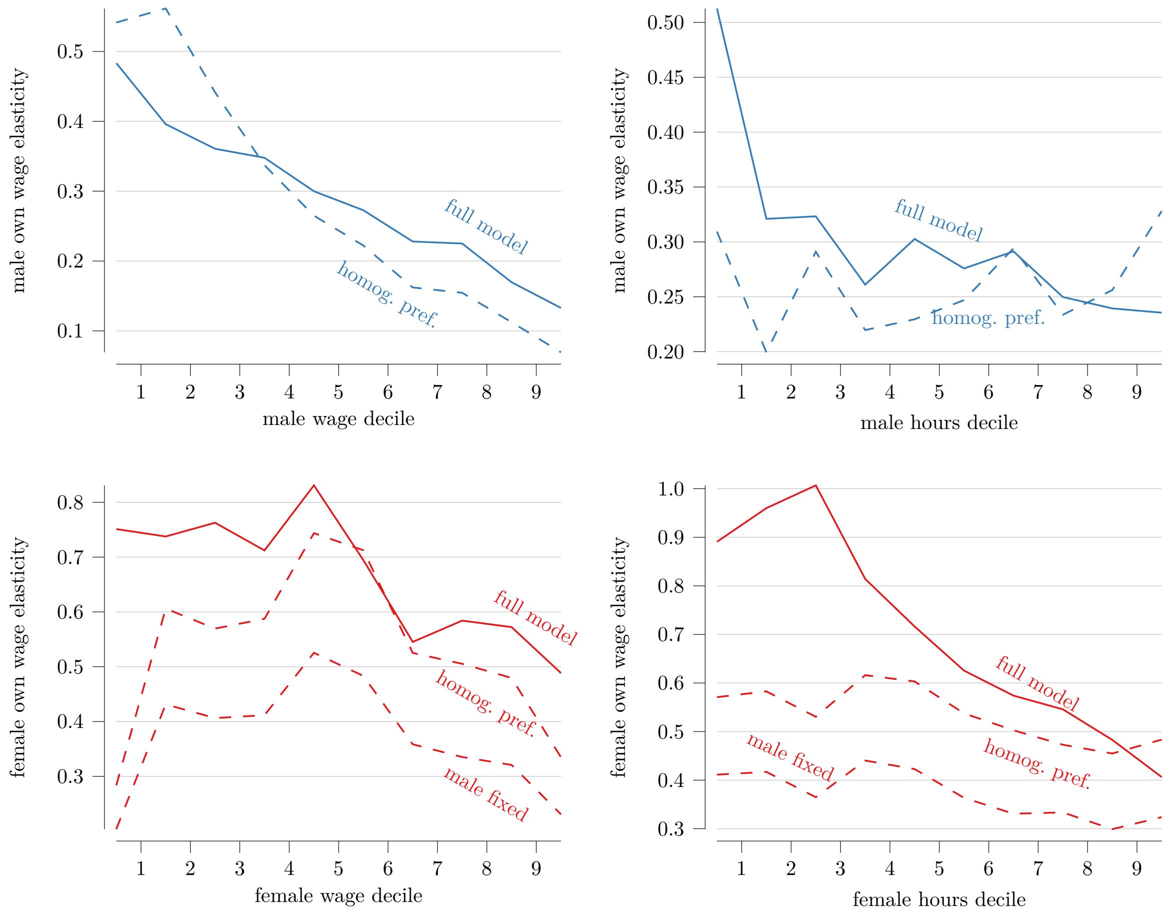

Figure 8: Elasticities by wage and hours deciles.

6.5 Robustness checks

In our baseline analysis, we have fixed γ at the value of 0.5. As we explain in Section 5.3, γ cannot be identified in our setup from allocation information alone. We do not pursue identification of γ in the context of this model, since we suspect an estimated value would be driven by incidental assumptions such as the tail shape of distributions in (12) which we parameterized with a multivariate normal for convenience. Instead, we consider the results to two alternative cases in which we decrease the male share in home production, γ, to 0.3 and increase it to 0.7, respectively.

Table 4, which should be compared to Table 3, shows the aggregate responses for $\gamma = 0.3$, $\gamma = 0.7$ using the 2002 wave of the data, and $\gamma = 0.5$ using the 2012 wave. According to the top and the middle panel of Table 4, varying γ does not noticeably change our main results. However, when comparing the bottom panel to Table 3, we observe a noticeable rise in the absolute values of almost all extensive margin elasticities. This holds for males and females in reaction to either wage rise. In addition, males' cross-wage elasticities along the intensive margin rise in absolute terms indicating that working males reduce their market hours relatively more strongly in reaction to their spouses' wage rise in 2012 compared to 2002. When comparing results across these two waves, one needs to keep in mind that during this decade, female labor force

2002 data, $\gamma = 0.3$

	increase male wages						increase female wages					
group	extensive (%)		intensive		total		extensive (%)		intensive		total	
	m	f	m	f	m	f	m	f	m	f	m	f
all	0.24	−0.74	0.29	−0.63	0.29	−0.64	−1.46	0.28	−0.29	0.66	−0.30	0.66
w Q1	0.51	−0.03	0.44	−0.72	0.44	−0.73	−1.56	0.20	−0.42	0.77	−0.44	0.84
w Q2	0.26	−0.64	0.34	−0.69	0.34	−0.70	−1.51	0.37	−0.33	0.76	−0.34	0.77
w Q3	0.15	−1.03	0.25	−0.67	0.25	−0.68	−1.58	0.29	−0.25	0.70	−0.27	0.70
w Q4	0.05	−1.25	0.17	−0.54	0.17	−0.55	−1.18	0.25	−0.17	0.55	−0.18	0.55
EE	0.00	−1.17	0.40	−0.67	0.40	−0.68	−2.21	0.00	−0.41	0.70	−0.43	0.70
EN	0.00	0.00	0.09	—	0.09	—	−0.47	0.65	−0.05	—	−0.06	—
NE	3.32	−0.87	—	−0.20	—	−0.21	0.00	0.00	—	0.28	—	0.28
NN	1.50	0.00	—	—	—	—	0.00	1.54	—	—	—	—

2002 data, $\gamma = 0.7$

	increase male wages						increase female wages					
group	extensive (%)		intensive		total		extensive (%)		intensive		total	
	m	f	m	f	m	f	m	f	m	f	m	f
all	0.25	−0.76	0.29	−0.57	0.29	−0.59	−1.39	0.27	−0.28	0.60	−0.30	0.60
w Q1	0.53	−0.03	0.43	−0.67	0.44	−0.68	−1.50	0.22	−0.42	0.71	−0.44	0.78
w Q2	0.27	−0.65	0.33	−0.62	0.33	−0.63	−1.46	0.36	−0.32	0.67	−0.33	0.68
w Q3	0.14	−1.10	0.24	−0.60	0.24	−0.61	−1.47	0.26	−0.24	0.62	−0.26	0.62
w Q4	0.06	−1.26	0.16	−0.52	0.17	−0.53	−1.14	0.24	−0.17	0.52	−0.18	0.53
EE	0.00	−1.21	0.39	−0.61	0.39	−0.62	−2.10	0.00	−0.40	0.63	−0.42	0.63
EN	0.00	0.00	0.09	—	0.09	—	−0.48	0.65	−0.05	—	−0.06	—
NE	3.40	−0.83	—	−0.20	—	−0.21	0.00	0.00	—	0.27	—	0.27
NN	1.49	0.00	—	—	—	—	0.00	1.44	—	—	—	—

2012 data ($\gamma = 0.5$)

	increase male wages						increase female wages					
group	extensive (%)		intensive		total		extensive (%)		intensive		total	
	m	f	m	f	m	f	m	f	m	f	m	f
all	0.30	−0.97	0.31	−0.65	0.32	−0.66	−2.00	0.41	−0.34	0.66	−0.36	0.66
w Q1	0.63	−0.40	0.43	−0.72	0.44	−0.74	−1.67	0.62	−0.45	0.71	−0.47	0.74
w Q2	0.31	−0.92	0.34	−0.74	0.35	−0.76	−2.24	0.44	−0.37	0.76	−0.39	0.77
w Q3	0.21	−1.26	0.29	−0.64	0.30	−0.66	−2.11	0.28	−0.32	0.66	−0.34	0.66
w Q4	0.06	−1.21	0.23	−0.56	0.23	−0.57	−1.96	0.34	−0.24	0.56	−0.26	0.56
EE	0.00	−1.34	0.39	−0.69	0.39	−0.70	−2.61	0.00	−0.42	0.70	−0.45	0.70
EN	0.00	0.00	0.09	—	0.09	—	−1.03	1.41	−0.09	—	−0.10	—
NE	4.03	−1.12	—	−0.21	—	−0.22	0.00	0.00	—	0.22	—	0.22
NN	1.71	0.00	—	—	—	—	0.00	2.08	—	—	—	—

Table 4: Robustness checks for aggregate responses. Compare to Table 3. Top and middle tables: results with $\gamma = 0.3$ and $\gamma = 0.7$. Bottom table: 2012 data wave, with $\gamma = 0.5$.

participation among the 25 to 54 years old rose from ca. 66 percent in 2002 to 74 percent in 2012 while that of males remained unchanged. This change is mirrored by the fact that by 2012, the share of dual-career couples had risen to 70 percent while that of traditional couples had declined to 21 percent with the remaining nine percent being shared equally among NE and NN couples.

7 Conclusion

We develop a model of time allocation making spouses within couples our unit of analysis. We model preference heterogeneity — in addition to wage heterogeneity — within and across couples and allow spouses to mutually insure against wage shocks by adjusting their time allocation. In this setting, all individuals endogenously sort into market work, or homework and leisure, yielding, as an equilibrium outcome, dual-career couples, those with only one spouse employed, and couples where neither partner works in the market. We estimate our model using Bayesian techniques and micro data from the 2001/02 wave of the German Time-Use Survey. Our sample contains actual couples — married or cohabiting — without young children where each spouse is of prime working-age.

The estimated model is consistent with our motivating observation that time-allocation is not only very diverse across couples of different labor market status, but also *within* couples of the same status. Wage heterogeneity alone cannot replicate this fact. We use the estimated model to study how males' and females' market hours change in reaction to a small change in one's own wage or the spouse's wage.

Our results clearly indicate that spousal interaction in time-use decisions matters and so does preference heterogeneity. This becomes apparent when looking at wage-elasticities of market hours from various perspectives. As for adjustments along the intensive margin, females' own- and cross-wage elasticities of hours worked on average are twice as large as comparable measures for males (0.66 vs. 0.29 in absolute values). They are largest in absolute value for dual-earner (EE) females and males (0.7 vs. 0.4), but drastically decrease to 0.28 for single-earner females (NE) and to a low 0.09 for single-earner males (EN). The adjustment patterns along the extensive margin are more diverse. Males are on average twice as likely as females to withdraw from market work when their spouses receive a positive wage shock (-1.46 compared to -.75), and this pattern prevails among dual-career couples (-2.21 vs. -1.19). However, for single earners, this cross-wage elasticity changes to -.48 for males and to -.77 for females. Females in (NE) couples are more likely to withdraw from market work than males in (EN) couples when their spouses receive a positive wage shock, because their male partners are much more inclined to start employment (3.25 vs. 0.68). Although all elasticities are measured for couples in Germany in 2001/02, our robustness checks show that the patterns of adjustment continue to hold a decade later, when female labor force participation had risen and so had the share of dual-career couples.

The quantitative importance of heterogeneous preferences for our results also becomes apparent when we replace them in our benchmark model by gender-specific homogeneous preferences. In that case own-wage elasticities along the intensive margin result that markedly differ from their counterparts in our benchmark model especially towards the margins of the wage distribution. E.g., for the lowest wage decile, the own-wage elasticity of females drops from 0.75 to 0.29, whereas for males it is eleven percentage points larger. When we also eliminate the mutual insurance option for males by holding their earnings constant when their spouses receive a wage shock, females' own-wage elasticities decline at all wage levels. At the median female wage this elasticity drops from 0.73 to 0.52. This result also underlines that allowing males to react to females' wage shocks matters quantitatively.

Our results show why modelling preference and wage heterogeneity at the spousal level matters for macroeconomics and especially for macroeconomic policy analysis that involves the labor market. Although mean or median wage-elasticities for men and women are similar and of plausible size across different model specifications, implying that the predicted average impact of a particular policy may be similar, the incidence of these policies obviously is very different for spouses depending on the couple's labor market status and on spouses' wage rates or hours worked. Our framework contains crucial ingredients for studying the implications of particular non-linear policies at the spousal level and their impact on labor supply. Furthermore, we believe that our rich analytical setting lends itself to studying important issues related to couples' time-allocation. For example, empirical evidence suggests that women in industrialized countries have steadily increased their educational achievements which has contributed to an improved assortative mating of partners by wages. This change most likely impacts spousal time-allocation, and also wage-elasticities in the cross-section.

References

Aguiar, Mark and Eric Hurst (2007). "Measuring Trends in Leisure: The Allocation of Time over Five Decades". In: *Quarterly Journal of Economics* 122.3, pp. 969–1006.

Apps, Patricia and Ray Rees (1988). "Taxation and the Household". In: *Journal of Public Economics* 35.3, pp. 355–369.

Attanasio, Orazio et al. (2018). "Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply". In: *Econometrica* 86.6, pp. 2049–2082.

Bernardo, José M (1996). "The concept of exchangeability and its applications". In: *Far East Journal of Mathematical Sciences* 4, pp. 111–122.

Betancourt, Michael (2017). "A Conceptual Introduction to Hamiltonian Monte Carlo". In: arXiv: 1701.02434v1 `[stat.ME]`.

Birinci, Serdar (2019). *Spousal labor supply response to job displacement and implications for optimal transfers*. Tech. rep. FRB St. Louis Working Paper 2019–020.

Blundell, Richard, Pierre-Andre Chiappori, and Costas Meghir (2005). "Collective Labor Supply with Children". In: *Journal of Political Economy* 113.6, pp. 1277–1306.

Blundell, Richard, Luigi Pistaferri, and Itay Saporta-Eksten (2016). "Consumption Inequality and Family Labor Supply". In: *American Economic Review* 106.2, pp. 387–435.

Browning, Martin and Pierre-Andre Chiappori (1998). "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests". In: *Econometrica* 66.6, pp. 1241–1278.

Chiappori, Pierre-André (1992). "Collective labor supply and welfare". In: *Journal of political Economy* 100.3, pp. 437–467.

Chiappori, Pierre-Andre (1988). "Rational Household Labor Supply". In: *Econometrica* 56.1, pp. 63–90.

Del Boca, Daniela and Christopher Flinn (1995). "Rationalizing Child Support Decisions". In: *American Economic Review* 85.5, pp. 1241–1263.

– (2012). "Endogenous Household Interaction". In: *Journal of Econometrics* 166.1, pp. 49–65.

Doepke, Matthias and Michèle Tertilt (2019). "Does female empowerment promote economic development?" In: *Journal of Economic Growth* 24.4, pp. 309–343.

Gelman, Andrew (2004). "Parameterization and Bayesian modeling". In: *Journal of the American Statistical Association* 99.466, pp. 537–545.

Gelman, Andrew, John B Carlin, et al. (2013). *Bayesian data analysis*. CRC Press.

Gelman, Andrew and Jennifer Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gobbi, Paula E (2018). "Childcare and commitment within households". In: *Journal of Economic Theory* 176, pp. 503–551.

Golosov, Mikhail et al. (2021). *How Americans Respond to Idiosyncratic and Exogenous Changes in Household Wealth and Unearned Income*. Tech. rep. University of Chicago.

Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin (2017). "Marriage, Labor Supply, and Home Production". In: *Econometrica* 85.6, pp. 1873–1919.

Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura (2012). "Taxation and Household Labor Supply". In: *Review of Economic Studies* 79.3, pp. 1113–1149.

Hoffman, Matthew D and Andrew Gelman (2014). "The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo." In: *Journal of Machine Learning Research* 15.1, pp. 1593–1623.

Hox, Joop J, Mirjam Moerbeek, and Rens Van de Schoot (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe (2009). "Generating random correlation matrices based on vines and extended onion method". In: *Journal of multivariate analysis* 100.9, pp. 1989–2001.

Lundberg, Shelly and Robert A Pollak (1994). "Noncooperative bargaining models of marriage". In: *The American Economic Review* 84.2, pp. 132–137.

Manser, Marilyn and Murray Brown (1980). "Marriage and Household Decision Making: A Bargaining Analysis". In: *International Economic Review* 21.1, pp. 31–40.

McElroy, Marjorie and Mary Horney (1981). "Nash-Bargained Household Decisions: Toward a Generalization of the Theory of Demand". In: *International Economic Review* 22.2, pp. 333–349.

Obermeier, Tim (2023). *Individual Welfare Analysis: A tale of consumption, time use and preference heterogeneity*. Tech. rep. Centre for Economic Performance Discussion Paper 1954.

Polson, Nicholas G, James G Scott, et al. (2012). "On the half-Cauchy prior for a global scale parameter". In: *Bayesian Analysis* 7.4, pp. 887–902.

Snijders, Tom AB and Roel J Bosker (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

Vehtari, Aki et al. (2021). "Rank-normalization, folding, and localization: An improved R for assessing convergence of MCMC". In: *Bayesian analysis* 1.1, pp. 1–28.

Wu, Chunzan and Dirk Krueger (2021). "Consumption insurance against wage risk: Family labor supply and optimal progressive income taxation". In: *American Economic Journal: Macroeconomics* 13.1, pp. 79–113.

Online appendices

A Additional graphs and tables for data

household type	German TUS	German Census
couples		
without kids	17.9	20.8
with kids above 6 years	35.2	38.1
with kids below 6 years	21.3	3.1
singles		
without kids	16.0	17.2
with kids	7.0	6.6
other	2.6	14.2

Table 5: Population shares by marital status. Numbers show percentages in population between 25 and 54 years old. German TUS refers to wave 2001/02, and German Census refers to year 2000.

B Summary of notation

	model setup
$j \in C$	indexes for couples
$i = m, f$	individual's index (male, female)
k	the "other" individual in a couple
α_i	preference parameter (consumption vs home and leisure), see (4)
β_i	preference parameter (home prod vs leisure, see (4))
γ_i	exponent in home production function, see (3)
ζ_i	random noise added to non-employment utility, see (18)
M	total non-wage income for couple *without* unemployment benefits
T	time endowment for each individual
n_i	market (work) hours
h_i	home production hours
ℓ_i	leisure hours
z	home consumption
c	market consumption
w_i	wages for individual
n_0	minimum number of working hours for employed
ρ	unemployment benefits as a fraction of maximum possible earnings
$U_i, \hat{U}_i(n_m, n_f)$	utility and indirect utility
	Bayesian model
μ_ω	intercept, (12)
X_m, X_f, X_c	individual covariates, couple-level covariates, (12)
B_m, B_f, B_c	regression coefficient on individual and coulpe-level covariates, (12)
σ_ω, L_ω	standard deviations and Cholesky factor of covariance, (12)
σ_ε	standard deviation of noise for hours observations, (10), (9)
D_j	total observed non-wage income *including* unemployment benefits,
σ_η	standard deviation of noise term for D_j, (8)
u_j	dummy for unemployment benefits being the main source of non-wage income
κ	parameter for mapping to u_j, (11)

	Mean	Std. Dev.	25%	50%	75%	Min.	Max.
				Full sample			
market work, female	4.22	3.41	0	4.67	7.17	0	12.3
market work, male	7.13	2.81	6.67	7.83	8.67	0	14
home prod., female	4.94	2.75	2.83	4.50	6.83	0.17	13.8
home prod., male	2.23	1.74	1	1.83	3	0.17	11
leisure, female	8.84	2.26	7.25	8.58	10.1	3.17	17.5
leisure, male	8.64	2.19	7.33	8.33	9.50	2.83	17.7
				EE			
market work, female	6.34	1.99	4.67	6.25	7.83	2.17	12.3
market work, male	7.92	1.63	7.08	8	8.83	2.67	14
home prod., female	3.70	1.89	2.25	3.50	5	0.17	8.83
home prod., male	1.98	1.30	0.92	1.83	2.83	0.17	8.83
leisure, female	7.96	1.67	6.83	7.92	8.92	3.17	15
leisure, male	8.09	1.61	7.08	8.17	9	2.83	13.5
				EN			
market work, female	0	0	0	0	0	0	0
market work, male	7.86	1.61	7	7.83	8.67	2.17	13.2
home prod., female	7.67	2.37	6.25	7.83	9.33	0.50	13.8
home prod., male	1.73	1.22	0.83	1.50	2.25	0.17	8.50
leisure, female	10.3	2.37	8.67	10.2	11.8	4.17	17.5
leisure, male	8.41	1.64	7.42	8.42	9.33	4.17	14.7
				NE			
market work, female	6.59	1.73	5.17	6.67	7.92	3.67	10.5
market work, male	0	0	0	0	0	0	0
home prod., female	2.94	1.67	1.75	2.92	4.17	0.17	7.50
home prod., male	5.47	2.75	3.58	5.42	7.50	0.33	11
leisure, female	8.47	1.48	7.50	8.42	9.08	5.17	13.3
leisure, male	12.5	2.75	10.5	12.6	14.4	7	17.7
				NN			
market work, female	0	0	0	0	0	0	0
market work, male	0	0	0	0	0	0	0
home prod., female	6.40	1.93	4.96	6.25	7.75	1.58	10.4
home prod., male	4.99	2.05	3.63	4.50	6.54	0.50	10.3
leisure, female	11.6	1.93	10.3	11.8	13.0	7.58	16.4
leisure, male	13.0	2.05	11.5	13.5	14.4	7.75	17.5

Table 6: Detailed descriptive statistics: daily core market work, total home production and leisure, in hours. Entire sample and by couple type. Not weighted with population weights.

C Algebraic details of the model solution

We first fix n_m and n_f, and maximize (4), substituting in the functional form (3). Now let $\tilde{M}_i = M + w_k n_k + \mathbf{1}_{n=0}\rho w_k T$, which the individual takes as given. Then the optimization

Variables	CMW, f	CMW, m	THP, f	THP, m	Leisure, f	Leisure, m
CMW, f	1.000					
CMW, m	0.098	1.000				
THP, f	-0.751	0.033	1.000			
THP, m	0.032	-0.626	-0.007	1.000		
Leisure, f	-0.597	-0.189	-0.081	-0.041	1.000	
Leisure, m	-0.152	-0.785	-0.036	0.008	0.274	1.000

Table 7: Correlation of time use variables. CMW denotes core market work, THP total home production, f denotes female and m male.

problem for the employed member i

$$\mathcal{U}_{E,i} = \max_{n_0 \le n_i \le T, 0 \le h_i \le T-n_i} \alpha_i \log\left(\tilde{M}_i + w_i T\right) + (1-\alpha_i)\left[\beta_i\gamma_i \log h_i + \beta_i\gamma_k \log h_k + (1-\beta_i)\log(T-n_i-h_i)\right]$$

while for the non-employed,

$$\mathcal{U}_{N,i} = \max_{0 \le h_i \le T} \alpha_i \log\left(\tilde{M}_i + \rho w_i T\right) + (1-\alpha_i)\left[\beta_i\gamma_i \log(h_i) + \beta_i\gamma_k \log(h_k) + (1-\beta_i)\log(T-n_i-h_i)\right]$$

Assume that n_m, n_f are fixed, and consider the part which depends on home production hours (and leisure),

$$H_i = \max_{0 \le h_i \le T} \beta_i\gamma_i \log(h_i) + (1-\beta_i)\log(T-n_i-h_i) \quad \text{for } i = m, f$$

This has the first order condition

$$h_i = \frac{\beta_i\gamma_i}{1-\beta_i+\beta_i\gamma_i}(T-n_i) \tag{14}$$

and the optimal value is

$$H_i = (1-\beta_i+\beta_i\gamma_i)\log(T-n_i) + \underbrace{\left[\beta_i\gamma_i \log(\beta_i\gamma_i) + (1-\beta_i)\log(1-\beta_i) + (1-\beta_i+\beta_i\gamma_i)\log(1-\beta_i+\beta_i\gamma_i)\right]}_{\text{constant}}$$

where the constant part does not play a role in the comparison of employed and non-employed states, so it can be ignored.

Now introduce

$$\phi_i = \frac{\alpha_i}{\alpha_i + (1-\alpha_i)(1-\beta_i+\beta_i\gamma_i)} \tag{15}$$

to simplify the algebra.

Dropping the constant terms, the utility of the employed can be rewritten as

$$\mathcal{U}_E(\tilde{M}_i, \phi_i, w_i) = \max_{T \ge n_i \ge n_0} \phi_i \log\left(\tilde{M}_i + n_i w_i\right) + (1-\phi_i)\log(T-n_i) \tag{16}$$

and similarly for the non-employed,

$$\mathcal{U}_N(\tilde{M}_i, \phi_i, w_i) = \phi_i \log\left(\tilde{M}_i + \rho T w_i\right) + (1-\phi_i)\log(T) \tag{17}$$

This means all the free time T and $\rho T w$ in addition to M is considered as a benefit. The utility of any member of a couple is given by

$$\mathcal{U}(\tilde{M}_i, \phi_i, w_i) = \max\left[\mathcal{U}_E(\tilde{M}_i, \phi_i, w_i), \mathcal{U}_N(\tilde{M}_i, \phi_i, w_i) + \zeta_i\right] \tag{18}$$

Lemma 1 (Characterization of individual responses.). *Given the model setup in* (16) *and* (17), *the following hold.*

1. (16) *has an interior solution* n_1 *that satisfies*

$$n_1(\tilde{M}_i,\phi_i,w_i)=\phi_i T-(1-\phi_i)\frac{\tilde{M}_i}{w_i}>n_0 \tag{19}$$

with utility

$$\mathcal{U}_E(\tilde{M}_i,\phi_i,w_i)=\underbrace{\phi_i\log(\phi_i)+(1-\phi_i)\log(1-\phi_i)}_{\equiv A(\phi_i)}+\log(\tilde{M}_i+Tw_i)-(1-\phi_i)\log(w_i) \tag{20}$$

whenever

$$\phi_i>\frac{\tilde{M}_i+n_0 w_i}{\tilde{M}_i+Tw_i}\quad \textit{or equivalently}\quad (\phi_i T-n_0)w_i>(1-\phi_i)\tilde{M}_i \tag{21}$$

Otherwise, the optimal choice is $n=n_0$ *given the constraint.*

2. $\mathcal{U}_E$ *is continuous, continuously differentiable except at* (21) *(at equality), always increasing in* $\tilde{M}_i$, *and increasing in* w_i.
3. $\mathcal{U}_N$ *is increasing in* $\tilde{M}_i$ *and* w_i.
4. $\mathcal{U}_E(\tilde{M}_i,\phi_i,w_i)-\mathcal{U}_N(\tilde{M}_i,\phi_i,w_i)$ *is decreasing in* $\tilde{M}_i$ *and increasing in* w_i *when* (5) *holds.*

Proof. 1. (19) and (21) follow from the optimization problem, then we plug in n_1 to obtain (20), to get

$$\begin{aligned}\mathcal{U}_E(\tilde{M}_i,\phi_i,w_i)&=\phi_i\log(\tilde{M}_i+\phi_i Tw_i-(1-\phi_i)\tilde{M}_i)+(1-\phi_i)\log(T-\phi_i T-(1-\phi_i)\tilde{M}_i/w_i)\\&=\phi_i(\log(\phi_i)+\log(\tilde{M}_i+Tw_i))+(1-\phi_i)(\log(1-\phi_i)+\log(T+\tilde{M}_i/w_i))\\&=\underbrace{\phi_i\log(\phi_i)+(1-\phi_i)\log(1-\phi_i)}_{\equiv\mathcal{A}}+\phi_i\log(\tilde{M}_i+Tw_i))+(1-\phi_i)\log(T+\tilde{M}_i/w_i))\\&=\mathcal{A}+\log(\tilde{M}_i+Tw_i)-(1-\phi_i)\log(w_i)\qquad\text{whenever (21) holds.}\end{aligned}$$

2. Note that

$$\mathcal{U}_E(\tilde{M}_i,\phi_i,w_i)=\begin{cases}\mathcal{A}+\log(\tilde{M}_i+Tw_i)-(1-\phi_i)\log(w_i) & \text{if (21) holds,}\\ \phi_i\log(\tilde{M}_i+n_0w_i)+(1-\phi_i)\log(T-n_0) & \text{otherwise.}\end{cases}$$

Now when $n_0w_i=\phi_i T-(1-\phi_i)\tilde{M}_i/w_i$, $n_1(\tilde{M}_i,\phi_i,w_i)=n_0$, ensuring continuity. Smooth differentiability follows trivially.

First, differentiate by $\tilde{M}_i$. Then

$$\frac{\partial\mathcal{U}_E}{\partial\tilde{M}_i}=\begin{cases}\frac{1}{\tilde{M}_i+Tw_i} & \text{if (21) holds,}\\ \frac{\phi_i}{\tilde{M}_i+n_0w_i} & \text{otherwise,}\end{cases}$$

so $\partial\mathcal{U}_E/\partial\tilde{M}_i>0$. Second, differentiate by w_i to obtain

$$\frac{\partial\mathcal{U}_E}{\partial w_i}=\begin{cases}\frac{T}{\tilde{M}_i+Tw_i}-\frac{1-\phi_i}{w_i} & \text{if (21) holds,}\\ \frac{\phi_i n_0}{\tilde{M}_i+n_0w_i} & \text{otherwise.}\end{cases}$$

But

$$\frac{T}{\tilde{M}_i+Tw_i}-\frac{1-\phi_i}{w_i}=\frac{\phi_i T-(1-\phi_i)\tilde{M}_i/w_i}{Tw_i+\tilde{M}_i}\geq 0$$

because of (21).

3. For $\tilde{M}_i$ it is easily seen that $\mathcal{U}_N$ is increasing, and for w_i, we can differentiate to obtain
$$\frac{\partial \mathcal{U}_N}{\partial w_i} = \frac{\phi_i \rho T}{\tilde{M}_i + \rho T w_i}$$
4. From the above, it follows that, as
$$\frac{\partial}{\partial \tilde{M}_i}(\mathcal{U}_E - \mathcal{U}_N) = \begin{cases} \frac{\rho T w_i - T\phi_i w_i + (1-\phi_i)\tilde{M}_i}{w_i(\rho T w_i + \tilde{M}_i)} & \text{if (21) holds,} \\ \frac{\phi_i}{\tilde{M}_i + n_0 w_i} - \frac{\phi_i}{\tilde{M}_i + \rho w_i T} & \text{otherwise.} \end{cases}$$
For both cases, (5) ensures the result.

Similarly,
$$\frac{\partial}{\partial w_i}(\mathcal{U}_E - \mathcal{U}_N) = \begin{cases} \frac{\tilde{M}_i}{w_i} \frac{\rho T w_i - T\phi_i w_i + (1-\phi_i)\tilde{M}_i}{w_i(\rho T w_i + \tilde{M}_i)} & \text{if (21) holds,} \\ \frac{\phi_i n_0}{\tilde{M}_i + n_0 w_i} - \frac{\phi_i \rho T}{\tilde{M}_i + \rho w_i T} & \text{otherwise.} \end{cases}$$
Again, for both cases, (5) ensures the result.

□

We use the above to find fixed points of (6) and (7). However, we have not shown that the fixed point is unique, and there is nothing in our model setup to ensure that it is. In fact, it is possible to get multiple equilibria: we observed this in couples with a high preference for consumption (i.e. high αs) *and* high wage w_i/M ratios. The intuition for this is that high αs make employment attractive, but high wages make non-employment attractive too, so a couple may end up with one member employed and the other not.[34] For estimation, multiple equilibria are not relevant since we observe a couple's employment status in the data. However, for the counterfactual results of Section 6, we need a practical selection rule, and choose the fixed point with the highest sum of utilities $\mathcal{U}(\tilde{M}_m, \phi_m, w_m) + \mathcal{U}(\tilde{M}_f, \phi_f, w_f)$. Anticipating results, note that we observe multiple fixed points for less than 0.5% of couples so this phenomenon does not affect the results in practice. Algorithm 1 summarizes our solution method.

Algorithm 1 (Equilibrium solution)**.** *Given model parameters γ_i, and couple parameters M, α_i, β_i, w_i, ζ_i,*
1. *calculate ϕ_i using* (15),
2. *for each combination of employment statuses (male, female) $s \in$* EE,NE,EN,NN, *calculate the hours choices n_i using Lemma 1,*
3. *discard the statuses s where the sign of $\zeta_i s$ is not compatible with $\mathcal{U}_E - \mathcal{U}_N$,*
4. *when multiple statuses remain, select the one with the highest total utility,*
5. *calculate home production hours from* (14), *and leisure hours as $\ell_i = T - n_i - h_i$.*

D MCMC diagnostics and prior-posterior comparisons

Figure 9 shows the effective sample size and the potential scale reduction for the MCMC run (5 chains, default NUTS warmup, 3000 in each chain samples after warmup). Both statistics indicate good mixing and convergence (low $\hat{R}$, sufficient effective sample size). We also checked

[34]Extensive simulation suggest that there are never more than two fixed points.

NUTS-specific statistic (e.g. divergence, reaching maximum tree depth) and they are do not suggest any problems with convergence either.

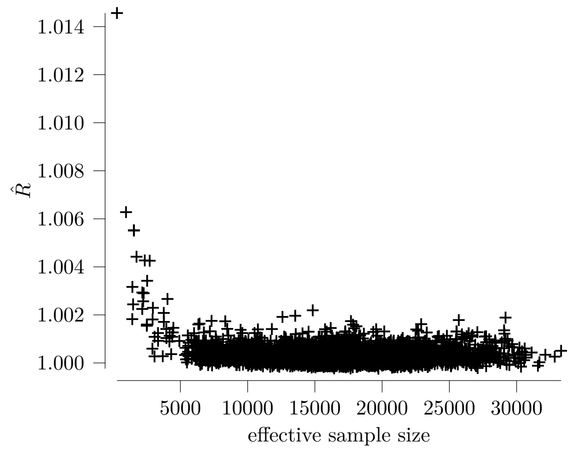

Figure 9: MCMC diagnostics. $\hat{R}$ (potential scale reduction) vs effective sample size (cf Vehtari et al. (2021)). 5×3000 samples after warmup.

Figure 10 compares the priors and the posteriors for cross-sectional parameters. Note the concentration of posteriors relative to the vague priors.

E Selected additional tables and figures

Table 10 shows the posterior quantiles for noise parameters. Table 11 shows the aggregate responses for increasing female wages, keeping male hours fixed and assigning gender-specific homogeneous preferences.

F Description of binning algorithm for plots and tables

We briefly describe the calculation used to produce plots and tables where we bin by quantiles of some variable.

1. Let $(x_i, y_{1,i}, y_{2,i}, \ldots)$, for $i = 1, \ldots$ denote some real-valued quantities, grouped by index i. To make things concrete, x_i can be male wages, $y_{1,i}$ male hours as predicted by the model, and $y_{2,i}$ counterfactual male hours for the same couple assuming that male wages increased by 10%, with i indexing both individuals and posterior draws, flattened to a single index for notational convenience.
2. Calculate $n+1$ quantiles $q_0, \ldots, q_n$ for x_i. For each i, let b_i be the bin index such that $q_{i-1} \le x_i \le q_i$, assuming some tie-breaking rule on the boundaries.
3. We then take the mean of x's and y's within each bin j. Formally, let $\bar{x}_j = \frac{\sum_{i|b_i=j} x_i}{\sum_{i|b_i=j} 1}$ be the mean of x_i in bin j, and similarly $\bar{y}_{k,j} = \frac{\sum_{i|b_i=j} y_{k,i}}{\sum_{i|b_i=j} 1}$ for $y_{1,\cdot}, \ldots$.
4. Finally, we calculate the relevant quantity, such as an elasticity or average hours response, using $\bar{x}_j, \bar{y}_{1,j}, \ldots$.

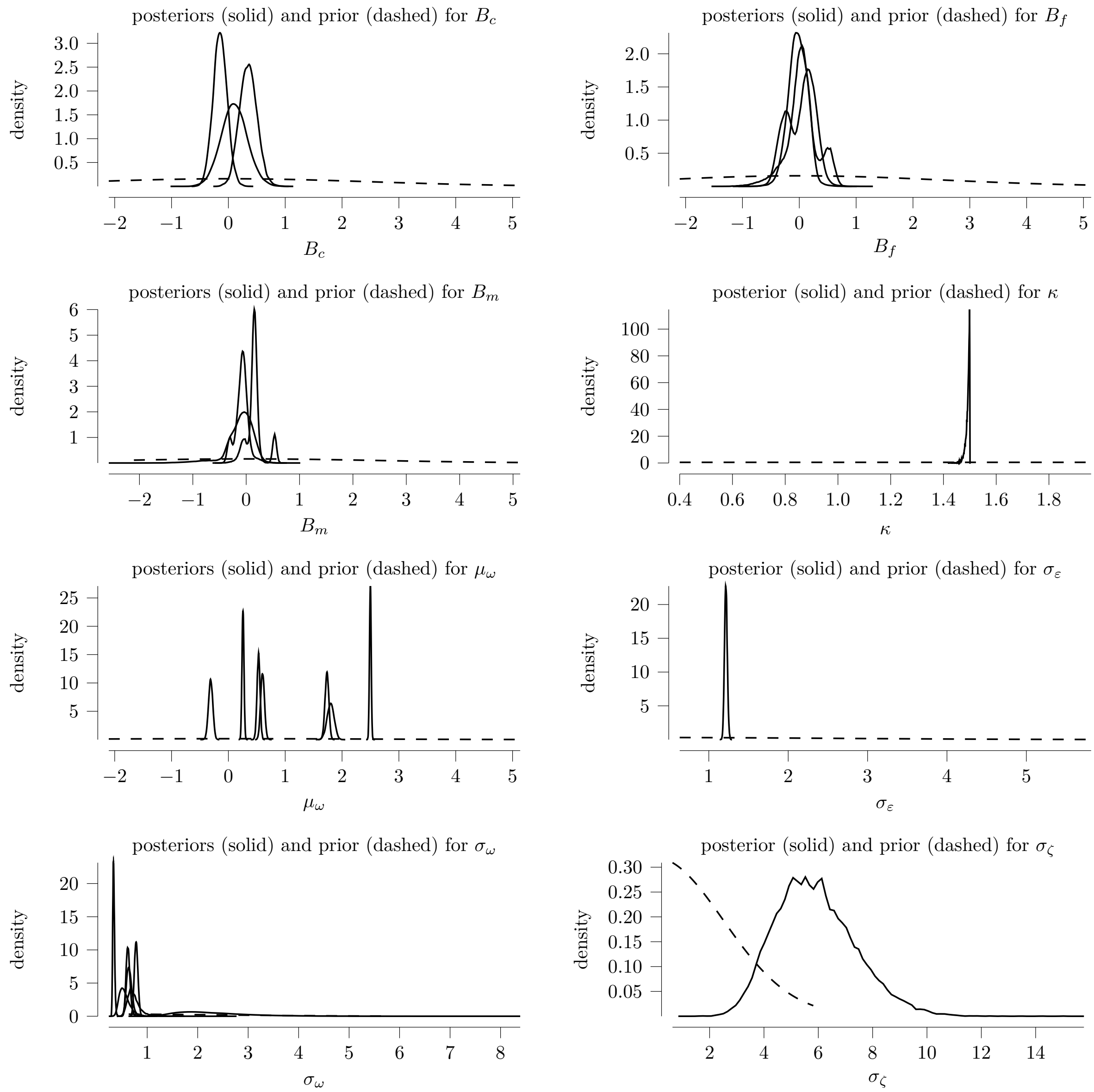

Figure 10: Comparison of priors (dashed) and posteriors (solid, kernel densities with cross-validated bandwidth) for cross-sectional parameters.

Alternatively, for some plots we form quantiles of the y's in each bin, in which case the last step is omitted.

Let $n'_{i;j,j}$ denote the new hours for each experiment for genders for $i = m, f$. We report the market hours changes $n'_{i;j,k} - n_{i;j,k}$, and similarly the elasticity of market hours with respect to wages, calculated as

$$\mathcal{E}_{i;j,k} = \frac{1}{1/10} \frac{n'_{i;j,k} - n_{i;j,k}}{n_{i;j,k}} \quad \text{whenever} \quad n_{i;j,k} > 0$$

which is only defined for members who were originally employed. Note that depending on the experiment we perform, these may be own- or cross elasticities.

	EE	EN	NE	NN
hourly wage, m [€]	13.21	27.90		
	(5.835)	(148.5)		
hourly wage, f [€]	11.51		10.00	
	(28.64)		(4.137)	
non-labor income [€]	227.5	371.8	1092.6	1654.0
	(429.9)	(514.8)	(881.2)	(733.7)
West [frac]	0.738	0.875	0.638	0.523
	(0.440)	(0.332)	(0.483)	(0.502)
married [frac]	0.913	0.982	0.915	0.943
	(0.282)	(0.135)	(0.281)	(0.233)
with kids [frac]	0.775	0.915	0.851	0.727
	(0.418)	(0.279)	(0.358)	(0.448)
age, m [years]	44.28	45.18	45.62	46.45
	(5.908)	(5.111)	(5.881)	(6.121)
age, f [years]	41.82	42.61	42.55	43.45
	(5.942)	(5.067)	(5.319)	(6.170)
university, m [frac]	0.199	0.258	0.128	0.0909
	(0.400)	(0.439)	(0.337)	(0.291)
university, f [frac]	0.152	0.107	0.0426	0.0455
	(0.359)	(0.310)	(0.204)	(0.211)
fh, m [frac]	0.155	0.148	0.149	0.0909
	(0.363)	(0.355)	(0.360)	(0.291)
fh, f [frac]	0.143	0.0738	0.128	0.0682
	(0.350)	(0.262)	(0.337)	(0.255)
sec. school I, m [frac]	0.279	0.310	0.213	0.205
	(0.449)	(0.463)	(0.414)	(0.408)
sec. school I, f [frac]	0.290	0.207	0.106	0.0682
	(0.454)	(0.406)	(0.312)	(0.255)
sec. school II, m [frac]	0.110	0.129	0.106	0.0227
	(0.313)	(0.336)	(0.312)	(0.151)
sec. school II, f [frac]	0.0750	0.0886	0.0851	0.0455
	(0.264)	(0.285)	(0.282)	(0.211)
sec. school III, m [frac]	0.304	0.288	0.298	0.386
	(0.460)	(0.454)	(0.462)	(0.493)
sec. school III, f [frac]	0.461	0.487	0.681	0.659
	(0.499)	(0.501)	(0.471)	(0.479)
Observations	1,146	542	94	88

Table 8: Couples by spouses' labor market status in the 2001/2002 wave. Table shows means of variables with standard deviations in parentheses, not weighted with population weights. Fractions are computed within couple type. Education groups do not sum to one for each couple type, remaining persons belong to other groups. *fh* refers to university of applied sciences, *sec. school I* refers to full university entrance degree (Abitur), *sec. school II* refers to limited university entrance degree (Fachabitur), *sec. school III* refers to lowest secondary schooling degree (Mittlere Reife).

income source	household type				total
	EE	EN	NE	NN	
capital income/property	0.2	0.7	2.1	4.5	0.6
employment	86.7	88.6	53.2	2.3	81.6
other public support	0.0	0.0	4.3	0.0	0.2
pension	0.2	0.0	19.1	22.7	2.1
self-employed/agriculture	12.4	10.7	6.4	2.3	11.1
social security	0.0	0.0	0.0	2.3	0.1
unemployment benefits	0.0	0.0	14.9	63.6	3.7

Table 9: Main source of income [%] by household labor market status. Notes: 2001/2002 sample. Together with missing values entries in each column add to 100%.

	q5%	q25%	q50%	q75%	q95%
σ_ζ	3.73	4.84	5.76	6.81	8.47
σ_ε	1.19	1.20	1.21	1.23	1.24
κ	1.48	1.49	1.49	1.50	1.50

Table 10: Posterior quantiles for noise parameters.

group	extensive (%)		intensive		total	
	male	female	male	female	male	female
all	0.00	0.23	0.00	0.49	0.00	0.50
w Q1	0.00	0.17	0.00	0.66	0.00	0.71
w Q2	0.00	0.30	0.00	0.61	0.00	0.61
w Q3	0.00	0.24	0.00	0.51	0.00	0.52
w Q4	0.00	0.18	0.00	0.38	0.00	0.39
EE	0.00	0.00	0.00	0.51	0.00	0.51
EN	0.00	0.49	0.00	—	0.00	—
NE	0.00	0.00	—	0.28	—	0.28
NN	0.00	1.55	—	—	—	—

Table 11: Aggregate responses to increasing female wages, keeping male hours fixed.