Marriage Market and Labor Market Sorting: Online Appendix

Paula Calvo Ilse Lindenlaub Ana Reynoso

July 21, 2023

Throughout this Online Appendix, we indicate figures, tables, and equations within this appendix by O.# ('O' for Online). In turn, figures, tables, and equations from the main paper are denoted by just 1,2,.... Figures, tables, and equations from the main appendix are denoted by A.#.

OA Empirical Evidence

OA.1 Marriage Market Sorting

Table O.1 reproduces Table 1, for the period 1990-1996. In turn, Tables O.2 and O.3 display the marital sorting parameters based on Eika, Mogstad, and Zafar (2019), whereby each cell reports the likelihood of a match between a man and a women with a certain level of education, relative to the likelihood that this type of match occurs under random matching. Following Eika et al. (2019), we compute the summary measure of sorting (reported in Footnote 5 of Section 2) as the weighted average of the marital sorting parameters along the diagonal of Table O.2 for 2010-2016 and of Table O.3 for 1990-1996.

Table O.1: Marriage Matching Frequencies by Education (1990-1996)

	Low Education Men	Medium Education Men	High Education Men
Low Education Women	**0.26**	0.16	0.03
Medium Education Women	0.14	**0.22**	0.11
High Education Women	0.01	0.02	**0.07**

Notes: Education groups are defined in Online Appendix OD.2.2. We consider the maximum level of education attained by each individual and keep only one observation per couple. Source: GSOEP.

Table O.2: Marital Sorting Parameters (2010-2016)

	Low Education Men	Medium Education Men	High Education Men
Low Education Women	**2.02**	0.43	0.38
Medium Education Women	1.42	**1.39**	1.20
High Education Women	0.36	0.33	**2.08**

Notes: Education groups are defined as in Online Appendix OD.2.2. Each cell reports the likelihood of a match between a man and a women with a certain level of education, relative to the likelihood that this type of match occurs under random matching. Source: GSOEP.

Table O.3: Marital Sorting Parameters (1990-1996)

	Low Education Men	Medium Education Men	High Education Men
Low Education Women	**1.39**	0.81	0.84
Medium Education Women	0.79	**1.24**	3.10
High Education Women	0.06	0.19	**3.69**

Notes: Education groups are defined as in Online Appendix OD.2.2. Each cell reports the likelihood of a match between a man and a women with a certain level of education, relative to the likelihood that this type of match occurs under random matching. Source: GSOEP.

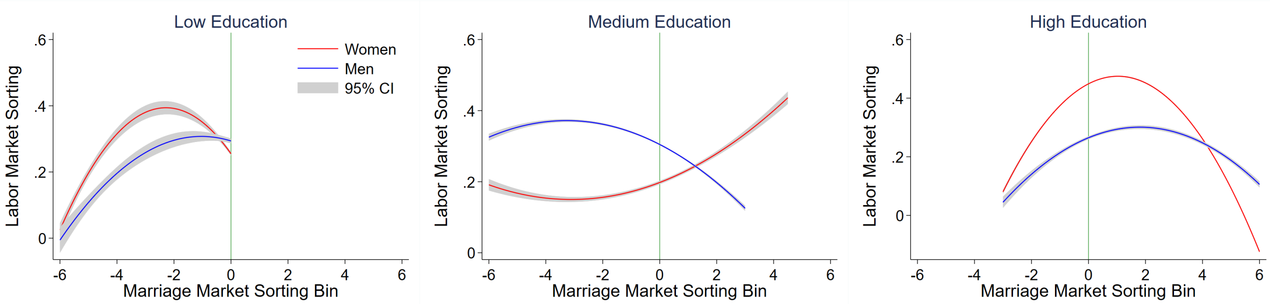
OA.2 Marriage Market and Labor Market Sorting

Table O.4: Labor Market Sorting and Marriage Market Sorting

	(1) Task Complexity	(2) Task Complexity	(3) Task Complexity
Educ.	0.166***	0.201***	0.188***
	(0.008)	(0.005)	(0.004)
Educ. x PAM	0.097***	0.056***	0.071***
	(0.010)	(0.007)	(0.006)
PAM	-0.180***	-0.129***	-0.142***
	(0.020)	(0.015)	(0.012)
Demographic Controls	Yes	Yes	Yes
State and Year FE	Yes	Yes	Yes
Sample	Women	Men	All
Period	2010-2016	2010-2016	2010-2016
Observations	5,145	5,983	11,128
R^2	0.350	0.417	0.377

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (1), Section 2. *Task Complexity* is measured as the percentile of the individual's occupation in the task complexity distribution. *Educ* is the individual's highest education level attained during the sample period. *PAM* is an indicator variable that takes value one when spouses have the same education level. We restrict the sample to individuals in couples. Demographic controls include age and presence of children in the household. Source: GSOEP and BIBB.

Figure O.1: Labor and Marriage Market Sorting, by Education Level



Notes: We reproduce Figure 1 (right) when splitting the sample by education level, as defined in Online Appendix OD.2.2. For each education level and gender, the 'zero' marriage market sorting bin pools couples in which the difference in years of education between partners is less than two years (in absolute value). Source: GSOEP and BIBB.

Table O.5: Labor Market Sorting and Marriage Market Sorting, by Education Level

	(1) Task Complexity	(2) Task Complexity	(3) Task Complexity
Years of Educ.	0.070***	0.023***	0.027***
	(0.010)	(0.004)	(0.004)
Years of Educ. x PAM	0.021*	0.026***	0.010**
	(0.012)	(0.005)	(0.005)
PAM	-0.223*	-0.328***	-0.148*
	(0.122)	(0.064)	(0.083)
Demographic Controls	Yes	Yes	Yes
State and Year FE	Yes	Yes	Yes
Sample	Low Educ	Mid Educ	High Educ
Period	2010-2016	2010-2016	2010-2016
Observations	2,680	5,121	3,058
R^2	0.125	0.057	0.102

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (1), Section 2, but by education level. *Task Complexity* is measured as in Table O.4. *Years of Educ* is measured as the individual's maximum years of education attained. *PAM* is an indicator variable that takes value one when spouses have the same education level, defined as a difference in years of education between the partners of less than two years (in absolute values). Each column corresponds to a different sample based on the education level of the individual, where education levels are defined in Online Appendix OD.2.2. We pool men and women. Other sample restrictions and demographic controls are as in Table O.4. Source: GSOEP and BIBB.

OA.3 Complementarity in Home Production Hours and in Labor Market Hours

OA.3.1 Complementarity in Home Production Hours

Table O.6: Complementarity in Home Production Hours

	(1) Male Hours	(2) Male Hours
Female Hours	0.192***	0.208***
	(0.018)	(0.019)
Demographic Controls	Yes	Yes
State and Year FE	Yes	Yes
Sample	All	Same Education
Period	2010-2016	2010-2016
Observations	7,124	4,142
R^2	0.169	0.180

Notes: Standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (2), Section 2. Male and female hours correspond to hours spent in home production and missing data on home hours are imputed, see Online Appendix OD.2.2 for details. Demographic controls: male education, male age, and presence of children in the household. For consistency with our sample in Table O.7, we pool observations from West/East Germany, consider couples with positive labor hours and do not impose sample restrictions on the marital history of individuals or occupation. Source: GSOEP.

OA.3.2 Complementarity in Labor Market Hours: Instrumental Variable Approach

As discussed in Section 2, we leverage variation across states and over time in the expansion of childcare availability in Germany, induced by a law ("Das Kinderförderungsgesetz") passed in December 2008. This law aimed to provide universal and subsidized childcare for children between 1 and 3 years old, by August 1, 2013 (see Müller and Wrohlich (2020) for more details about this law and its implementation). We use this variation to instrument for female labor market hours ($FemaleHours_{cts}$) in regression (2) and causally estimate the effect of changes in female labor hours on the labor hours of her male partner. Our instrument is the share of children between 1 to 3 years old in state s and year t enrolled in childcare ($ShareChildcareSlots_{ts}$), interacted with an indicator that takes value 1 when there is a child in that age group in household c, year t and state s ($ChildHH_{cts}$).[1] Our first stage is given by:

$$\begin{aligned}\text{FemaleHours}_{cts} =& \beta_0 + \beta_1 \text{ShareChildcareSlots}_{ts} \\ & + \beta_2 \text{ShareChildcareSlots}_{ts} \times \text{ChildHH}_{cts} + X_{ct}\boldsymbol{\Gamma} + \delta_t + \delta_s + \epsilon_{cts},\end{aligned}$$

where vector X_{ct} includes the same set of controls as regression (2) in Section 2 and δ_t and δ_s capture year and state fixed effects.

Our identification assumption is that the increase in childcare availability for small children has a direct impact on female labor market hours, but only affects male labor hours indirectly, through changes in his partner's hours. Several pieces of evidence support this exclusion restriction: First, gender norms in Germany are such that women still take on most of the burden of childcare, independently of the partner's education, and adjust their labor supply in response to the arrival of a child. As a result, we expect changes in childcare availability to mostly affect female time allocation. Second, we find that the presence of a small child in the household has a strong negative impact on female labor hours, but no significant impact on male labor hours. Therefore, policies that allow households to outsource childcare would have a *direct* impact on the mother's (but not on the father's) labor market hours. Finally, note that our exclusion restriction is in line with previous work on the impact of childcare expansion in Germany on labor market outcomes, which only emphasizes *female* outcomes, with positive effects on maternal employment (Boll and Lagemann, 2019; Müller and Wrohlich, 2020; Bauernschuster and Schlotter, 2015).[2]

Columns 3 and 4 of Table O.7 contain the IV results. Column 5 reports the first stage.

[1] We use data at the state and year level on the number of children between 1 and 3 years old enrolled in childcare and the total number of children in that age group, obtained from the German Federal Statistical Office (*Statistisches Bundesamt*). We then construct the *share* of children enrolled in childcare by state and year. Under the assumptions of excess demand for childcare slots and full take-up of available slots, this corresponds to the share of slots offered in each year and state (Müller and Wrohlich, 2020).

[2] Müller and Wrohlich (2020) use the same policy change to estimate the impact of childcare availability on maternal time allocation, as in our first stage. However, they use the German Micro-Census and their empirical strategy also differs.

Table O.7: Complementarity in Labor Market Hours

	OLS Male Hours	OLS Male Hours	IV Male Hours	IV Male Hours	First Stage Female Hours
Female Hours	-0.006	0.004	0.351**	0.505**	
	(0.017)	(0.021)	(0.170)	(0.237)	
Share Childcare Slots					16.054***
					(3.553)
Share Childcare Slots × Child in HH					4.934***
					(1.433)
Demographic Controls	Yes	Yes	Yes	Yes	Yes
State and Year FE	Yes	Yes	Yes	Yes	Yes
Sample	All	Same Education	All	Same Education	All
Period	2006-2016	2006-2016	2006-2016	2006-2016	2006-2016
Observations	10,533	6,124	10,533	6,124	10,533
R^2	0.017	0.022			0.257

Notes: Standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (2), Section 2. Male and female hours correspond to labor market hours, as defined in Online Appendix OD.2.2. Columns (3) and (4) instrument female labor market hours with the share of available daycare slots for children 1 to 3 years old, and its interaction with the presence of a child of that age group in the household. Columns (2) and (4) restrict the sample to couples in which both spouses have the same level of education. Demographic controls are as in Table O.6, but we also control for the presence of children between 1 to 3 years old. In all regressions, we pool observations from West and East Germany to capture more of the regional variation in childcare availability. We extend our sample period to 2006-2016, in line with the timing of the childcare policy rollout. To increase the number of observations in our IV regressions (here and also in Table O.9), we do not impose restrictions on marital histories or occupations. We focus on couples in which both partners work in the labor market. Source: GSOEP.

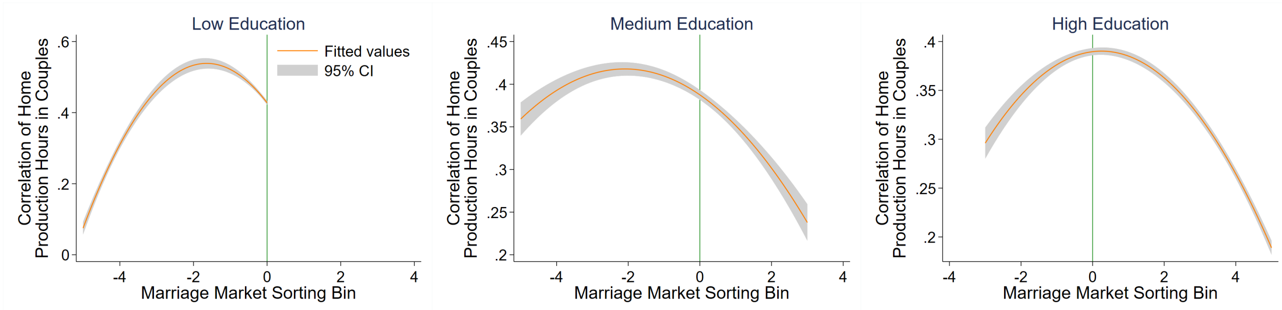
OA.3.3 Complementarity in Home Production Hours, By Education Level

Table O.8: Complementarity in Home Production Hours, by Male Partner's Education Level

	(1) Male Hours	(2) Male Hours	(3) Male Hours	(4) Male Hours	(5) Male Hours	(6) Male Hours
Female Hours	0.234***	0.235***	0.198***	0.220***	0.145***	0.168***
	(0.041)	(0.050)	(0.017)	(0.022)	(0.020)	(0.035)
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
State and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Low Educ.	Low Educ.	Medium Educ.	Medium Educ.	High Educ.	High Educ.
	Male	Both	Male	Both	Male	Both
Period	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016	2010-2016
Observations	1,619	564	3,556	2,513	1,927	1,065
R^2	0.194	0.236	0.181	0.199	0.168	0.208

Notes: Standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (2), Section 2. Male and female hours correspond to hours spent in home production, as defined in Online Appendix OD.2.2. Results in Columns (1), (3) and (5) replicate column (1) in Table O.6 but condition on the education level of the male partner (Low, Medium and High Education, defined as in Online Appendix OD.2.2). Columns (2), (4) and (6) replicate column (2) in Table O.6, i.e. they condition on both partners having the same education level. Demographic controls and sample restrictions are as in Table O.6. Source: GSOEP.

Figure O.2: Complementarities in Home Production Hours, by Male Partner's Education Level



Notes: We reproduce Figure 2 (left), when splitting the sample by education level, as defined in Online Appendix OD.2.2. For each education level and gender, the 'zero' marriage market sorting bin pools couples in which the difference in years of education between partners is less than two years (in absolute value). Source: GSOEP.

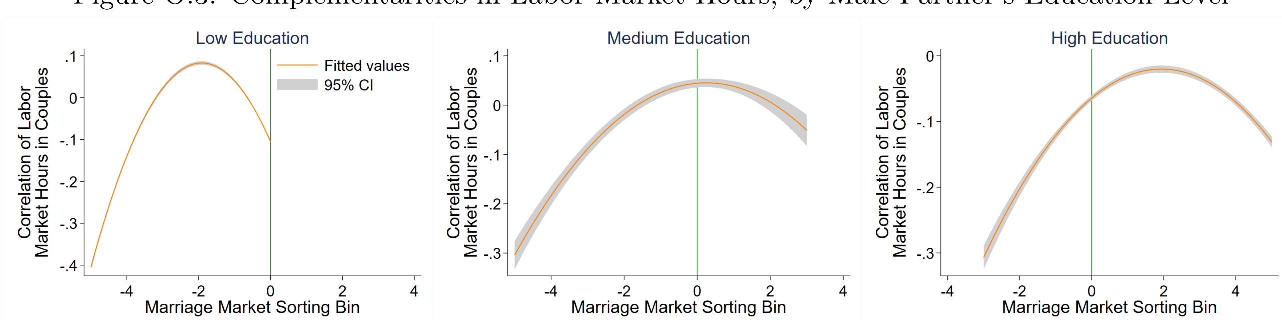
OA.3.4 Complementarity in Labor Market Hours, By Education Level

Table O.9: Complementarity in Labor Market Hours, by Male Partner's Education Level (IV Results)

	(1) Male Hours	(2) Male Hours	(3) Male Hours	(4) Male Hours	(5) Male Hours	(6) Male Hours
Female Hours	0.249	-0.406	0.692*	0.699***	0.548	0.555
	(0.329)	(0.293)	(0.401)	(0.267)	(0.459)	(0.489)
Demographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
State and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Low Educ.	Low Educ.	Medium Educ.	Medium Educ.	High Educ.	High Educ.
	Male	Both	Male	Both	Male	Both
Period	2006-2016	2006-2016	2006-2016	2006-2016	2006-2016	2006-2016
Observations	2,460	860	5,197	3,684	2,849	1,580

Notes: Standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The table is based on regression (2), Section 2. Male and female hours correspond to labor market hours, as defined in Online Appendix OD.2.2. Columns (1), (3) and (5) replicate column (3) of Table O.7, but condition on the education level of the male partner (Low, Medium and High Education, defined as in Online Appendix OD.2.2). Columns (2), (4) and (6) replicate column (4) of Table O.7, i.e., they condition on both partners being in the same education group. Demographic controls, instruments and additional sample restrictions are as in Table O.7. Source: GSOEP.

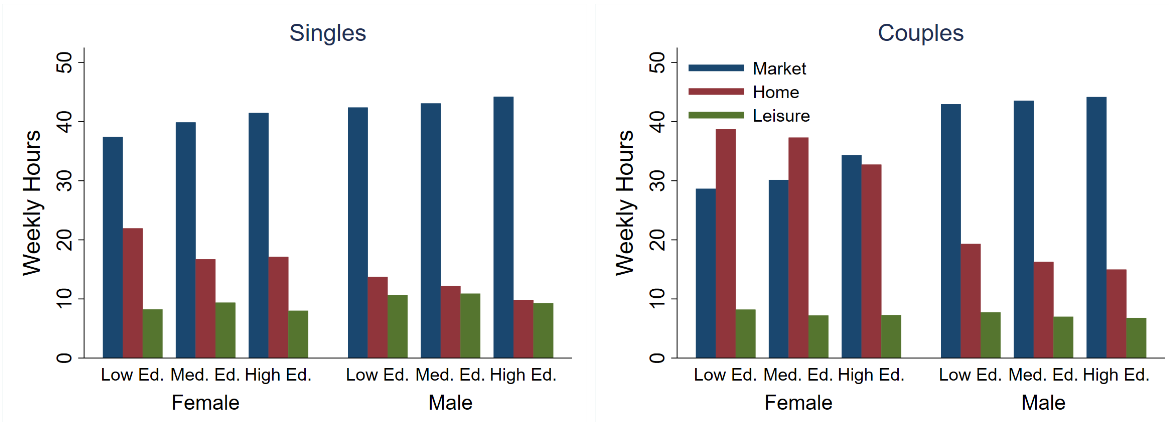
Figure O.3: Complementarities in Labor Market Hours, by Male Partner's Education Level



Notes: We reproduce Figure 2 (right), when splitting the sample by education level, as defined in Online Appendix OD.2.2. We pool the data as in Figure O.2. Source: GSOEP.

OA.4 Time Allocation by Marital Status and Education

Figure O.4: Weekly Mean Hours by Gender, Education and Marital Status

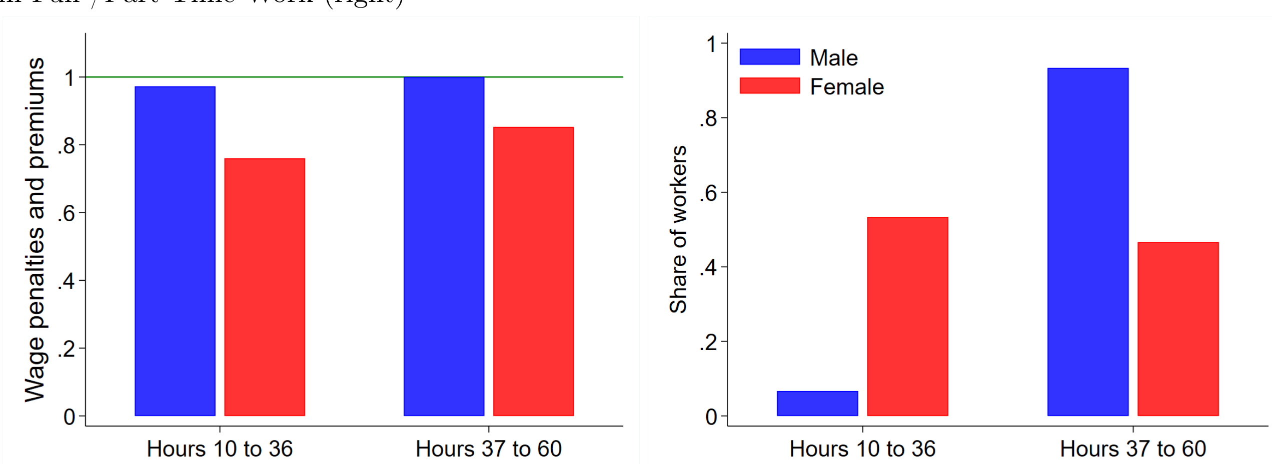


Notes: The figure reports average weekly hours by individuals in the labor market ('market'), in home production ('home') and spent on leisure activities ('leisure'), by gender, education and marital status. Source: GSOEP.

OA.5 Time Allocation and Wages

The left panel of Figure O.5 shows the wage penalties of various groups *relative* to men working full-time (the blue bar, whose height equals one), where full-time is defined as working between 37.5 and 60 hours per week. The height of the remaining bars in Figure O.5 (left) reflects the estimates of a regression of the logarithm of hourly wages on an indicator for part-time work, a gender indicator and the interaction between both. We control for age, marital status, birth place and education. While full-time women have a wage penalty of 14.7 percentage points relative to full-time men, women who work part-time have an even larger penalty of 24 percentage points. Figure O.5 (right) shows that while less than 10% of employed men work part-time, more than 50% of employed women do so.

Figure O.5: Part-Time Wage Penalties Relative to Men Who Work Full-Time (left); Share of Workers in Full-/Part-Time Work (right)



Notes: The left panel is based on a log wage regression on part-time work and gender indicators (plus controls), described in Online Appendix OA.5. The right panel indicates the share of men and women in part-/full-time work. Source: GSOEP.

OB Theory

OB.1 TU Representation and Monotone Equilibrium

TU Representation and the Generalized Quasi-Linear Class. A broad class of utility functions makes our model TU-representable, according to which households' overall split between public and private consumption is independent of how private consumption is shared. It is well-known (Chiappori and Gugl, 2020) that the class of Generalized Quasi-Linear (GQL) utility functions is necessary and sufficient for the TU representation of the type of problem we are dealing with. The GQL is given by

$$u(c,p) = F(\alpha(p)c + \beta(p)),$$

where functions F, α, β are strictly increasing and β is strictly concave. Some concrete examples that fall into the GQL class and thus yield the TU representation in our model are:

$$u(c,p) = F(c+p), \quad \text{for any } F' > 0 \tag{O.1}$$

$$u(c,p) = \log(cp) \tag{O.2}$$

$$u(c,p) = \alpha(p)m(c) \quad \text{for any } m \text{ s.t. } m(xy) = m(x)m(y)$$

$$u(c,p) = \frac{c^{1-\gamma}}{1-\gamma}\frac{p^{1-\theta}}{1-\theta} = \frac{1}{(1-\gamma)(1-\theta)}\left(cp^{\frac{1-\theta}{1-\gamma}}\right)^{1-\gamma}, \gamma \geq 0, \theta \geq 0.$$

This highlights that many common utility functions lead to the TU representation in our model, including those that allow for flexible degrees of risk aversion.

Monotone Equilibrium Beyond the Quasi-Linear Class. The reason why we focus on quasi-linear utility (O.1) within the GQL class is twofold: First, this assumption increases tractability because it guarantees the TU property. Second, and most importantly, it also allows us to obtain clean analytical (sufficient) conditions for the monotone equilibrium, which captures well the features of the data we document in Section 2 of the paper. For several other functional forms in the GQL class, the conditions for monotone equilibrium would be more involved because they encapsulate forces toward household specialization and therefore toward the *non*-monotone equilibrium of our model.

For instance, consider log-utility (O.2). The complication with log-utility is that private and public consumption goods are *complements* in the utility function, generating a force towards household specialization (one partner works a lot at home so that p is high and the other one works a lot in the market so that c is also high), which pushes *against* the monotone equilibrium that—by definition—has the requirement that spouses' hours are complements and therefore co-move.

Nevertheless, more general utility functions from the GQL class (such as log-utility) do lead to predictions consistent with monotone equilibrium (and thus with the data) when considering a natural extension of our framework. This extension is to allow couples to purchase some of the public good in the market using their wages. We now denote by $\tilde{p} = p + c_p$ the public good, where c_p is the public good expenditure in the market while p is home-produced, as before. The idea is that the consumption good that is purchased in the market can not only be used for private consumption but also as an input for the public good. We assume for illustration that the public good produced at home and the one purchased in the market are perfect substitutes. This gives us analytical tractability and is plausible given the high degree of substitutability between home production and market purchases found in the literature (e.g., Chang and Schorfheide, 2003 and Greenwood, Guner, and Vandenbroucke, 2017).

We now take a closer look at our model with log-utility under this extension: We start with the household decision problem. The collective decision problem of any potential couple (x_m, x_f) that forms in the marriage market is given by

$$
\begin{aligned}
& \max_{c_p, c_f, c_m, h_m, h_f} \log(c_m \tilde{p}) \\
s.t. \qquad & c_p + c_m + c_f - w(\tilde{x}_m) - w(\tilde{x}_f) = 0 \\
& \log(c_f \tilde{p}) \geq \overline{v},
\end{aligned}
$$

where we normalized prices to 1, as is often assumed in this literature (see, e.g., many of the applications in Browning, Chiappori, and Weiss, 2014).

Plugging the constraints into the objective, and using a monotone transformation, gives

$$
\begin{aligned}
& \max_{c_p, h_m, h_f} \log\left(\left(w(\tilde{x}_m) + w(\tilde{x}_f) - c_p - \frac{\exp(\overline{v})}{\tilde{p}}\right)\tilde{p}\right) \\
\Leftrightarrow \quad & \max_{c_p, h_m, h_f} \big(w(\tilde{x}_m) + w(\tilde{x}_f) - c_p\big)\big(c_p + p(1-h_m, 1-h_f)\big).
\end{aligned}
$$

Taking the FOC w.r.t. c_p gives: $c_p = (w(\tilde{x}_m) + w(\tilde{x}_f) - p(1-h_m, 1-h_f))/2$. Plugging c_p back into the objective (and applying another monotone transformation) gives:

$$
\max_{h_m, h_f} w(\tilde{x}_m) + w(\tilde{x}_f) + p(1-h_m, 1-h_f),
$$

which closely resembles our objective function from the baseline model. Thus, in the household stage, the conditions for hours to be increasing in own and partner's type remain the same. In the marriage market stage, the conditions for positive sorting are also the same. The conditions on primitives for monotone equilibrium are therefore *identical* to those in the baseline model with quasi-linear preferences.

The intuition for this result is straightforward: Substitutability in the public good $\tilde{p}$ between the component produced at home using hours, p, and the one that is purchased with wages, c_p, *undoes* the complementarity between p and c_m—or between p and c_f—in the utility function. So even though under log-utility, private and public consumption (c_m and $\tilde{p}$) are complements (and thus wages and $\tilde{p}$ are complements, pushing towards non-monotone equilibrium), the substitutability between p and c_p in $\tilde{p}$ counteracts these forces. This re-introduces a substitutability between wages (that purchase c_p and also c_m) and p. Therefore, if p is complementary in spouses' home production time, this setting gives rise to the monotone equilibrium, in which spouses' hours are positively correlated, as in the baseline model.

As a result, what our simple functional form (quasi-linearity) captures is in essence a more general utility function from the GQL class when allowing for substitutability in the public good between the component produced at home and the component purchased in the market.

OB.2 Regular Equilibrium

We define a *regular equilibrium* as follows.

Definition O1 (Regular Equilibrium)**.** *An equilibrium is regular if household problem (5) has a unique solution that is interior and continuous in (x_m, x_f), and $\tilde{N}$ is atomless.*

At the cost of additional technical detail, we could justify the regularity properties of Definition O1 in terms of primitives.[3] However, the approach we follow (i.e., starting immediately from regular equilibrium) allows us to focus on the monotone structure of the model in a more direct way.

OB.3 Connecting the Properties of Monotone Equilibrium to the Stylized Facts

We connect the properties of monotone equilibrium with our stylized facts from Section 2 in a *qualitative* way, before accurately replicating them in the quantitative analysis of Section 4.5.

Marriage Market Sorting. The property of positive sorting in the marriage market resembles our empirical finding of positive sorting on partners' education in Table 1.

Labor Market Sorting. In a monotone equilibrium, more skilled individuals work more in the labor market than at home compared with less skilled individuals, a feature that is reinforced by the fact that high skilled individuals have more skilled partners. As a result, more skilled individuals (i.e. those with higher x) have higher effective types $\tilde{x}$, and thereby obtain more productive labor market matches: There is positive sorting in the labor market in (x, y), which captures the positive correlation between workers' education and jobs' task complexity in the data (Figure 1, left, solid lines).

Marriage Market and Labor Market Sorting. The unique feature of our model is the link between labor and marriage market equilibrium and, in particular, labor and marriage sorting. This link becomes most transparent when highlighting how labor market matching function μ depends on the marriage market matching function. Consider the total derivative $(\mu)_{x_f}$ (for $(\mu)_{x_m}$, this is similar), which—when positive—indicates PAM on the labor market in skills and skill requirements (x, y):

$$(\mu)_{x_f} = \mu' \left(e_{x_f} + e_h \left(\frac{\partial h_f}{\partial x_m} \eta' + \frac{\partial h_f}{\partial x_f} \right) \right), \tag{O.3}$$

[3] Here we provide a sketch: The property of an interior solution of the household problem can be justified based on Inada conditions on p. In turn, we can specify conditions on the objective function of the household problem that render a unique solution (which also implies stability of the equilibrium in a tatonnment sense). Uniqueness, in turn, allows for the application of the Implicit Function Theorem, which guarantees differentiability and therefore continuity of the hours functions in (x_m, x_f). Finally, an atomless $\tilde{N}$ can be guaranteed if, in addition, the effective type function e is assumed to be Morse (i.e., a function with only isolated critical points).

which is based on $\mu(\tilde{x}_f)$ with $\tilde{x}_f = e(x_f, h_f(\eta(x_f), x_f))$. Equation (O.3) illustrates how labor market matching, μ, depends on marriage market matching, η. When marriage market sorting is positive, $\eta' > 0$, then higher x_f are matched to higher y, $(\mu)_{x_f} > 0$ (given that the hours of spouses are complementary $\partial h_f / \partial x_m > 0$). The intuition is straightforward. PAM in the marriage market induces women with higher x_f to have a better partner $x_m = \eta(x_f)$ and therefore to work more hours. This translates into a higher effective type $\tilde{x}_f$ and thus a better labor market match $y = \mu(\tilde{x}_f)$, compared with when marriage market sorting is not positive. In a stylized way, this property of the monotone equilibrium is related to our empirical fact that labor market sorting is stronger for positively sorted couples (Figure 1, right).

The Role of Hours. In a monotone equilibrium, labor hours are complementary within couples: Increasing, say, female skills, not only pushes up her own labor hours but also induces her partner to work more. As a result, partners' hours comove. There are two drivers behind this result. First, for a given male partner type x_m (i.e., if marriage matching was *exogenous*), an increase in female skills increases her labor hours. But this reduces her home hours, which induces her partner to also work less at home and more in the market due to the home production complementarity, $p_{\ell_m \ell_f} > 0$. As a result, *both* partners increase their labor hours as the female skill improves. Second, this comovement of spouses' hours is reinforced under *endogenous* marriage market sorting: Under PAM, an increase in female skill x_f leads to a better partner $x_m = \eta(x_f)$, who by himself puts in more labor hours and fewer home hours. And since $p_{\ell_m \ell_f} > 0$, the wife adjusts hours in the same direction (fewer home hours and more labor hours), strengthening the comovement of hours within the couple. Thus, PAM in the marriage market *fuels* the complementarity of hours within couples—a feature we observe in the data (Figure 2).

Finally, an interesting feature of a monotone equilibrium is that it can be consistent with a gender gap in labor market sorting: If the home production technology induces women to spend relatively more time at home (e.g., if they are relatively more productive at home), then men will be 'better' matched in the labor market compared with women of the same skill. Thus, our competitive model can generate a gender gap in sorting and wages even in the absence of discrimination or differential frictions.

To see this, consider labor market sorting in terms of skills and firm productivity (x_i, y) and how it varies across gender $i \in \{f, m\}$. Consider a man and a woman with $x_f = x_m$. We say that x_m is 'better sorted' than x_f if $\mu(e(x_m, h_m)) > \mu(e(x_f, h_f))$. Then, for each man and woman of equal skills, $x_f = x_m$, man x_m is better sorted if he works more hours in the labor market, $h_m(x_m, \eta^{-1}(x_m)) > h_f(\eta(x_f), x_f)$, which will help rationalize our finding in the data on the gender gap in labor market sorting (solid lines Figure 1, left). But controlling for hours worked, $h_m(x_m, \eta^{-1}(x_m)) = h_f(\eta(x_f), x_f)$, closes the sorting gap in the model and considerably shrinks it in the data (dashed lines Figure 1, left).

OB.4 Non-Monotone Equilibrium

We here state the result on sufficient conditions for the non-monotone equilibrium.

Proposition O1 (Non-Monotone Equilibrium)**.** *Assume p is strictly submodular in (ℓ_m, ℓ_f), and z*

is strictly supermodular in $(\tilde{x}, y)$ and convex in $\tilde{x}$ for each y. Then any regular equilibrium is non-monotone.

Proof. The proof is analogous to that of Proposition 1, which is why we shorten the arguments.

LABOR MARKET PROPERTIES. For the argument for positive labor market sorting, see Proposition 1.

MARRIAGE MARKET PROPERTIES. Turning to the marriage stage, consider any couple (x_m, x_f) who jointly chooses hours h_m and h_f. The household's problem is

$$\Phi(x_m, x_f) = \max_{h_m, h_f \in [0,1]} \left(w(e(x_m, h_m)) + w(e(x_f, h_f)) + 2p(1 - h_m, 1 - h_f) \right),$$

where we replaced $\tilde{x}_m = e(x_m, h_m)$ and $\tilde{x}_f = e(x_f, h_f)$. If hours are strictly *decreasing* in the partner's type, then $\Phi(x_m, x_f)$ is strictly submodular and thus *negative* sorting emerges in the marriage market.

PROPERTIES OF THE HOURS FUNCTIONS. We will show that h_m strictly increases in x_m and strictly *decreases* in x_f. To do so, write the couple's problem as follows:

$$\max_{h_m \in [0,1]} \left(w(e(x_m, h_m)) + \max_{h_f \in [0,1]} \left(w(e(x_f, h_f)) + 2p(1 - h_m, 1 - h_f) \right) \right).$$

That is, we again split the joint maximization w.r.t. (h_m, h_f) into two maximization problems.

Let V be the value of the inner maximization problem, that is

$$V(x_f, h_m) = \max_{h_f \in [0,1]} \big(w(e(x_f, h_f)) + 2p(1 - h_m, 1 - h_f) \big).$$

We will first focus on the outer maximization problem

$$\max_{h_m \in [0,1]} \Big(w(e(x_m, h_m)) + V(x_f, h_m) \Big),$$

taking as given the function V. The objective function is additively separable in (x_m, h_m) and (x_f, h_m). Given the properties of e and w, we have that $w(e(\cdot, \cdot))$ is strictly supermodular in (x_m, h_m) since it is the composition of a convex function with a strictly supermodular one, implying that h_m is increasing in x_m. And if V is strictly *submodular* in (x_f, h_m), it will follow that h_m is decreasing in x_f (we will verify the submodularity of V below).

To show that h_m is *strictly* increasing in x_m, we follow the same argument as in the Proposition 1. And to show that h_m is *strictly* decreasing in x_f, we also follow this line of argument, only taking into account a submodular (instead of supermodular) V.

Let us now consider the inner maximization problem:

$$V(x_f, h_m) = \max_{h_f \in [0,1]} \left(w(e(x_f, h_f)) + 2p(1 - h_m, 1 - h_f) \right).$$

Note that the first term of the objective function is again strictly supermodular in (x_f, h_f) (since $w(e(\cdot,\cdot))$ is the composition of a supermodular and a convex function); and the second one is strictly submodular in (h_m, h_f) (under our premise that p is strictly submodular in (ℓ_m, ℓ_f)). So, any optimal h_f in this problem is increasing in x_f and decreasing in h_m. And since h_m is strictly increasing in x_m (see above), we obtain that h_f is decreasing in x_m as well (by taking the composition of functions).

We now show, as above, that given that the solution of the inner maximization problem is interior for all (x_f, h_m)—again ensured by the Inada conditions on p—it must be strictly increasing in x_f and strictly decreasing in h_m. 'Strictly increasing' follows from the same argument as in the proof of Proposition 1. In turn, 'strictly decreasing' also follows from the same line of argument, only taking into account that p is strictly submodular. Thus, h_f is strictly increasing in x_f and strictly decreasing in h_m.

To complete the proof, it remains to show that V is strictly submodular in (x_f, h_m) and differentiable in h_m. Note that

$$V(x_f, h_m) = w(e(x_f, h_f(x_f, h_m))) + 2p(1 - h_m, 1 - h_f(x_f, h_m)).$$

By the Envelope Theorem (Milgrom and Segal, 2002, Corollary 4(iii)), V is differentiable in h_m with $V_{h_m}(x_f, h_m) = -2p_{\ell_m}(1 - h_m, 1 - h_f(x_f, h_m))$. Since p is strictly submodular and h_f strictly increasing in x_f, it follows that V is strictly submodular. Hence, the premise made above when analyzing the choice of h_m in the outer maximization problem holds.

The remaining part of the proof is analogue to that of Proposition 1. Specifically, since hours are strictly monotonic in each attribute (x_f, x_m), it follows that $\tilde{N}$ is indeed atomless. □

OB.5 Monotone Equilibrium in the Quantitative Extension of the Model

We now show that the key properties of our baseline model (Propositions 1 and O1 of the paper) are preserved in the quantitative extension of our model, when adjusted for the stochastic nature of hours and matching. Since Propositions 1 and O1 are flip sides of each other, we here focus on generalizing Proposition 1. We will show that under similar conditions as in the baseline model, the properties of monotone equilibrium hold *on average*.

We will maintain the following assumptions (see also Section 4.1). Marriage taste shocks and labor supply shocks follow type-I extreme-value distributions:

$$\begin{aligned} \beta^s &\sim \text{Type I}(0, \sigma_\beta) \qquad \text{for } s \in \{\mathcal{S} \cup \emptyset\} \\ \delta^{h^t} &\sim \text{Type I}(0, \sigma_\delta) \qquad \text{for } h^t \in \mathcal{H} \text{ and } t \in \{M, U\}. \end{aligned}$$

where

$$\delta^{h^t} = \begin{cases} \delta^{h_i}, i \in \{f, m\} & \text{if } t = U \\ \delta^{h_f} + \delta^{h_m} & \text{if } t = M. \end{cases}$$

That is, when making hours choices, a decision-making unit—either a married or a single household—

draws *only one* labor supply shock for their time allocation, δ^{h^t}, which is extreme-value distributed.

Proposition O2. *If p is strictly supermodular in (ℓ_m, ℓ_f) and z is strictly supermodular in $(\tilde{s}, y)$ and convex $\tilde{s}$, then the properties of monotone equilibrium are satisfied on average, i.e.*

1. *labor market: there is positive sorting between effective human capital $\tilde{s}$ and firm types y;*
2. *households: on average, labor hours h_i are increasing in own type s_i and in partner's type s_j, for $i, j \in \{f, m\}$ and $i \neq j$;*
3. *marriage market: on average, there is positive sorting in the sense that higher s_f match with higher s_m.*

Proof.

Part 1. Recall the firm's problem

$$\max_{\tilde{s}} \quad z(\tilde{s}, y) - w(\tilde{s}).$$

Based on well-known arguments, given that z is assumed to be strictly supermodular in $(\tilde{s}, y)$, the optimal matching satisfies positive sorting in $(\tilde{s}, y)$.

Part 2. We want to show that $\widehat{\Pi}_m(h_m|s_m)$ and $\widehat{\Pi}_f(h_f|s_m)$ are decreasing in s_m, where

$$\widehat{\Pi}_m(h_m|s_m) \equiv \sum_{s_f} \Pi_m(h_m|s_m, s_f)\eta_{s_m}(s_f)$$
$$\widehat{\Pi}_f(h_f|s_m) \equiv \sum_{s_f} \Pi_f(h_f|s_m, s_f)\eta_{s_m}(s_f)$$

are marginal cdfs of male and female hours conditional on male (own and partner) type, so that higher male s_m-types are associated with stochastically higher labor market hours both for them and their partner (the argument for hours being increasing in female types s_f is analogous and omitted). Note that $\eta_{s_m}(s_f)$ is the conditional marriage probability of s_m choosing s_f, and $\Pi_m(h_m|s_m, s_f)$ and $\Pi_f(h_f|s_m, s_f)$ are the marginal cdfs of male and female hours, conditional on partners' types,

$$\Pi_m(h_m|s_m, s_f) \equiv \sum_{\tilde{h}_m < h_m} \sum_{\tilde{h}_f} \pi_{(s_f, s_m)}(\tilde{h}_f, \tilde{h}_m)$$
$$\Pi_f(h_f|s_m, s_f) \equiv \sum_{\tilde{h}_f < h_f} \sum_{\tilde{h}_m} \pi_{(s_f, s_m)}(\tilde{h}_f, \tilde{h}_m),$$

obtained from the probability that spouses (s_f, s_m) choose the joint hours allocation (h_f, h_m), $\pi_{(s_f,s_m)}(h_f, h_m)$ (see (O.9), Online Appendix OC.2):

$$\pi_{(s_f,s_m)}(h_f, h_m) \quad = \quad \frac{\exp(\overline{u}_{\mathbf{s}}(\mathbf{h})/\sigma_\delta)}{\sum_{\tilde{\mathbf{h}} \in \{\mathcal{H} \cup \emptyset\}^2} \exp(\overline{u}_{\mathbf{s}}(\tilde{\mathbf{h}})/\sigma_\delta)}.$$

We dropped the household type superscript t to reduce notation (we exclusively focus on couples here). We will now derive the conditions under which $\widehat{\Pi}_m(h_m|s_m)$ is decreasing in s_m step-by-step.

As we are interested in the impact of s_m on $\widehat{\Pi}_m(h_m|s_m)$, we apply the discrete chain rule to obtain:

$$\Delta_{s_m}\widehat{\Pi}_m(h_m|s_m) = \sum_{s_f}\Pi_m(h_m|s_m,s_f)\Delta_{s_m}\eta(s_f|s_m) + \sum_{s_f}\Delta_{s_m}\Pi_m(h_m|s_m,s_f)\eta(s_f|s_m) \tag{O.4}$$

where we denote the discrete derivative of a function $f(n)$ by $\Delta_n f(n) = f(n+1) - f(n)$. We want to establish conditions under which $\Delta_{s_m}\widehat{\Pi}_f(h_f|s_m) \leq 0$ in (O.4).

We can further simplify this expression by applying summation by parts to the first term. To do so, let's index the different female types by k, so $s_{f_k}, k \in \{0,...,n\}$. Then:

$$\begin{aligned}\sum_{s_f}\Pi_m(h_m|s_m,s_f)\Delta_{s_m}\eta(s_f|s_m) &= \sum_{k=0}^{n}\Pi_m(h_m|s_m,s_{f_k})\Delta_{s_m}\eta(s_{f_k}|s_m)\\ &= \Pi_m(h_m|s_m,s_{f_n})\sum_{k=0}^{n}\Delta_{s_m}\eta(s_{f_k}|s_m) - \sum_{j=0}^{n-1}\Delta_{s_{f_j}}\Pi_m(h_m|s_m,s_{f_j})\sum_{k=0}^{j}\Delta_{s_m}\eta(s_{f_k}|s_m)\end{aligned}$$

where in the first term, we have

$$\sum_{k=0}^{n}\Delta_{s_m}\eta(s_{f_k}|s_m) = \sum_{k=0}^{n}\eta(s_{f_k}|s_{m+1}) - \sum_{k=0}^{n}\eta(s_{f_k}|s_m) = 0$$

and thus the first term vanishes due to the standard property of cdfs. Therefore,

$$\sum_{s_f}\Pi_m(h_m|s_m,s_f)\Delta_{s_m}\eta(s_f|s_m) = -\sum_{j=0}^{n-1}\Delta_{s_{f_j}}\Pi_m(h_m|s_m,s_{f_j})\sum_{k=0}^{j}\Delta_{s_m}\eta(s_{f_k}|s_m)$$

where

$$\sum_{k=0}^{j}\Delta_{s_m}\eta(s_{f_k}|s_m) = \sum_{k=0}^{j}\eta(s_{f_k}|s_{m+1}) - \sum_{k=0}^{j}\eta(s_{f_k}|s_m) \leq 0$$

if the cdf $H(s_f|s_m) \equiv \sum_{k=0}^{j}\eta(s_{f_k}|s_m)$ is decreasing in s_m—something we will verify below—so that higher s_m are matched with higher s_f in the FOSD sense.

Thus, the first term in (O.4) is negative if $\Delta_{s_f}\Pi_m(h_m|s_m,s_f) \leq 0$ for all s_f (and given s_m); and the second term is negative if $\Delta_{s_m}\Pi_m(h_m|s_m,s_f) \leq 0$ for all s_m (and given s_f). To derive conditions under which this holds (i.e. under which $\Pi_m(h_m|s_m,s_f)$ is decreasing in s_m (given s_f) and in s_f (given s_m)) it suffices to show conditions under which the *marginal* pmf of male hours $\sum_{\tilde{h}_f}\pi_{(s_f,s_m)}(\tilde{h}_f,h_m)$ satisfies the monotone likelihood property (or, equivalently, is log-supermodular) in (h_m,s_m) (for fixed s_f), and in (h_m,s_f) (for fixed s_m), as these properties imply FOSD of the marginal cdf $\Pi_m(h_m|s_m,s_f)$ in male and female types, respectively.

The denominator of the joint pmf, $\pi_{(s_f,s_m)}(h_f,h_m)$, does not depend on the specific hours bundle (h_f,h_m), and so the denominator of the marginal pmf, $\sum_{\tilde{h}_f}\pi_{(s_f,s_m)}(\tilde{h}_f,h_m)$ does not depend on it either. We can thus focus on the numerator of $\sum_{\tilde{h}_f}\pi_{(s_f,s_m)}(\tilde{h}_f,h_m)$ when establishing log-supermodularity. That is, we aim to show under which conditions:

$$\sum_{\tilde{h}_f}\exp(\overline{u}_{(s_f,s_m)}(\tilde{h}_f,h_m)/\sigma_\delta)$$

is log-supermodular in (h_m,s_m) (for fixed s_f), and in (h_m,s_f) (for fixed s_m). For this it suffices that the log-transformed summand is supermodular pairwise,[4] meaning in (h_m,s_m), (h_m,h_f), (h_f,s_m) (for fixed s_f) and in (h_m,s_f), (h_m,h_f), (h_f,s_f) (for fixed s_m), where

$$\log(\exp(\overline{u}_{(s_f,s_m)}(h_f,h_m)/\sigma_\delta)) = (w\,(\psi s_f h_f) + w(s_m h_m) + 2p^M(1-h_m,1-h_f))/\sigma_\delta.$$

Supermodularity in (h_f,s_m) and in (h_m,s_f) is trivially satisfied (at equality). Supermodularity in (h_m,h_f) holds under the premise of strictly supermodular p. Finally, supermodularity in (h_m,s_m) and (h_f,s_f), holds if the wage function is supermodular in these pairs, which—as in the baseline model—is true if z is strictly supermodular and convex, as assumed.

A similar argument and analogous conditions establish that $\widehat{\Pi}_f(h_f|s_m)$ is decreasing in s_m. We have thus shown that under the premise of the proposition (and if marriage sorting is positive in a stochastic sense, to which we will turn next), labor hours are stochastically increasing in own and partner's type.

Part 3. We seek conditions under which matching probability $\eta(s_f,s_m)$ is log-supermodular in (s_f,s_m), or equivalently, conditions that ensure the monotone likelihood ratio property of $\eta(s_f,s_m)$. The important implication will be that higher s_m-type men are matched to higher s_f-type women in the FOSD sense or, equivalently—using again the notation $H(s_f|s_m)\equiv\sum_{k=0}^{j}\eta(s_{f_k}|s_m)$ for the probability that a man with s_m marries a woman of type weakly below s_f—that $H(s_f|s_m)$ is decreasing in s_m.

Log-supermodularity of $\eta(s_f,s_m)$ is ensured if

$$\frac{\eta(s''_f,s''_m)}{\eta(s'_f,s''_m)}\geq\frac{\eta(s''_f,s'_m)}{\eta(s'_f,s'_m)} \tag{O.5}$$

for all $s''_f > s'_f$ and $s''_m > s'_m$. Plug into this inequality the expression for $\eta(s_f,s_m)$ as a function of match surplus, obtained from (A.7):

$$\frac{\exp\left(\frac{1}{2\sigma_\beta}surplus(s''_f,s''_m)\right)}{\exp\left(\frac{1}{2\sigma_\beta}surplus(s'_f,s''_m)\right)}\geq\frac{\exp\left(\frac{1}{2\sigma_\beta}surplus(s''_f,s'_m)\right)}{\exp\left(\frac{1}{2\sigma_\beta}surplus(s'_f,s'_m)\right)}.$$

For (O.5) to hold it thus suffices that $\exp\left(\frac{1}{2\sigma_\beta}surplus(s_f,s_m)\right)$ is log-supermodular in (s_f,s_m) or that

[4] This step uses the fact that log-supermodularity is preserved under discrete sums, as shown by de Clippel, Eliaz, and Rozen (2014), Proof of Theorem 4, Step 3.

the match surplus, $surplus(s_f, s_m)$, is supermodular in (s_f, s_m).

Recall that match surplus is given by:

$$surplus(s_m, s_f) := \sigma_\delta \Bigg[\kappa + \log\Big(\sum_{\mathbf{h}^M \in \mathcal{H}^2} \exp\left(\overline{u}^M_{\mathbf{s}}(\mathbf{h}^M)/\sigma_\delta \right) \Big) \\ - \Bigg(2\kappa + \log\Big(\sum_{\mathbf{h}^U \in \mathcal{H}} \exp\left(\overline{u}^U_{s_m}(\mathbf{h}^U)/\sigma_\delta \right) \Big) + \log\Big(\sum_{\mathbf{h}^U \in \mathcal{H}} \exp\left(\overline{u}^U_{s_f}(\mathbf{h}^U)/\sigma_\delta \right) \Big) \Bigg) \Bigg].$$

The last two sums depend on only either s_m or s_f, so we can focus on the first sum, which depends on both $\mathbf{s} = (s_f, s_m)$. That first sum is supermodular in (s_m, s_f) if

$$\log\Big(\sum_{\mathbf{h}^M \in \mathcal{H}^2} \exp\left(\overline{u}^M_{\mathbf{s}}(\mathbf{h}^M)/\sigma_\delta \right) \Big) = \log\Big(\sum_{(h_f, h_m) \in \mathcal{H}^2} \exp\left\{ \frac{w(s_m h_m) + w(\psi s_f h_f) + 2p^M(1-h_m, 1-h_f)}{\sigma_\delta} \right\} \Big)$$

is supermodular, or if

$$\sum_{(h_f, h_m) \in \mathcal{H}^2} \exp\left\{ \frac{w(s_m h_m) + w(\psi s_f h_f) + 2p^M(1-h_m, 1-h_f)}{\sigma_\delta} \right\}$$

is log-supermodular in (s_f, s_m). This is the case if the summand

$$\exp\left\{ \frac{w(s_m h_m) + w(\psi s_f h_f) + 2p^M(1-h_m, 1-h_f)}{\sigma_\delta} \right\}$$

is log-supermodular *pairwise*, or if $w(s_m h_m) + w(\psi s_f h_f) + 2p^M(1-h_m, 1-h_f)$ is supermodular pairwise in $(s_f, s_m), (s_f, h_f), (s_f, h_m), (s_m, h_f), (s_m, h_m), (h_f, h_m)$.[5] Thus, $\eta(s_f, s_m)$ is log-supermodular—and therefore higher s_f women match with higher s_m men in the FOSD sense and thus on average—if the wage function is supermodular in (s_i, h_i) (which again is satisfied if z is supermodular and weakly convex) and if the home production function p is supermodular in (ℓ_m, ℓ_f). □

OC Computation

The solution of our model consists of solving for a fixed point in the wage function (as a function of effective types). That is, we find the market-clearing wage function that induces households, which form in the marriage market, to optimally supply labor (pinning down their effective types) such that, when optimally sorting into firms in the labor market, this gives rise to that exact same wage function.

In practice, we first solve for the optimal matching in the marriage market and households' labor supply choices for a *given* wage function. Given the induced labor supply decisions, individuals optimally match with firms on the labor market. Sorting in the labor market gives rise to a new wage function that

[5]This step again uses the fact that log-supermodularity is preserved under discrete sums, as shown by de Clippel, Eliaz, and Rozen (2014), Proof of Theorem 4, Step 3.

supports this particular matching. Given this new wage function, individuals make new marriage and labor supply decisions that again affect wages in the labor market. We iterate between the problem of households on the one hand and the problem of workers and firms on the other until the wage function converges (until a fixed point in the wage function is found).

We next describe the solution in each decision stage, starting backwards from the labor market and then going to household and marriage problems. Finally, we outline the algorithm to find the fixed point and investigate the equilibrium properties of existence and uniqueness numerically.

OC.1 Partial Equilibrium in the Labor Market ([lpe])

First, we show how to solve for the matching and wage functions in the labor market, (μ, w). Consider our exogenous distribution of firms, $y \sim G$, and *any given* distribution of effective types, $\tilde{s} \sim \tilde{N}_s$, i.e., we take marriage and household outcomes (η, h_f, h_m) as given. Note that even though $\tilde{N}_s$ is an endogenous object in our model, firms take this distribution as given.

To solve for the optimal matching between firms and workers note that the supermodular production function $z(\tilde{s}, y)$ induces positive assortative between y and $\tilde{s}$, so matching function μ is increasing in $\tilde{s}$. Moreover, the wage function w is derived from the (discrete version) of firms' optimality condition (4), evaluated at the optimal matching μ. Since G and $\tilde{N}_s$ are discrete, we approximate the integral in (4) numerically, using trapezoidal integration.

OC.2 Optimal Household Choices ([hh])

Second, we derive the solution to the household problem that yields spouses' optimal labor supply (h_f, h_m) and the distribution of effective types $\tilde{N}_s$, taking labor market and marriage market outcomes (μ, w, η) as given.

Individuals arrive at the household stage either as singles with human capital s_i or in a couple with human capital bundle (s_m, s_f). We denote the household human capital type by the bundle $\mathbf{s} = (s_m, s_f) \in \{\mathcal{S} \cup \emptyset\}^2$ where, e.g., $(s_f, \emptyset)$ denotes the household of single woman of type s_f.

Couples solve problem (A.1) and singles solve problem (A.2), see Appendix B. Plugging the constraints into the objective function and using the transferable utility property (based on quasi-linear preferences with F being the identity function as assumed in estimation), the collective problem of couple (s_m, s_f), after labor supply shocks realize, is given by:

$$\max_{h_m, h_f} w(\tilde{s}_m) + w(\tilde{s}_f) + 2p^M(1 - h_m, 1 - h_f) + \delta^{h_m} + \delta^{h_f}, \tag{O.6}$$

where $w(\tilde{s}_m)$ and $w(\tilde{s}_f)$ depend on hours through the effective human capital types (7).

Similarly, the problem of a single woman of type s_f, after realization of her labor supply shocks, is

$$\max_{h_f} w(\tilde{s}_f) + p^U(1-h_f) + \delta^{h_f}, \tag{O.7}$$

and analogously for a single men of type s_m who choose h_m.

To derive aggregate labor supply and the distribution of effective types $\tilde{N}_s$, we need to introduce some notation. We denote the hours that a decision maker can choose by $\mathbf{h} \in \{\mathcal{H} \cup \emptyset\}^2 := \{\{0, ..., 1\} \cup \emptyset\}^2$ (where $\emptyset$ indicates the hours of the non-existing partner when the individual is single). We then denote by $\mathbf{h}^t$ the hours alternative chosen by a decision maker of type $t \in \{M, U\}$:

$$\mathbf{h}^t = \begin{cases} (h_i, \emptyset), i \in \{f, m\} & \text{if } t = U \\ (h_f, h_m) & \text{if } t = M, \end{cases}$$

where type $t = U$ indicates *single* (or Unmarried) and type $t = M$ indicates *married*.

Also, we denote the economic utility associated with hours alternative $\mathbf{h}^t$ of household type $t \in \{M, U\}$ with human capital type $\mathbf{s} \in \{\mathcal{S} \cup \emptyset\}^2$ by $\overline{u}^t_{\mathbf{s}}(\mathbf{h}^t)$, where

$$\overline{u}^t_{\mathbf{s}}(\mathbf{h}^t) = \begin{cases} w(\tilde{s}_i) + p^U(1-h_i), i \in \{f, m\} & \text{if } t = U \\ w(\tilde{s}_m) + w(\tilde{s}_f) + 2p^M(1-h_m, 1-h_f) & \text{if } t = M. \end{cases} \tag{O.8}$$

We obtain the optimal private consumption and labor supply (c_m, c_f, h_m, h_f) for each household by solving problems (O.6) and (O.7). Given our assumption that the labor supply shock distribution is Type-I extreme value (see Section 4.1), we then obtain the probability that household type $t \in \{M, U\}$ with human capital type $\mathbf{s} \in \{\mathcal{S} \cup \emptyset\}^2$ chooses hours alternative $\mathbf{h} \in \{\mathcal{H} \cup \emptyset\}^2$ as:

$$\pi^t_{\mathbf{s}}(\mathbf{h}^t) \quad = \quad \frac{\exp(\overline{u}^t_{\mathbf{s}}(\mathbf{h}^t)/\sigma_\delta)}{\sum_{\tilde{\mathbf{h}}^t \in \{\mathcal{H} \cup \emptyset\}^2} \exp(\overline{u}^t_{\mathbf{s}}(\tilde{\mathbf{h}}^t)/\sigma_\delta)}. \tag{O.9}$$

Again denoting the fraction of households with human capital bundle $\mathbf{s}$ by $\eta(\mathbf{s})$, the fraction of households that has human capital composition $\mathbf{s}$ *and* chooses hours alternative $\mathbf{h}$ is given by $\eta(\mathbf{s}) \times \pi^t_{\mathbf{s}}(\mathbf{h}^t)$.

From this distribution of *household* labor supply we back out the distribution of *individual* labor supply. To do so, we compute the fraction of men and women of each *individual* human capital type, s_i, $i \in \{f, m\}$, in household $\mathbf{s}$, optimally choosing each individual hours alternative h_i associated with household labor supply $\mathbf{h}$. Given the distribution of individual labor supply, we can compute the distribution of effective skills, $\tilde{N}_s$. First, note that the support of this distribution is obtained by applying functional forms (7) for any combination of *individual* hours and skill types. Second, to each point in the support of $\tilde{s}$, we attach the corresponding mass using the individual labor supply distribution obtained above as well as the exogenously given distribution of human capital N_s.

OC.3 Partial Equilibrium in the Marriage Market ([mpe])

In the marriage stage, individuals draw idiosyncratic taste shocks for partners and singlehood, β_i^s, with $i \in \{f, m\}$, from a type-I extreme value distribution (see Section 4.1). At this stage, labor supply shocks are not yet realized. As a result, the *ex ante* economic value from a marriage involving types (s_m, s_f) is the expected value of (O.6); and the ex-ante economic value from singlehood is the expected value of (O.7). In both cases, the expectation is taken over the distribution of δ-shocks. Denoting the utility transfer from a male spouse with s_m to a female spouse with s_f by $v(s_m, s_f)$, the values of being married (economic plus non-economic) for a female type s_f and a male type s_m in couple (s_m, s_f) are given by

$$
\begin{aligned}
& v(s_m, s_f) + \beta_f^{s_m} \\
\Phi(s_m, s_f, v(s_m, s_f)) + \beta_m^{s_f} &:= \mathbb{E}_\delta \Big\{ \max_{h_m, h_f} w(\tilde{s}_m) + w(\tilde{s}_f) + 2p^M(1-h_m, 1-h_f) + \delta^{h_m} + \delta^{h_f} \Big\} + \beta_m^{s_f} - v(s_m, s_f) \\
&= \sigma_\delta \Big[\kappa + \log \Big(\sum_{\mathbf{h}^M \in \mathcal{H}^2} \exp\{ \overline{u}_{\mathbf{s}}^M(\mathbf{h}^M)/\sigma_\delta \} \Big) \Big] + \beta_m^{s_f} - v(s_m, s_f),
\end{aligned}
$$

where $\kappa = 0.57722$ is the Euler constant and $\mathbb{E}_\delta$ is the expectation taken over the distribution of δ-shocks.

In turn, we denote the values of being single for woman s_f and man s_m respectively by:

$$
\begin{aligned}
v(\emptyset, s_f) + \beta_f^{\emptyset} &:= \sigma_\delta \Big[\kappa + \log \Big(\sum_{\mathbf{h}^U \in \mathcal{H}} \exp\{ \overline{u}_{(\emptyset, s_f)}^U(\mathbf{h}^U)/\sigma_\delta \} \Big) \Big] + \beta_f^{\emptyset} \\
\Phi(s_m, \emptyset, v(s_m, \emptyset)) + \beta_m^{\emptyset} &:= \sigma_\delta \Big[\kappa + \log \Big(\sum_{\mathbf{h}^U \in \mathcal{H}} \exp\{ \overline{u}_{(s_m, \emptyset)}^U(\mathbf{h}^U)/\sigma_\delta \} \Big) \Big] + \beta_m^{\emptyset},
\end{aligned}
$$

where we note that $v(s_m, \emptyset) = 0$ and—with some abuse of notation—$v(\emptyset, s_f) \neq 0$ as specified, which has no longer the interpretation of a transfer. Every man of type s_m and every woman of type s_f then choose the skill type of their partner or to remain single in order to maximize their value on the marriage market:

$$
\begin{aligned}
\max \quad & \{ \max_{s_f \in \mathcal{S}} \Phi(s_m, s_f, v(s_m, s_f)) + \beta_m^{s_f}, \Phi(s_m, \emptyset, v(s_m, \emptyset)) + \beta_m^{\emptyset} \} \\
\max \quad & \{ \max_{s_m \in \mathcal{S}} v(s_m, s_f) + \beta_f^{s_m}, v(\emptyset, s_f) + \beta_f^{\emptyset} \}.
\end{aligned}
$$

In practice, using the transferable utility property of our model, we solve for the *optimal* marriage matching by maximizing the total sum of marital values across all individuals in the economy, using a linear program:

$$
\begin{aligned}
\max_{\eta(s_m, s_f) \in [0,1]} \quad & \sum_{(s_m, s_f) \in \{\mathcal{S} \cup \emptyset\}^2} \eta(s_m, s_f) \times \big(\tilde{\Phi}(s_m, s_f) + \tilde{\beta} \big) \\
s.t. \quad & \sum_{s_m \in \mathcal{S}} \eta(s_m, s_f) = n_{s_f} \quad \text{and} \quad \sum_{s_f \in \mathcal{S}} \eta(s_m, s_f) = n_{s_m},
\end{aligned}
$$

where $\eta(s_m, s_f)$ denotes the mass of households with human capital $(s_m, s_f) \in \{\mathcal{S} \cup \emptyset\}^2$; n_{s_m} and n_{s_f} are the exogenous probability mass functions for male and female human capital corresponding to cdf N_s; $\tilde{\Phi}(s_m, s_f)$ denotes the economic value from marriage for the different types of households, $\tilde{\Phi}(s_m, s_f) \in \{\Phi(s_m, s_f, v(s_m, s_f)) + v(s_m, s_f), v(s_f, \emptyset), \Phi(s_m, \emptyset, v(s_m, \emptyset))\}$; and $\tilde{\beta}$ denotes $\beta_f^{s_m} + \beta_m^{s_f}$ for couples and $\beta_i^{\emptyset}$ $(i = \{f, m\})$ for singles. The constraints this linear program imposes are that the marginal distributions of η are consistent with the exogenous human capital distribution n_{s_i} for both genders.

We obtain the equilibrium matching in the marriage market, η, by solving this linear program, taking prices and allocations in households and the labor market, $(w, \mu, h_f, h_m, \tilde{N}_s)$, as given.

OC.4 General Equilibrium of the Model

Once we have derived the solution of each stage taking the output from the other stages as given, we solve for the *general equilibrium* of the model by searching for the prices, allocations, and assignments such that all markets are simultaneously in equilibrium. We start with an arbitrary initial wage function. In the household stage, each *potential* household takes those wages as given and makes their labor supply choices, after observing the realization of their labor supply shocks. These optimal labor supply choices are then used by each individual in the marriage market to compute the expected value of singlehood and marriage with different partners, leading to marriage choices and thus household formation. In turn, the hours choices of *formed* households give rise to a distribution of effective types. With this distribution in hand, we go back to the labor market stage, where we optimally match worker effective types with firm productivities. This labor market matching gives rise to a new wage function supporting this allocation. Given this new wage function, we solve and update the household and marriage problems and iterate until convergence, i.e. until we have found a fixed point in the wage function.

Trembling Effective Types. A challenge in the search for the equilibrium is that each household type needs to know the wage for *any* hours choice in order to make its optimal labor supply choices. However, it may be the case that at a given iteration of our fixed point algorithm, the wage function is such that certain levels of hours are *not* chosen by some household types. Therefore, in the next iteration, agents would face a wage function that only maps *realized* effective types to a wage (i.e., a wage function 'with gaps in the support'), see Online Appendix OC.1. The problem then is that agents do not know the payoff from all potential hours choices when they try to make their optimal choice.

To address this issue, we develop a trembling strategy. We draw a small random sample of women and men and force them to supply a suboptimal amount of hours from the set of unchosen hours in each iteration. In practice, for each group of women with skill type s_f and each group of men with skill type s_m, we track their optimal choices for a given wage function and determine the hours that were *not* chosen with positive probability. We then draw a 1% random sample of women and men within each of those skill types (the 'tremblers') and assign them uniformly to the unchosen hours. Finally, we construct the distribution of effective types $\tilde{N}_s$ by taking into account both 'trembling' effective types and 'realized' effective types.

Fixed Point Algorithm. To solve for the general equilibrium, we denote by $\tilde{N}_s^*$ the distribution of *realized* effective types (based on *optimal* hours choices, as opposed to *trembling* hours choices). Similarly, we denote by w^* the wage as a function of *realized* effective types only, while the *full support* wage function is denoted by w. The fixed point algorithm we designed to solve for the equilibrium is as follows:

0. Initiate a round-zero ($r = 0$) wage function for all possible effective types, w^0.

At any round $r \geq 1$

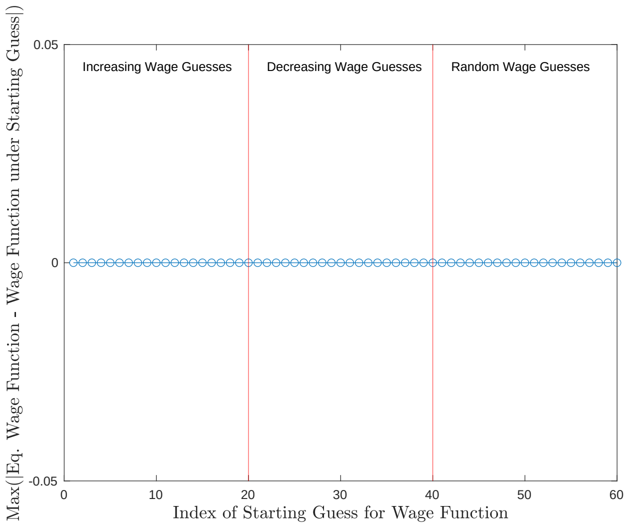
1. Input w^{r-1} and solve [hh] and [mpe]. Update $\tilde{N}_s^{*r}$.
2. Input $\tilde{N}_s^{*r}$ and solve [lpe]. Update w^{*r}.
3. Update w^r:
   (a) We determine w^{*r} from step 2. above.
   (b) Simultaneously, we fill in the wage for effective types that did not realize at round r by solving step 2. for *trembling types*. Along with (a) this yields w^r.
4. Move to round $r + 1$ by going back to step 1. and continue iterating until the wage function converges element-by-element, that is for each $\tilde{s}$, $w^r(\tilde{s}) - w^{r-1}(\tilde{s}) < \epsilon$ for small $\epsilon > 0$.
5. (OUTPUT) Compute the general equilibrium as the tuple of outputs from [hh], [mpe], and [lpe] at the round where the wage function w^r converged.

OC.5 Numerical Existence and Uniqueness of Equilibrium

Existence. Starting from a guess for the wage function, we always converged with a very high level of accuracy to a fixed point in wages. At convergence, the wage function that results from choices in the labor market coincides with the wage function that agents took as given when making marriage, household and labor market choices. This suggests that our numerical solution method indeed finds an equilibrium.

Uniqueness. Our numerical investigations suggest that the equilibrium we found is unique, whereby the equilibrium wage function does *not* depend on the initial starting guess. We systematically investigated this for different classes of (random) initial guesses with certain properties: (i) initial guesses for a wage function that is increasing in hours and skills; (ii) initial guesses for a wage function that is decreasing in hours and skills; (iii) completely random initial guesses for the wage function. In each class, we randomly draw 20 wage functions that comply with the properties of the class. We then compare element-by-element the convergent wage function under each initial guess with our equilibrium wage function. In Figure O.6, we plot the maximal element-wise deviation (i.e., the absolute value of the difference) between the two: Independently of the initial guess, the convergent wage function is always exactly the same as our equilibrium wage function, with a maximal element-wise deviation of zero.

Figure O.6: Difference Between Our Equilibrium Wage Function and the Convergent Wage Function Resulting from 60 Alternative Initial Guesses



OD Data and Sample Construction

OD.1 Data Sources

In this section, we provide details on our three sources of data:

GSOEP. The main dataset used for the empirical analysis (Section 2) and the estimation (Section 4.4) is the German Socio Economic Panel (SOEP, 2019), a household survey conducted by the German Institute of Economic Research (Deutsches Institut fuer Wirtschaftsforschung) starting in 1984. The core study of the GSOEP surveys about 25,000 individuals living in 15,000 households each year. All individuals aged 16 and older respond to the individual questionnaire. The head of household additionally answers a household questionnaire. This survey is longitudinal in nature, and collects rich information on demographics (such as marital status, education, fertility, family background, etc.), labor market variables (including hours worked, wages, and occupation), and detailed time-use information. Important for us, the GSOEP contains the same information for both the head of household and their partner (whether married or cohabiting).

Throughout our analysis, we focus on West Germany. Our baseline period is 2010-2016. We also consider an earlier period, 1990-1996, for over-time comparisons. We do not use data from before 1990 because key time-use variables are missing.

BIBB. Our main data source for measuring occupation types is the BIBB Employment Survey collected in 2012 by the German Federal Institute of Vocational Training (Bundesinstitut fuer Berufsbildung—BIBB), and the German Federal Institute for Occupational Safety and Health, see Hall et al. (2020) and Rohrbach-Schmidt and Hall (2020). The survey is representative of the German employed population. It contains data on task usage (self-reported by individuals on a discrete ordinal scale) for 1,235 occupations, defined at the 4-digit level (based on variable: `kldb92`). We merge this information at the occupational level with the GSOEP.

GTUS. For our empirical analysis on time use and its changes over time in Section 6, we complement the GSOEP with the German Time Use Survey (GTUS, 2023), see Merz and Ehling (1999) and Stuckemeier and Kuehnen (2013). The GTUS is administered by the German Federal Statistical Office. The survey assesses information on time use (work, leisure, home production, education, social engagements, training, etc.) and demographic characteristics of a representative sample of individuals and households, which allows us to link couples. It has three waves: 1991/92, 2001/02, 2012/13. In line with the time periods of our main sample from GSOEP, we focus on the first wave (1991/92) and the third wave (2012/13). The first wave surveys 6,400 households while the third one surveys 5,000 households.

The GTUS provides the most detailed and accurate time-use data for Germany. The data are collected using household and individual questionnaires and a time-use diary. The diary is filled out for several days: two days in 1991/92, and three days in 2012/13 (two weekdays and a weekend day). The time-use data is reported in short time intervals (5-minute intervals for 1991/92 and in 10-minute intervals for 2012/13). As our main goal is to illustrate properties of home production and their changes over time, we focus on the time-use categories related to *home production*. We harmonize the available home production categories (Childcare, House Chores, Pets, Shopping, Household Organization, Meals, Textiles, Repairs, and Care, including commuting times for each category) across waves in order to make the analysis comparable over time. We aggregate the detailed time-use data to 'hours per day'.

We impose the same demographic restrictions as in our main GSOEP sample (Online Appendix OD.2.1).

OD.2 Sample for Empirical Facts

We now describe the sample restrictions and the variables used for our empirical facts in Section 2.

OD.2.1 Sample Restrictions

For our *Main Sample*, we pool observations from the period 2010-2016, from the original GSOEP samples and their refreshments.[6] For most of the analysis, we restrict our attention to West Germany.[7]

We impose the following demographic restrictions: We keep all individuals in private households, either singles or heterosexual couples (married or cohabiting). We restrict our analysis to individuals who either never married or are in their first marriage.[8] We keep individuals who are in their prime working age, 22-55 years old.

Regarding the labor market, we exclude from our sample those individuals who are self-employed or still in school, those working in odd occupations (identified with the occupational code `kldb92`$\geq$ 9711),

[6] We exclude from our analysis the migrants and refugees samples, the oversampling of low income individuals and single parents, and the oversampling of high income earners.

[7] We drop Berlin from the sample since it cannot be unambiguously assigned to East or West Germany. This is standard in the literature, see for example Heise and Porzio (2019).

[8] Since cohabitants are defined as never married, we may include some cohabiting relationships that are not the first ones.

and those who are employed but with missing occupational code.

We impose these restrictions at the individual level. This implies that when we analyze individual outcomes, one partner of a given couple could be in the sample, while the other partner is not. For the analysis of couple outcomes, we keep couples in which both partners fulfill our sample restrictions.

OD.2.2 Variable Description

We now describe the variables we use in the empirical analysis.

1. **Education:** We classify individuals into three education levels: a) 'Low Education' includes individuals with either only a high school degree (13 years of schooling) or those with middle school degrees (9 or 10 years) plus some basic vocational training (< 11 years of schooling); b) 'Medium Education' includes those with either high school or middle school degrees and vocational training, with $>= 11$ years of schooling; c) 'High Education' includes those with a college degree or more. Education levels are defined based on the ISCED-97 classification. Alternatively, we use years of education as our schooling measure, and we left-truncate this variable at 10 years of education.
2. **Marriage Market Sorting:** For our graphical analysis at the individual level, we define marriage market sorting bins by the difference between the years of education of an individual and the years of education of their partner. For the graphical analysis at the couple level, we define marriage market sorting bins as the difference between the years of education of the male partner and the years of education of the female partner. Beyond the graphical analysis, we measure marriage market sorting as the correlation between the education levels of partners.
3. **Labor Market Sorting:** We define labor market sorting as the correlation between the individual's years of education and their matched job characteristic—the task complexity of the individual's occupation (defined in Online Appendix OE.4).
4. **Hours:**
   (a) **Labor Market Hours**: We define labor market hours as the number of self-reported hours that an individual works in a given week (including overtime). We winsorize hours to 10 and 60 hours, at the bottom and the top, respectively.
   (b) **Home Production Hours:** We measure home production hours as the weekly time an individual allocates to the following activities: childcare, housework (which includes household chores such as cooking, cleaning, etc.), running errands, repairs of the house or car, and garden work. Since home hours are measured on a typical weekday, we multiply them by five, for consistency with market hours. We impute missing data on home hours using information on labor hours, assuming a total weekly time budget of 70h for work at home and in the labor market.
   (c) **Leisure Hours:** We measure leisure hours as the weekly hours allocated to hobbies and other leisure activities (measured as daily leisure hours $\times$ 5).

OD.3 Estimation Sample

In this section, we describe the sample restrictions and variables in our estimation sample.

OD.3.1 Sample Restrictions

In order to construct our *Baseline Estimation Sample*, we use data from West Germany for the period 2010-2016 of the original GSOEP and its refreshments (as discussed above). For our *Past Estimation Sample*, we use data from West Germany for the period 1990-1996.[9] We drop those individuals that appear in both periods. We then apply the following restrictions, similar to our *Main Sample* above:

1. **Age Restrictions**: We restrict our attention to individuals between 22 and 55 years old. We keep couples in which both partners are within this age range.
2. **Marital Status Restrictions**: We focus on individuals who are either single or in heterosexual couples (married or cohabiting). We restrict our analysis to their first marital spell, as defined in Online Appendix OD.2.1. We drop observations from periods after the first marriage ended, or for which the end date of the first marriage cannot be identified. We drop individuals for whom we can identify more than one spouse/partner during the sample period.
3. **Labor Market Restrictions:**
   (a) We exclude from our sample observations corresponding to individuals working in odd occupations (`kldb92`$\geq$ 9711) or that are employed but have missing occupation codes.
   (b) We drop employed individuals with missing data on hourly wages.
   (c) We drop the self-employed and individuals who are still in school, as defined by their typical occupation (see Online Appendix OD.3.2).
4. **Additional Restrictions:**
   (a) We exclude observations from individuals to whom we cannot assign a human capital type. The estimation of human capital types is discussed in Online Appendix OE.3.
   (b) We drop individuals in couples for which information on the spouse/partner is always missing.

OD.3.2 Definition of Typical Occupation, Hours, Wages and Marital Status

In this section, we explain our methodology to create measures of individual-level labor market and demographic variables (which we refer to as 'typical outcomes'). This allows us to bridge the dynamic features of the data with the static nature of our model.

Typical Occupation: We use the following rules to assign a typical occupation to an individual.

1. If they appear in the sample only once, or if they appear more than once but always report the same occupation, we assign to them that unique occupation.
2. If they appear in the sample more than once, and in at least one of those years they are in a *'non-labor market state'* (i.e., 'self-employed', 'studying' or 'not-employed'), we proceed as follows:

[9] We drop individuals observed both in West and East Germany while in their typical occupation, see below.

(a) If they were either 'not-employed' or 'self-employed' or 'studying' for strictly more than half of the time they appear in the sample, we consider that state as their typical occupation.

(b) If they were in one of these states (e.g., 'self-employed') exactly half of the time they appear in the sample, and spent the other half in the other two states (e.g., 'studying' and 'not-employed'), we assign them to the state in which they spent half the time ('self-employed').

(c) If (a) and (b) do not hold, but they spent more than 75% of their time in the 'non-labor market states' combined, then we assign them to the state with the longest duration.

3. If we observe an individual multiple times, but only in a single occupation (and they hold this occupation for more than 25% of the time), we assign them to that unique occupation.

4. If we observe an individual in more than one occupation during the sample period, we construct the difference in percentiles of task complexity between their highest and lowest-ranked occupation (where occupations are ranked as described in Online Appendix OE.4). We then proceed as follows:

   (a) When the difference in percentiles is ≥ 0.1, we assign them to their highest ranked occupation.

   (b) If the difference is < 0.1, we assign them to the occupation with the longest tenure. If there is a tie, we assign them to the highest ranked occupation among those with equal tenure.

5. After applying these rules, we drop from the sample individuals whose typical occupation is 'self-employment' or 'studying'.

Typical Labor Market Hours: We define typical labor market hours as the average self-reported work hours (including overtime) over the years an individual was in their typical occupation, defined above. We winsorize labor hours to 10 and 60 hours, at the bottom and the top, respectively. For those individuals whose typical occupation is 'not employed', typical labor hours are set to zero.

Typical Hourly Wage: For each individual, we define their typical wage as the average real hourly wage over the years they worked in their typical occupation.[10]

Typical Home Production Hours: We construct typical home production hours (defined in Online Appendix OD.2) as the average home production hours over the years an individual was in their typical occupation. We impute missing data for home hours as discussed in Online Appendix OD.2.

Typical Marital Status: We define the typical marital status based on the following rules:

1. If the individual had only one marital status during the sample period, we consider that marital status as the typical one.
2. If the individual switched from being single to being married during the sample period, we assign the marital status observed when employed in their typical occupation. If they were observed as both single and married while in their typical occupation, we assign them to marriage.
3. We exclude from the sample individuals that report more than one spouse during the sample period. We also exclude their partners.

[10] We construct hourly wages based on inflation adjusted monthly earnings, divided by monthly hours (constructed as weekly hours times 4.3). Hourly wages are trimmed at the bottom and top 1% percentile. The data for inflation adjustment comes from the OECD: https://data.oecd.org/price/inflation-cpi.htm

Typical Child: In Section 6/Online Appendix OG.4, we re-estimate our model on two different subsamples: those with and those without children. To assign individuals to either subsample, we define their 'typical child' status based on the presence of children under 18 years old in the household, during at least one of the years in which they are in their typical occupation. If there are discrepancies between husband and wife (4% of cases), we classify households based on the 'typical child' status of the male partner. For singles, we randomly distribute them to either subsample so that the share of singles in each subsample equals the share of singles in the whole sample.

Following these rules, our *Baseline Estimation Sample* (West Germany, 2010-2016) has 3,857 individuals living in 2,326 households. Of these households, 1,531 are couples and 795 are singles (418 are single women and 377 are single men).

Our *Past Estimation Sample* (West Germany, 1990-1996) consists of 2,336 individuals in 1,294 households, of which 1,042 are couples and 252 are singles (117 single women and 135 single men).

OE Estimation

OE.1 Construction of Data Moments

Our main estimation targets 16 moments defined in Table O.10. In this section, we provide details on how we construct these moments in the data and in the model.

For moment M1, we compute the female to male labor force participation ratio, including both married and single individuals. Both in the data and in the model, we define labor force participation as a dichotomous variable that takes value 1 when an individual works positive hours, and zero otherwise.

For moment M2, we compute the ratio of female full-time workers to male full-time workers, including both married and single individuals. Both in the data and in the model, we define 'full-time' as working more than 44% of the available time. This is equivalent to working more than 37.5 hours per week in the data, and to working more hours than described by the fifth entry in the hours grid in the model.[11]

For M3 and M4, we compute the married to single ratio in labor force participation, separately for women and men.

We compute M5 as the correlation between the time that female and male partners spend in home production, where home production hours are constructed as the share of total available time spent in home production.[12] We restrict our attention to individuals in couples.

To construct wage moments M6-M9, we use data on hourly wages for all employed individuals, whether they are single or in a couple. See Table O.10 for more details.

[11] In the data, we determine the total available time by the 95th percentile of the distribution of total time spent working (home production plus labor market work). This is 85 weekly hours for the period 2010-2016.

[12] We define the total available time as in footnote 11 to deal with the fact that different individuals report different total hours allocated to home production and market work. When individuals report more than 85 hours of home production per week, we say that their share of time allocated to home production equals 1.

Moment M10 is constructed as the correlation of partners' human capital types, i.e., of their s-types (see Online Appendix OE.3 for the details on the human capital estimation).

Moments M11 and M12 measure the gender wage gap by (s, h)-types. We use two (s, h)-type combinations: all individuals of either s-type 3 or s-type 4 (see columns 3 and 4 of Table O.13) that work full-time in the labor market.

For M13-M14, we compute the labor force participation rate of women in couples where both partners have a similar human capital type. For M13, we pool couples in which both partners are either of s-type 3 or 4 (corresponding to columns 3 and 4 in Table O.13). For M14, we pool couples in which both partners are either of s-type 5 or 6 (corresponding to columns 5 and 6 in Table O.13).

Finally, for M15 and M16, we compute the labor force participation rate among single women. M15 considers single women of s-types 3 or 4. M16 pools single women of s-types 5 or 6.

Table O.10: Moments

Moment Description	**Definition**
Labor Force Participation Female to Male Ratio (M1)	$\frac{\mathbb{P}(h_f>0)}{\mathbb{P}(h_m>0)}$
Full-time Work Female to Male Ratio (M2)	$\frac{\mathbb{P}(h_f=\hat{h})}{\mathbb{P}(h_m=\hat{h})}, \hat{h} \geq 37.5$
Labor Force Participation Married to Single Ratio, by Gender (M3-M4)	$\frac{\mathbb{P}(h_i>0\|Married)}{\mathbb{P}(h_i>0\|Single)}, i \in \{f, m\}$
Correlation of Spouses' Home Production Hours (M5)	$corr(1-h_f, 1-h_m)$
Mean Hourly Wage (M6)	$\mathbb{E}[w]$
Variance of Hourly Wage (M7)	$Var[w]$
Overall (90-10) Wage Inequality (M8)	$\frac{w^{90}}{w^{10}}$
Upper Tail (90-50) Wage Inequality (M9)	$\frac{w^{90}}{w^{50}}$
Correlation between Spouses' Human Capital Types (M10)	$corr(s_m, s_f)$
Gender Wage Gap by Effective Type (M11-M12)	$\frac{\mathbb{E}[w(h_i s_i)\|i=m,h_i=\hat{h},s_i=\hat{s}]-\mathbb{E}[w(s_i h_i)\|i=f,h_i=\hat{h},s_i=\hat{s}]}{\mathbb{E}[w(h_i s_i)\|i=m,h_i=\hat{h},s_i=\hat{s}]}$
Female Labor Force Participation by Couples' Human Capital Type (M13-M14)	$\mathbb{P}(h_f > 0\|s_f = s_m = \hat{s})$
Labor Force Participation of Single Women by Human Capital Type (M15-M16)	$\mathbb{P}(h_f > 0\|Single, s_f = \hat{s})$

OE.2 Simulated Methods of Moments

To estimate the model parameters $\Lambda = (\theta, \rho, A_p, \gamma_1, \gamma_2, A_z, \sigma_\delta, \sigma_\beta, \psi)$, we apply the method of simulated moments (McFadden, 1989; Pakes and Pollard, 1989). For any vector of parameters, Λ, the model produces the 16 moments, $mom_{sim}(\Lambda)$, that will also be computed in the data, mom_{data}. We first use a global search and then a local algorithm to find the parameter values that minimize the distance between simulated and observed moments. Formally, the vector $\hat{\Lambda}$ solves

$$\hat{\Lambda} = \arg\min_{\Lambda} \quad [mom_{sim}(\Lambda) - mom_{data}]' \mathcal{V} [mom_{sim}(\Lambda) - mom_{data}],$$

where $\mathcal{V}$ is the inverse of the diagonal of the covariance matrix of the data moments.

OE.3 Estimation of Worker Types

OE.3.1 Sample Selection, Methodology and Key Variables

Sample Selection. Our sample consists of individuals in the GSOEP from 1984-2018, who are between 20 and 60 years old and are either married/cohabiting or single. We exclude individual-year observations corresponding to self-employment or work in odd occupations (`kldb92`$\geq$ 9711). We also exclude observations with missing information on education or with missing (not zero) labor force experience. Our panel consists of around 212,000 person-year observations.

Methodology. We aim to estimate the distribution of workers' human capital N_s. To do so, we need to estimate the returns to education to obtain a proxy for x in our model, as well as individual unobserved heterogeneity, capturing ν in our model. We use our theoretical wage function as guidance for specifying the empirical log hourly wage of individual i at time t (where t is a year in our sample) as

$$\ln w_{it} = \nu_i + \sum_{ed \in \{med, high\}} \alpha^{ed} x_{it}^{ed} + \beta_1 h_{it} + \beta_2 h_{it}^2 + \beta_z' Z_{it} + \delta_s + \delta_t + \epsilon_{it}, \tag{O.10}$$

where x_{it}^{ed} are indicator variables for the education group of an individual (capturing x in our model). Coefficient α^{ed} gives the return to education $ed \in \{med, high\}$ in terms of log wage units, where $0 < \alpha^{med} < \alpha^{high}$ would indicate positive returns to education. In turn, ν_i is an individual fixed effect that captures unobserved time-invariant ability, with model counterpart ν. Further, h_{it} denotes weekly labor hours (capturing h in our model). Finally, Z_{it} are time-varying controls for the individual, δ_s and δ_t are state and time fixed effects, and ϵ_{it} is a mean-zero error term.[13] We implement (O.10) as a panel wage regression. Below we explain how we deal with the following three challenges in the estimation: selection into labor force participation, endogeneity of hours worked, and agents who hold jobs for less than two years in our sample.

[13] We do not include occupation fixed effects since in our model, conditional on $\tilde{s}$ (which we control for by including regressors (x, ν, h)), the wage does not depend on occupation y in our competitive equilibrium. But even doing so—which we have done for robustness—does not significantly change the impact of x or ν on the hourly wage.

Key Variables. For *weekly hours*, we use reported actual hours, see Online Appendix OD.2.2. For *labor force experience*, we use the reported labor force experience and we impute it by potential experience if this information is missing. The three *education* groups (low, medium, high) are defined in Online Appendix OD.2.2. Our wage variable is *log hourly wages*, inflation-adjusted in terms of 2016 Euros. For the definition of 'demographic cells' in the selection stage, we additionally use the presence of *children below 3 years old* in the household, *age bins* (≤ 25, > 25 and ≤ 40, > 40 and ≤ 50, > 50) and *state* of residence.

OE.3.2 Selection Equation

To account for selection into labor force participation in the wage regression, we first run a selection regression. To do so, we need an instrument that affects participation but is excluded from wage regression (O.10). Since the variation in participation in our sample is mainly driven by women, we use the 'progressiveness' in an individual's narrowly defined demographic cell, proxied by the share of females working in that cell. Our cells are defined by a combination of state of residency, year, age bin, and an indicator of whether a child under the age of 3 is in the household.[14] This strategy builds on the literature on grouping instruments (Angrist, 1991 and Blundell, Duncan, and Meghir, 1998) which proposes to assign individuals to groups and use group-specific averages to instrument for endogenous variables at the individual level.[15] When defining this variable for a particular individual, we employ the 'leave-one-out' method to avoid counting the individual's labor force participation when computing this statistic. We further drop cells with less than five observations, yielding around 2,500 cells in total. Our assumption on this IV (i.e., our progressiveness variable) is that the following exclusion restriction holds: 'progressiveness of an individual's demographic cell' only affects wages through labor force participation but not in other ways.

While in our empirical selection equation, labor force participation is driven by the progressiveness of the demographic cell, in the model this selection is determined by labor supply shocks. We argue, however, that these sources of selection are related. Low realizations of the labor supply shocks that lead to non-participation in the model partly reflect obstacles to the labor force participation of women. These hurdles are likely specific to locations and demographic groups, and can stem from institutional factors (e.g., the low regional availability of childcare slots); or from supply side forces (e.g., low availability of immigrant workers, as in Cortés and Tessada, 2011); or from conservative norms and culture prevailing in a location and therefore among the peers whose opinions affect own labor force participation choices (see Fernández and Fogli, 2009, Maurin and Moschion, 2009, Fogli and Veldkamp, 2011,

[14] German children start kindergarden at age 3. Before that, they are mostly at home (in 2013, only 29% of children aged 0-2 were in daycare (OECD, 2016)), so age 3 is an important threshold affecting maternal labor force participation.

[15] The idea is that group-specific means capture the aggregate/institutional *common* environment a group faces, which explains members' behavior but is uncorrelated to unobserved idiosyncratic errors. In light of this literature, our cell definition is closest to that in Kalíšková (2020) who studies the impact of tax benefits on female labor supply. To instrument for individual participation tax benefits she employs the average tax benefits when out of work, by education, presence of children by age, marital status, country and year. Relatedly, Burns and Ziliak (2016) identify the impact of marginal tax rates on taxable income, using cohort-education-state average tax rates to instrument for person-specific tax rates.

Mota, Patacchini, and Rosenthal, 2016, Fortin, 2015, Olivetti, Patacchini, and Zenou, 2020). To the extent that these barriers affect labor supply choices of most women in a certain location/demographic group, our IV is a good proxy for the labor supply shocks in the model. Specifically, we think of the individual's labor supply shock in our model as being positively correlated with our progressiveness IV in the data, bridging selection in model and data.

We run the following probit selection regression:

$$emp_{it} = \alpha_0 + \gamma \; share_{jit} + \sum_{ed \in \{med, high\}} \alpha^{ed} x_{it}^{ed} + \beta_z' Z_{it} + \delta_s + \delta_t + \epsilon_{it}, \tag{O.11}$$

where the dependent variable indicates whether individual i is employed at time t. Variable $share_{jit}$ is the progressiveness measure in the demographic cell j of individual i at time t (given by the share of women working in the cell, see above), x_{it}^{ed} is an education indicator ('low education' is the reference group), Z_{it} is a vector of individual controls (linear and quadratic labor force experience in years, household size). Finally, δ_s and δ_t are state and year fixed effects and ϵ_{it} is a mean-zero error term.

The results are in Table O.11, where we label the variable $share_{jit}$ by *Share of Working Women in Cell*. We find a strong positive effect of the share of women working in a demographic cell on the labor force participation of an individual in that cell.

Table O.11: Selection Regression

	Employed
Share of Working Women in Cell	1.355***
	(0.028)
Experience	0.092***
	(0.002)
Experience^2	-0.002***
	(0.000)
Medium Educ	0.453***
	(0.009)
High Educ	0.851***
	(0.011)
HH Size	-0.046***
	(0.005)
Constant	-0.918***
	(0.036)
State and Year FE	Yes
Observations	212,894
Pseudo R^2	0.144

Notes: Standard errors clustered at cell level in parentheses. ***p<0.01, **p<0.05, *p<0.1. Results are based on regression (O.11). A *Cell* is defined by state, age, year and presence of children under 3 in the household. Source: GSOEP.

OE.3.3 Panel Wage Regression

One potential issue for our specification of the wage regression is that there may be unobserved confounding factors that impact both hours and wages. While we deal with time-invariant unobserved heterogeneity using the panel regression with individual fixed effects, time-varying unobserved heterogeneity (see Section 2 for a discussion) could still be problematic. To address this concern, we instrument own labor hours by partner's labor hours in regression (O.10). More specifically, we instrument own labor hours and labor hours squared by (i) labor hours of the partner, (ii) labor hours of the partner squared and (iii) whether a partner is present (where we set the partner's hours to zero in two cases, either if the partner is present but not working or if no partner is present).

Through the lens of our model this instrument is valid: First, there is a systematic relationship between the hours worked of an individual and the hours worked by their partner. Second, the partner's labor hours impact own wage *only* through own labor hours.

We drop observations for individuals whose partner reports to be employed but has zero reported labor hours, or whose partner has missing employment and hours information. Since we include individual fixed effects we also drop observations for individuals who are in our panel for only a single year. We further restrict the sample to those who are employed and have non-missing hourly wages. This leads to a sample of 133,214 person-year observations. Based on our model wage function, we choose the following regression specification:

$$\ln w_{it} = \nu_i + \sum_{ed \in \{med, high\}} \alpha^{ed} x_{it}^{ed} + \beta_0 IMR_{it} + \beta_1 h_{it} + \beta_2 h_{it}^2 + \beta_z' Z_{it} + \delta_s + \delta_t + \epsilon_{it} \tag{O.12}$$

where ν_i is a individual fixed effect, IMR_{it} is the inverse mills ratio of individual i in year t from the selection probit regression (O.11), h_{it} are weekly hours worked, and x_{it}^{ed}, Z_{it}, δ_s and δ_t are as in (O.11). Note that we could not include h_{it} into selection equation (O.11) since $h_{it} = 0$ versus $h_{it} > 0$ is a perfect predictor of employment. Nevertheless, based on our model, it is important to control for hours worked in the hourly wage regression (O.12).

Table O.12 contains the results. Columns (1) and (2) report the first stage regressions for two variables to be instrumented: weekly hours and weekly hours squared. Column (3) contains the results of the second stage regression. Based on the F-statistics, the three IV's for the hours worked variables (partner's hours, partner's hours squared and partner present) are not subject to the weak instrument problem. Regarding the second stage, we note that the inverse mills ratio is positive and significant, indicating that individuals are positively selected into working and not controlling for selection here would have biased the coefficients upward. Moreover, we note that weekly hours worked have a positive effect on hourly wages, justifying our model assumption that hours affect productivity and thus hourly wages. In particular, increasing hours from 30 to 40 hours per week yields an hourly wage return of around 4%, computed from the estimated coefficients in Column (3) as $(40 \times 0.119 - 40^2 \times 0.00164) - (30 \times 0.119 - 30^2 \times 0.00164) = 0.042$.

OE.3.4 Imputation

Based on these results, we are able to obtain x-types and ν-types (and thus s-types) for around 17,000 individuals. We impute fixed effects of the remaining ones (around 11,600 individuals) based on the *multiple imputation* approach. As auxiliary variables in this imputation we choose covariates that are most correlated with the individual fixed effects (such as *education*, *gender* and *full time labor force participation*). After imposing the sample restrictions described in Online Appendix OD.3.1, our final estimation sample (for period, 2010-2016) has 3,857 unique individuals, 24% of them have imputed ν_i.

Table O.12: Wage Regression

	(1) Weekly Hours Worked	(2) Weekly Hours Worked^2	(3) Log Hourly Wage
Partner's Weekly Hours Worked	-0.0574***	-4.6070***	
	(0.0048)	(0.3894)	
Partner's Weekly Hours Worked^2	0.0009***	0.0796***	
	(0.0001)	(0.0086)	
Partner Present	1.1326***	60.8468***	
	(0.1638)	(12.5684)	
Experience	0.3765***	28.5312***	0.0332***
	(0.0251)	(1.7180)	(0.0018)
Experience^2	-0.0046***	-0.3615***	-0.0006***
	(0.0004)	(0.0298)	(0.0000)
Medium Educ	0.2424	27.2826*	0.0264**
	(0.2157)	(15.8949)	(0.0103)
High Educ	4.2893***	299.7916***	0.1954***
	(0.4577)	(30.9288)	(0.0244)
HH Size	-1.2799***	-80.3213***	0.0218***
	(0.0510)	(3.4953)	(0.0047)
Inverse Mills Ratio	-3.5185***	-215.1606***	0.1334***
	(0.2547)	(18.3766)	(0.0197)
Weekly Hours Worked			0.1192***
			(0.0156)
Weekly Hours Worked^2			-0.0016***
			(0.0002)
State and Year FE	Yes	Yes	Yes
Observations	133,214	133,214	133,214
F	69.918	59.519	162.902
R^2			-0.386

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the *Cell* level are in parentheses, where *Cell* is defined by state, year, age, and presence of children under the age of 3 in the household. The *Inverse Mills Ratio* is defined for each individual i in year t and based on the selection probit regression (O.11). Column (3) reports the IV version of regression (O.12), where we instrument *Weekly Hours Worked* and *Weekly Hours Worked*^{2} by the partner's weekly hours worked (linear and squared) and whether the partner is present. The results of the first stage are in columns (1) and (2). Source: GSOEP.

OE.3.5 Distribution of Workers' Human Capital

We then divide individuals into our three education bins and assess within each bin whether an individual has a low (below median) or high (above median) fixed effect, so there are two subgroups $j \in \{1,2\}$ in each education bin. We compute the subgroup fixed effect, $\overline{\nu}_j^{ed}$, as the mean of the individual fixed effects in subgroup j of education bin ed. Finally, we compute the human capital type for each individual i as $s_i = \alpha^{ed} x_i^{ed} + \overline{\nu}_{ji}^{ed}$ where $\overline{\nu}_{ji}^{ed}$ is the mean fixed effect of individual i's group j. We obtain six s-types. The resulting human capital distribution (s-types by education group) is displayed in Table O.13.

Table O.13: Worker Distribution of s-Types by Education

	s-type (ed, $\overline{\nu}^{ed}$)						
Education	(low,low)	(med,low)	(low,high)	(med,high)	(high,low)	(high,high)	Total
Low Educ	523	0	601	0	0	0	1,124
Medium Educ	0	617	0	1,046	0	0	1,663
High Educ	0	0	0	0	447	623	1,070
Total	523	617	601	1,046	447	623	3,857

Notes: The education groups (*Low, Medium, High*) are defined in Online Appendix OD.2.2. The table displays six s-types, ordered by their value of $s = \alpha^{ed} x^{ed} + \overline{\nu}_j^{ed}$ for $j \in \{1,2\}$ and $ed \in \{low, med, high\}$. Parameters α^{ed} are estimated in regression (O.12) and $\overline{\nu}_j^{ed}$ is the mean of individual fixed effects in education group ed and subgroup j, also based on (O.12). Source: GSOEP.

OE.4 Estimation of Job Types

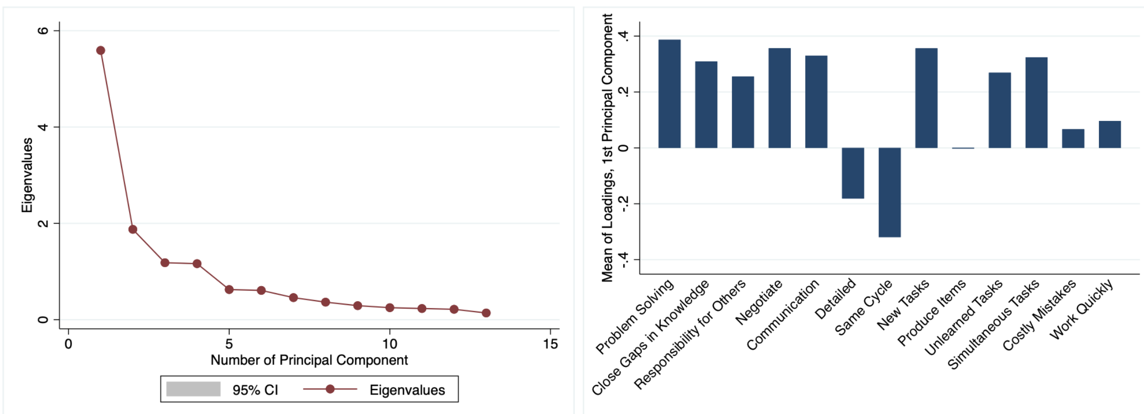
Sample Selection. Our main data source for measuring occupation types is the BIBB/BAuA Employment Survey 2012 (see Online Appendix OD.1 for details). In order to reduce the problem of noisy reporting, we drop occupations for which task information is based on less than 5 individuals. We are left with task data for 613 occupations. Our structural estimation exercise will be based on the subset (608 occupations) that can be merged to the occupations held by individuals in our GSOEP sample, using the four-digit occupational code `kldb92` from the German Classification of Occupations 1992.

Task Data. The BIBB contains data on how intensely different tasks are performed at each of our 4-digit occupations. The set of tasks is comprised of: Detailed Work, Same Cycle, New Tasks, Improve Process, Produce Items, Tasks not Learned, Simultaneous Tasks, Consequence of Mistakes, Reach Limits, Work Quickly, Problem Solving, Difficult Decisions, Close Gaps of Knowledge, Responsibility for Others, Negotiate, Communicate. See Online Appendix OD.1 for details.

Model Selection Stage. We merge the task data from the BIBB into occupations held by individuals in the GSOEP. As with the estimation of worker types, we here use the entire GSOEP panel (here: pooled). We run a Lasso regression of log hourly wages on the task descriptors listed in the last paragraph in order to systematically select the tasks that matter for pay. This procedure selects 13 tasks (all tasks except: Improve Process, Consequence of Mistakes and Difficult Decisions).

Principal Component Analysis (PCA). To reduce the occupational type to a single dimension, we collapse the information of the 13 selected tasks into a single measure using a standard dimension reduction technique (PCA). We then select the first principal component, which captures the most variation of the underlying task variables in the sample of employed workers in the GSOEP (43%). Based on the loadings on the underlying task descriptors, we interpret this component as *task complexity*: All task variables have positive loadings except *detailed work* and *same cycle*, which arguably are the only tasks in the dataset that indicate routine work. Figure O.7 shows a scree-plot with eigenvalues of the different principal components (left) and the loadings of the first principal component (right).

Figure O.7: Principal Component Analysis



Distribution of Jobs' Task Complexity. We then compute the mean of this first principal component by occupation and denote it by $\hat{y}$. Once matched to our *Main Sample* (see Online Appendix OD.3), we define our final measure of occupational or job type as the *rank* of the occupation in the task complexity distribution, i.e. $y = \widehat{G}(\hat{y})$, where $\widehat{G}$ is the cdf of $\hat{y}$. So $y \sim G$ where $G = U[0, 1]$. We use this transformation because the occupational task data has an ordinal interpretation. Examples of occupations in the top 5% of the G distribution include engineers and programmers. Examples of occupations in the bottom 5% include janitors and cleaners.

OF Quantitative Analysis: The Key Drivers of Inequality

To highlight the main forces behind inequality, we conduct *comparative statics* exercises in our estimated model with respect to the key drivers of inequality (ρ, θ, ψ).

The Effect of ρ. We are interested in the effects on inequality when ρ becomes even more negative (i.e., if home hours become even more complementary).

Figure O.8, first row, plots the effect of ρ on different inequality measures: gender wage gap (panel a), within- and between-household income inequality (panel b), and overall household income inequality (panel c). It shows that a decline in ρ (moving from the right to the left on the x-axis) *decreases* the gender gap significantly. Starting from our estimate $\rho = -0.54$ and decreasing this parameter to -2 decreases the gender wage gap by almost 13%. This is due to a direct effect of complementarities on hours and several indirect effects through sorting: First, as the complementarity in home production between partners increases, complementarity between spouses' types in the marriage value becomes stronger, resulting in more positive assortative matching (panel d). Both increased marriage sorting (indirectly) and stronger complementarities in home production (directly) induce spouses to better align their hours. Women increase their labor hours while men decrease theirs, leading to a smaller gender gap in labor hours (panel e) and thus productivity, which puts downward pressure on the wage gap. Moreover, because women 'improve' a sorting-relevant attribute (labor hours) relative to men, the gender gap in labor market sorting declines (panel f), reducing the gender wage gap even further.

Figure O.8(c) then shows that overall income inequality *declines* with stronger complementarities. This decline is driven by the decrease in within-household inequality (mirroring the decline in the gender wage gap), which dominates the increase in between-inequality stemming from stronger marriage sorting.

THE EFFECT OF θ. Eliminating the gender gap in home productivity (reducing θ from the estimate $\theta = 0.82$ to $\theta = 0.5$) would cut the gender wage gap by almost half, see Figure O.9(a): Making home productivity more equal increases the incentive for positive marriage sorting (panel d) and pushes against household specialization with a smaller gender gap in labor hours (panel e). This positively affects women's wages directly and also indirectly, through a smaller labor market sorting gap (panel f). Interestingly, overall household income inequality decreases as men and women become similarly productive at home, Figure O.9(c). Here, this is driven by a decrease in within-household inequality, which dominates the increase in between-household inequality driven by a rise in marriage sorting (Figure O.9(b)).

THE EFFECT OF ψ. Last, Figure O.10(a) shows that eliminating the female labor market wedge ψ (i.e., increasing ψ from our estimate $\psi = 0.85$ to $\psi = 1$) would reduce the gender gap by about 25%. There is a direct positive effect of ψ on female productivity—and thus wages—but also several indirect effects: First, the wife's labor hours increase in productivity ψ relative to the husband's, reducing the gender hours gap (panel e) and thus the gender wage gap. Second, the reduction in the gender hours gap leads to a decline in the labor market sorting gap (panel f), further curbing the gender wage gap. Third, smaller gender disparities in the labor market are associated with an increase in marriage market sorting (panel d), since in a world in which men and women are more equal, the motive for positive sorting strengthens. The increase in marriage sorting reinforces the drop in both hours and labor sorting gaps, further dampening the gender wage gap. Finally, Figures O.10(b) and (c) show that an increase in ψ leads to lower within-household inequality, but higher between-household inequality.

In sum, we derive several insights: First, eliminating asymmetries in productivity across gender (whether at home through $\theta \to 0.5$ or at work through $\psi \to 1$) reduces the gender wage gap. But this is not the only way to mitigate gender disparities: An increase in home production complementarity (i.e., a decrease in ρ, the key parameter of our model in shaping equilibrium) has qualitatively similar effects. Second, a decline in the gender wage gap tends to go hand in hand with a decline in gender gaps in labor hours and labor market sorting and with an increase in marriage market sorting. Third, while the effect of these parameters on overall income inequality depends on the exercise, in all cases the gender wage gap comoves positively with within-household inequality but negatively with between-household inequality.

Figure O.8: a. Gender Wage Gap, b. Household Income Variance Decomposition, c. Household Income Variance, d. Marriage Market Sorting, e. Gender Gap Labor Hours, f. Gender Gap Labor Sorting.

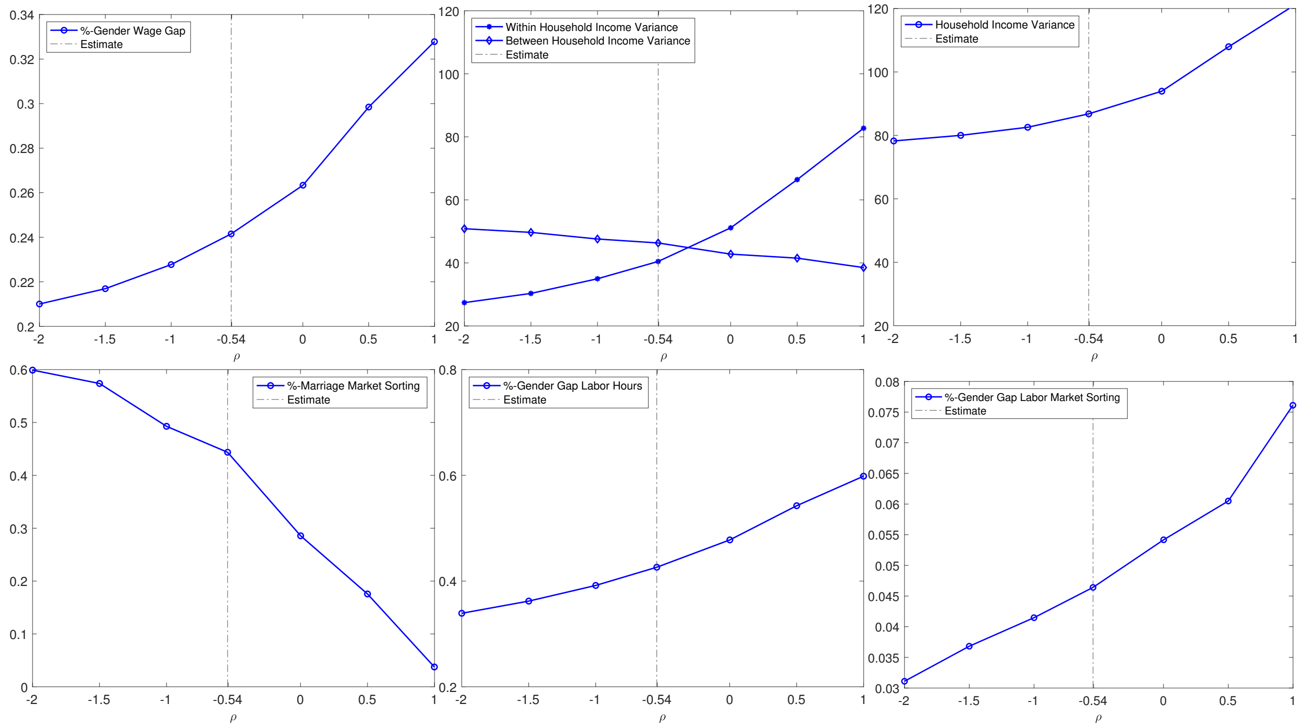


Figure O.9: a. Gender Wage Gap, b. Household Income Variance Decomposition, c. Household Income Variance, d. Marriage Market Sorting, e. Gender Gap Labor Hours, f. Gender Gap Labor Sorting.

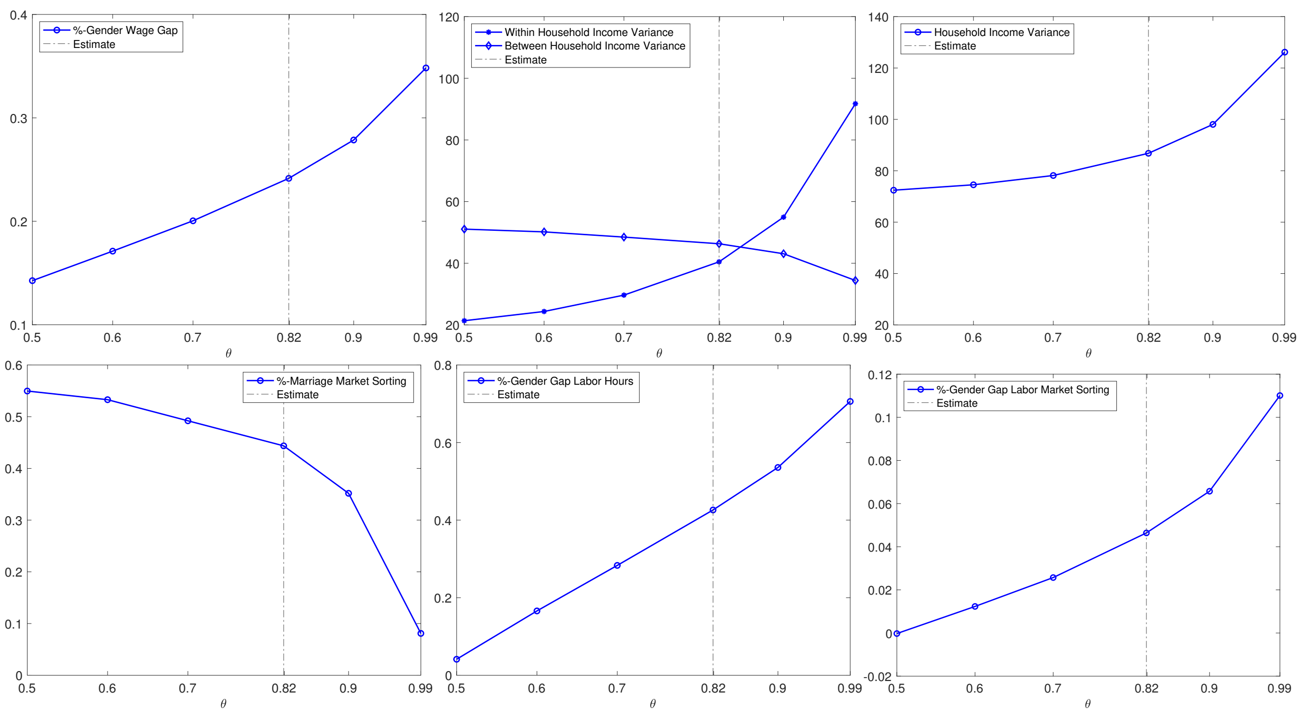
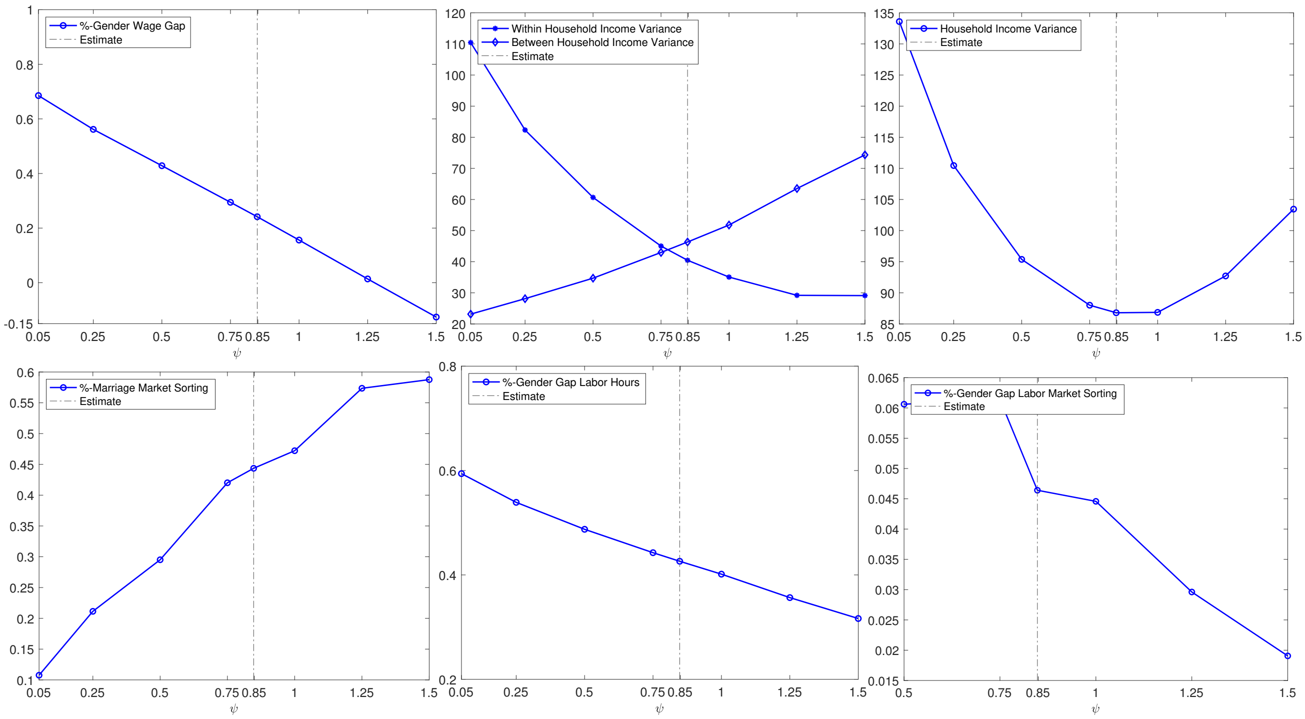


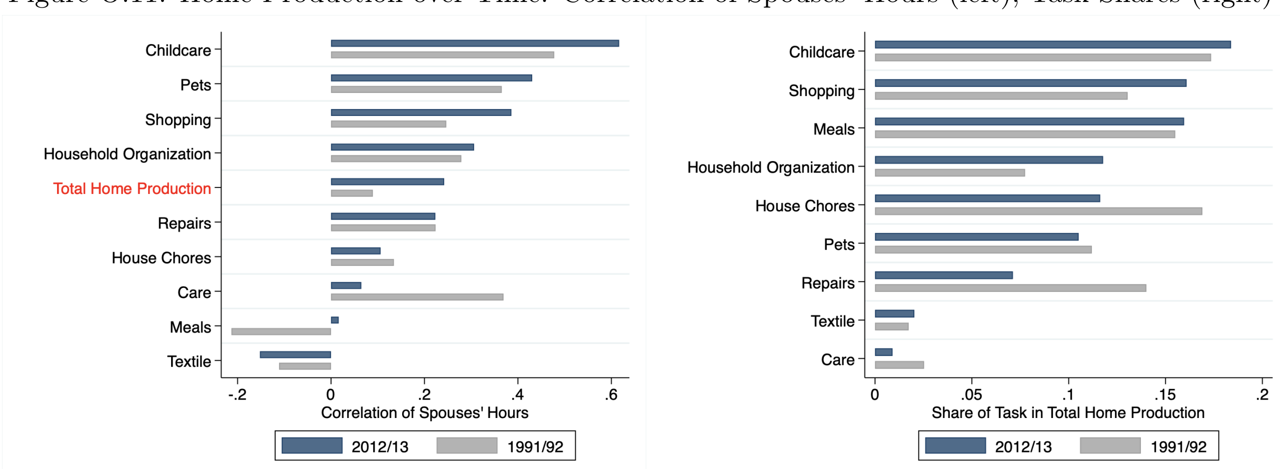
Figure O.10: a. Gender Wage Gap, b. Household Income Variance Decomposition, c. Household Income Variance, d. Marriage Market Sorting, e. Gender Gap Labor Hours, f. Gender Gap Labor Sorting.



OG The Sources of Home Production Complementarities

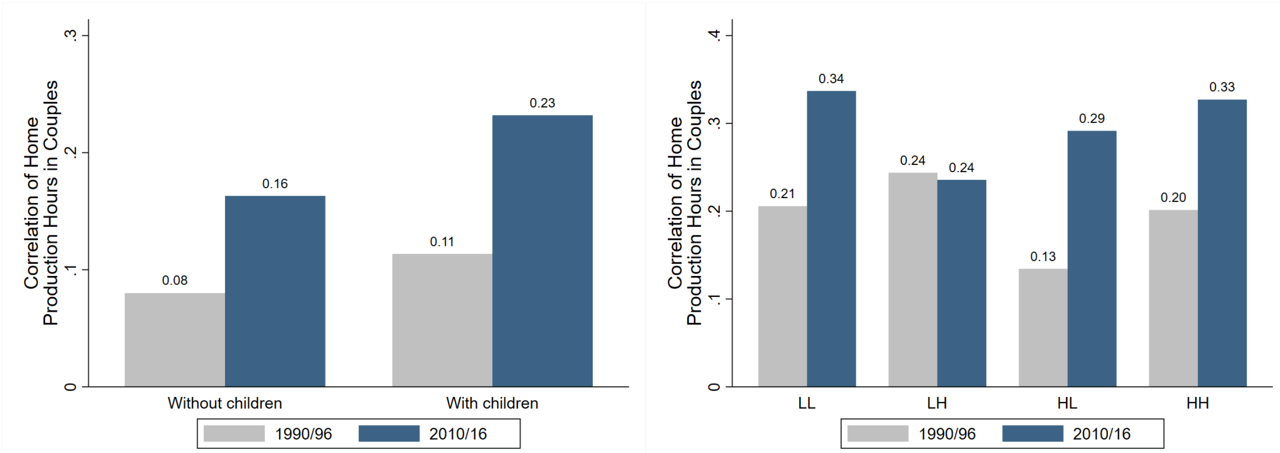
OG.1 Motivating Evidence

Figure O.11: Home Production over Time: Correlation of Spouses' Hours (left); Task Shares (right)



Notes: The left panel reports the correlation of spouses' hours by detailed home production tasks. The right panel reports the share of each task (in terms of time) in overall home production. Source: GTUS.

Figure O.12: Correlation of Spouses' Hours by Couple Type: Children (left); Education (right)



Notes: In the left panel, couples with children are defined as those who have at least one child under 18 years old in the household. In the right panel, we define four types of couples, based on the education of the male and the female partners. 'L' stands for low-education (which pools all individuals with less than a college degree) and 'H' represents high education (including all individuals with a college degree or more). The first type entry refers to the husband while the second one refers to the wife. Source: GSOEP.

OG.2 Statistical Decomposition: By Detailed Home Production Tasks

Cross-Section. Equation (8) in Section 6.1 decomposes the aggregate correlation of home production time between spouses into the weighted sum of correlations of $n = 9$ detailed home production tasks. In the GTUS, these nine tasks are: Childcare, House Chores, Pets, Shopping, Household Organization, Meals, Textiles, Repairs and Care.

Over Time. Based on (8), we can further decompose the over-time *changes* in the aggregate home production correlation, using a shift-share decomposition:

$$\begin{aligned}\Delta_{t,t-1}\,\mathrm{Corr}\,(\ell_f,\ell_m) &\equiv \mathrm{Corr}\,(\ell_f,\ell_m)_t - \mathrm{Corr}\,(\ell_f,\ell_m)_{t-1} \\ &= \sum_{j=1}^{n}\sum_{k=1}^{n} \underbrace{\overline{\mathrm{Corr}\,(\ell_{fj},\ell_{mk})}(V_{jk,t}-V_{jk,t-1})}_{\text{Change in Task Weights}} + \underbrace{\overline{V}_{jk}(\mathrm{Corr}\,(\ell_{fj},\ell_{mk})_t - \mathrm{Corr}\,(\ell_{fj},\ell_{mk})_{t-1})}_{\text{Change in Hours Correlation Within and Across Tasks}}, \qquad \text{(O.13)}\end{aligned}$$

where a 'bar' indicates the arithmetic average of a variable over the two time periods.[16]

Results. Implementing the over-time decomposition, we obtain the following results.

Table O.14: Decomposing Changes in the Aggregate Home Production Correlation (Tasks)

Overall Change	Weights (%)	Correlations(%)
0.15	27.8	72.2

Notes: This decomposition is based on (O.13). Column 1 gives $\Delta_{t,t-1}\,\mathrm{Corr}\,(\ell_f,\ell_m)$, while columns 2 and 3 report the first and the second terms on the RHS of (O.13). Source: GTUS.

[16] That is,

$$\overline{\mathrm{Corr}\,(\ell_{fj},\ell_{mk})} = \frac{1}{2}(\mathrm{Corr}\,(\ell_{fj},\ell_{mk})_t + \mathrm{Corr}\,(\ell_{fj},\ell_{mk})_{t-1}) \quad \text{and} \quad \overline{V}_{jk} = \frac{1}{2}\big(V_{jk,t}+V_{jk,t-1}\big).$$

OG.3 Statistical Decomposition: By Couple Types

Cross-Section. Let $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, ..., \mathcal{G}_r\}$ be the set of couple types (e.g., for the groups 'couples with children' and 'couples without children' we have $r = 2$ and so $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2\}$). Interpreting $\mathcal{G}$ as a random variable, the Law of Total Covariance gives:

$$\operatorname{Corr}(\ell_f, \ell_m) = \frac{\operatorname{Cov}(\ell_f, \ell_m)}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}} = \frac{\mathbb{E}[\operatorname{Cov}(\ell_f, \ell_m|\mathcal{G})] + \operatorname{Cov}(\mathbb{E}[\ell_f|\mathcal{G}], \mathbb{E}[\ell_m|\mathcal{G}])}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}}.$$

Spelling this out, we obtain:

$$\begin{aligned}\operatorname{Corr}(\ell_f, \ell_m) = &\left(\sum_{i=1}^{r} \mathbb{P}(\mathcal{G}_i)\operatorname{Corr}(\ell_f, \ell_m|\mathcal{G}_i)\frac{\sqrt{\operatorname{Var}(\ell_f|\mathcal{G}_i)\operatorname{Var}(\ell_m|\mathcal{G}_i)}}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}}\right) \\ &+ \operatorname{Corr}(\mathbb{E}[\ell_f|\mathcal{G}], \mathbb{E}[\ell_m|\mathcal{G}])\frac{\sqrt{\operatorname{Var}(\mathbb{E}[\ell_f|\mathcal{G}])\operatorname{Var}(\mathbb{E}[\ell_m|\mathcal{G}])}}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}},\end{aligned}$$

where $\mathbb{P}(\mathcal{G}_i)$ is the proportion of couple type i in the population and where we will denote the 'composite weight' of each type i by:

$$\mathcal{V}_i := \mathbb{P}(\mathcal{G}_i)\frac{\sqrt{\operatorname{Var}(\ell_f|\mathcal{G}_i)\operatorname{Var}(\ell_m|\mathcal{G}_i)}}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}}.$$

Thus, the decomposition of the aggregate home production correlation is given by:

$$\operatorname{Corr}(\ell_f, \ell_m) = \underbrace{\left(\sum_{i=1}^{r}\operatorname{Corr}(\ell_f, \ell_m|\mathcal{G}_i)\,\mathcal{V}_i\right)}_{\text{Within-Group Component}} + \underbrace{\operatorname{Corr}(\mathbb{E}[\ell_f|\mathcal{G}], \mathbb{E}[\ell_m|\mathcal{G}])\frac{\sqrt{\operatorname{Var}(\mathbb{E}[\ell_f|\mathcal{G}])\operatorname{Var}(\mathbb{E}[\ell_m|\mathcal{G}])}}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}}}_{\text{Between-Group Component}} \tag{O.14}$$

Over Time. The over-change in total home production correlation can then be written as:

$$\begin{aligned}\Delta_{t,t-1}\operatorname{Corr}(\ell_f, \ell_m) = &\underbrace{\sum_{i=1}^{r}\left(\overline{(\operatorname{Corr}(\ell_f, \ell_m)\,|\mathcal{G}_i)}\underbrace{(\mathcal{V}_{i,t} - \mathcal{V}_{i,t-1})}_{\text{Change in Group Weights}} + \overline{\mathcal{V}_i}\underbrace{(\operatorname{Corr}(\ell_f, \ell_m|\mathcal{G}_i)_t - \operatorname{Corr}(\ell_f, \ell_m|\mathcal{G}_i)_{t-1})}_{\text{Change in Within-Group Correlation}}\right)}_{\text{Change in Within-Group Component}} \\ &+ \underbrace{\Delta_{t,t-1}\left(\operatorname{Corr}(\mathbb{E}[\ell_f|\mathcal{G}], \mathbb{E}[\ell_m|\mathcal{G}])\frac{\sqrt{\operatorname{Var}(\mathbb{E}[\ell_f|\mathcal{G}])\operatorname{Var}(\mathbb{E}[\ell_m|\mathcal{G}])}}{\sqrt{\operatorname{Var}(\ell_f)\operatorname{Var}(\ell_m)}}\right)}_{\text{Change in Between-Group Component}}\end{aligned} \tag{O.15}$$

Results: Couples With and Without Children. We consider two groups $r = 2$ with $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2\} = \{\text{Couples with Children, Couples without Children}\}$. Implementing over-time decomposition (O.15) with these two couple types, we obtain the following results.

Table O.15: Decomposing Changes in Home Production Correlation: Couples with and without Children

Overall Change	Between (%)	Within(%)
0.12	17.9	82.1

Notes: This decomposition is based on (O.15). Column 1 gives $\Delta_{t,t-1}\,\text{Corr}\left(\ell_f, \ell_m\right)$, while columns 2 and 3 report the second and the first lines on the RHS of (O.15), respectively. Source: GSOEP.

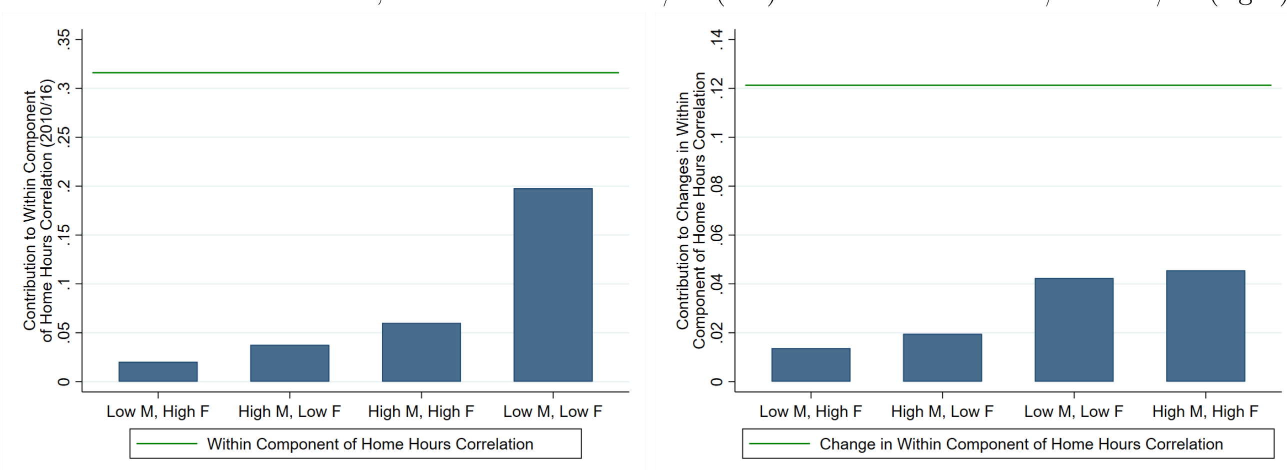
Results: Couples of Different Education. We consider four groups $r = 4$ with $\mathcal{G} = \{\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4\} = \{LL, LH, HL, HH\}$, where the first (second) entry refers to the education of the husband (wife), and L indicates low education and H high education. Implementing over-time decomposition (O.15) with these four couple types, we obtain the following results.

Table O.16: Decomposing Changes in Home Production Correlation: Couples of Different Education

Overall Change	Between (%)	Within(%)
0.12	1.9	98.1

Notes: This decomposition is based on (O.15). Column 1 gives $\Delta_{t,t-1}\,\text{Corr}\left(\ell_f, \ell_m\right)$, while columns 2 and 3 report the second and the first lines on the RHS of (O.15), respectively. Source: GSOEP.

Figure O.13: Contribution of Different Education Groups to the Within-Component of the Aggregate Home Production Correlation, Cross-Section of 2010/16 (left) and Over Time 1990/96-2010/16 (right)



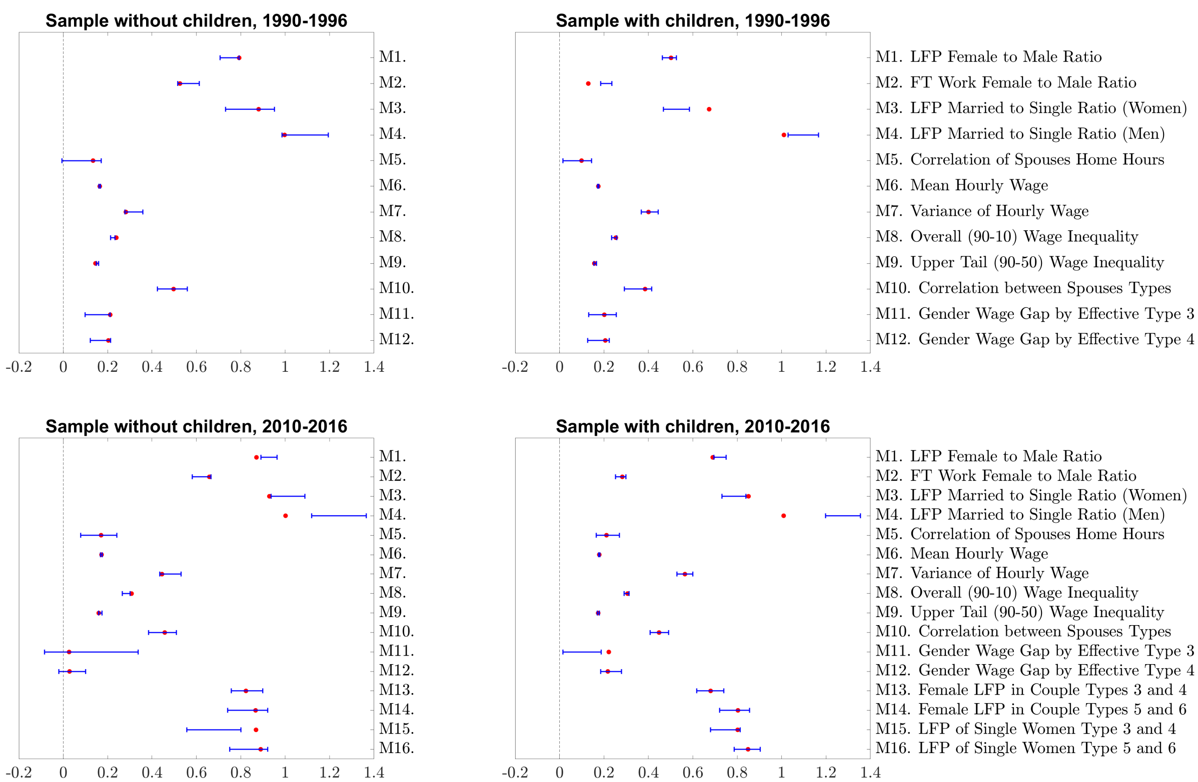
Notes: Left panel plots the contribution $\text{Corr}\left(\ell_f, \ell_m|\mathcal{G}_i\right)\mathcal{V}_i$ for each group i to the total within-group component, see (O.14). Right panel plots the over-time *changes* in the contribution $\text{Corr}\left(\ell_f, \ell_m|\mathcal{G}_i\right)\mathcal{V}_i$, that is $\overline{(\text{Corr}\left(\ell_f, \ell_m\right)|\mathcal{G}_i)}(\mathcal{V}_{i,t} - \mathcal{V}_{i,t-1}) + \overline{\mathcal{V}_i}(\text{Corr}\left(\ell_f, \ell_m|\mathcal{G}_i\right)_t - \text{Corr}\left(\ell_f, \ell_m|\mathcal{G}_i\right)_{t-1})$, for each group i, see (O.15). 'Low M' ('Low F') means that the husband (wife) of the couple has low education, and so on. Since the between-component is negligible, both in the cross-sectional decomposition (O.14) (where it accounts for only 0.005% of the aggregate correlation) and in the over-time decomposition (O.15) (where it accounts for only 1.9% of changes in the aggregate correlation), we focus on the contributions of different couple types to the *within*-component of the aggregate home production correlation. Education groups are defined in the notes to Figure O.12. Source: GSOEP.

OG.4 Model Re-Estimation: Heterogeneous Home Production by Couple Type

Couples With and Without Children. We assess whether the model parameters—especially, the home production technology—differ by the presence of children in the household. To this end, we first divide our sample into two groups, based on whether children ‘typically’ live in the household (see Online Appendix OD.3.2 for details). We then estimate our model on each sample, where we compute and target moments for each sample separately.[17]

Table O.17 shows the estimates and their standard errors for both samples and both periods, 1990-1996 and 2010-2016 (i.e., we have four estimations). Focusing on home production parameters in the 2010-2016 period, the sample with children features a significantly higher relative productivity of women in home production, with $\theta = 0.91$ (versus $\theta = 0.76$ for the sample without children). Moreover, complementarities in home production are most pronounced for the sample with children with $\rho = -0.39$ (versus $\rho = 0.01$ for the sample without children). Using a standard Wald test, we reject the null hypothesis that ρ is equal across samples with and without children at the 5% level, and similarly for θ. Additionally, the sample with children features the largest increase in complementarities over time, based on ρ evolving from 0.16 in 1990-1996 to -0.39 in 2010-2016. Finally, Figure O.14 shows that these four models fit the targeted moments well, with the vast majority of model-produced moments falling within the confidence intervals of their data counterparts.

Figure O.14: Model Fit: Model Moments (red) with Data Confidence Intervals (blue)



Notes: The red dots indicate the level of the model moments while the blue bars are their corresponding data 90%-confidence intervals, computed from a bootstrap sample. We rescaled moments $M6 - M9$ to be able to plot all moments in the same graph.

[17] To compute the moments involving home production, we define hours in home production as time spent on housework for couples with no children, and as time spent on housework *and* childcare, for the sample of couples with children.

Table O.17: Estimated Parameters by Presence of Children: 1990-1996 versus 2010-2016

	Sample without children				Sample with children			
	1990-1996		**2010-2016**		**1990-1996**		**2010-2016**	
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
θ	0.76	0.07	0.76	0.06	0.92	0.03	0.91	0.02
ρ	0.23	0.36	0.01	0.19	0.16	0.07	-0.39	0.07
A_p	47.88	7.08	40.75	10.53	38.65	2.42	33.67	1.84
γ_1	0.22	0.17	0.33	0.16	0.50	0.07	0.62	0.05
γ_2	0.17	0.34	0.29	0.27	0.20	0.13	0.20	0.05
A_z	49.87	36.42	54.80	23.20	37.55	3.76	41.12	1.58
ψ	0.60	0.05	0.99	0.06	0.78	0.03	0.75	0.02
σ_β	0.01	0.00	0.03	0.00	0.02	0.01	0.07	0.00

Notes: s.e. denotes standard errors. See Section 4.4.2 for a description of how these standard errors are computed.

Couples of Different Skill Composition. We also estimate a modified version of our model, in which we allow the parameters of home production to depend on the skill composition of households. To do so, we categorize individuals into low-skilled (denoted by L) and high-skilled (denoted by H). L-types are individuals with the three lowest skill levels (in terms of s, see columns 1-3 in Table O.13) and H-types are those with the three highest levels (columns 4-6 in Table O.13). Based on this categorization, we observe four types of couples, which we denote by $\mathcal{G}$:

$$\mathcal{G} \in \{LL, LH, HL, HH\}.$$

In this specification of the model, the public good production function is assumed to be CES and its parameters are allowed to be heterogeneous across types of couples $\mathcal{G}$:[18]

$$p^M(1-h_m, 1-h_f) = A_p\Big[\theta_{\mathcal{G}}\big(1-h_f\big)^{\rho_{\mathcal{G}}} + (1-\theta_{\mathcal{G}})\big(1-h_m\big)^{\rho_{\mathcal{G}}}\Big]^{\frac{1}{\rho_{\mathcal{G}}}}. \tag{O.16}$$

For estimation, we discipline the heterogeneous parameters of home production by targeting the moments that are related to time allocation choices *by type* of couple or single (moments M1 to M5).

Table O.18 shows the results of estimating this model in both periods, 1990-1996 and 2010-2016, when home production is specified as in (O.16).[19] Our results indicate that the parameters of home production do not vary significantly with the skill composition of households. For example, in 2010-2016, the relative productivity of women at home is similar across all types of couples, with the estimated $\theta_{\mathcal{G}}$ ranging from 0.65 to 0.82. Pairwise Wald tests confirm that the difference in $\theta_{\mathcal{G}}$ between any two

[18]Analogously, the home production function for singles depends on whether the individual is of type L or H, where only the relative female productivity, θ, varies with skill. We restrict θ of single women of type L (H) to equal the simple average of θ_{LL} and θ_{HL} (θ_{LH} and θ_{HH}), i.e., the average of $\theta_{\mathcal{G}}$ of married households involving women of type L (H).

[19]Note that we restrict home production TFP to be the same across couple types as it reflects access to technology, which arguably does not vary across couples' skill groups.

types of couples is not statistically significant at the 5% level. Likewise, all types of couples feature quantitatively similar home production complementarities, with the estimate of $\rho_{\mathcal{G}}$ ranging between -0.94 and -0.36 across groups. The pairwise differences between these parameters are not statistically significant. Even though this model fits the data well (as shown in Figure O.15), we cannot reject the statistical hypothesis that a restricted version of this model—which imposes that home production parameters are the same across groups—represents the data better.[20]

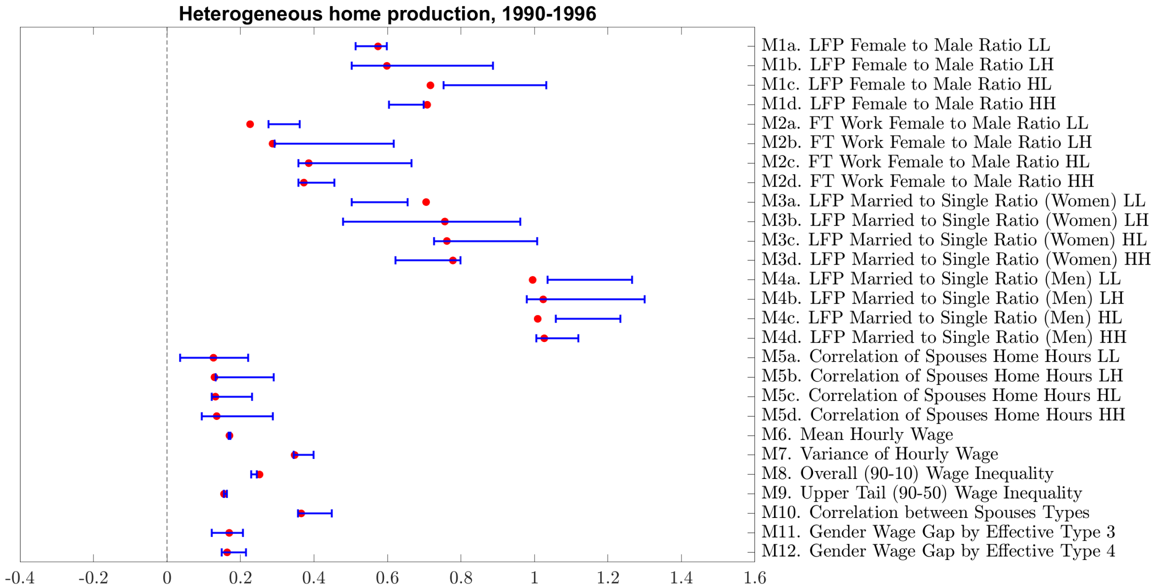
Table O.18: Estimated Parameters by Couples' Skill Type: 1990-1996 versus 2010-2016

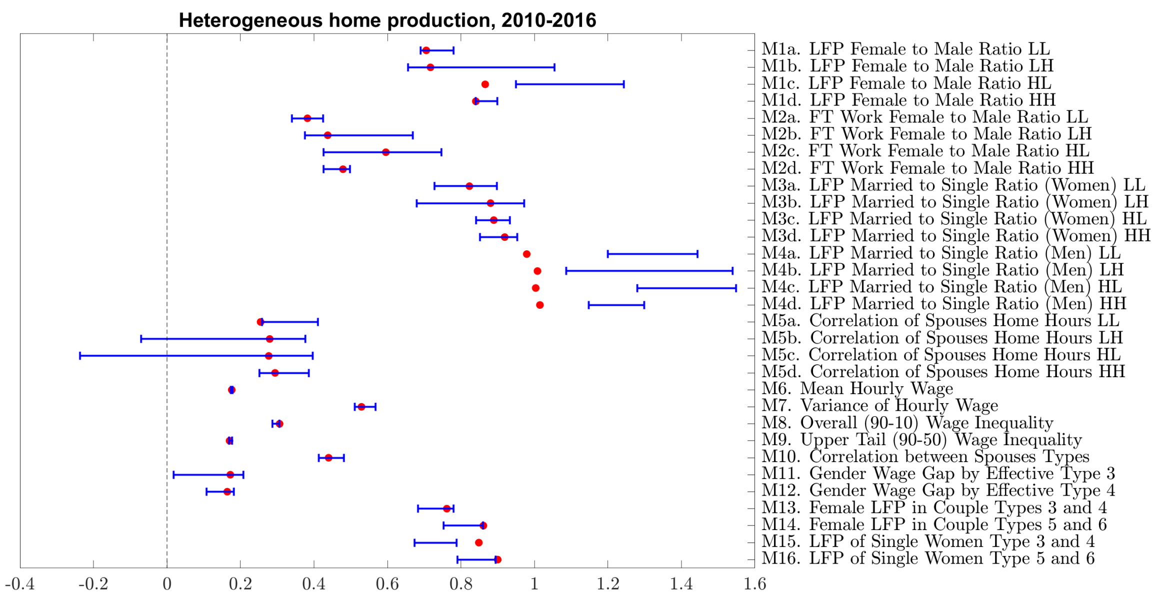
	1990-1996		**2010-2016**	
	Estimate	s.e.	Estimate	s.e.
θ_{LL}	0.87	0.00	0.82	0.00
θ_{LH}	0.88	0.00	0.82	0.06
θ_{HL}	0.61	0.44	0.65	0.13
θ_{HH}	0.88	0.02	0.82	0.03
ρ_{LL}	0.01	0.19	-0.61	0.39
ρ_{LH}	0.01	0.23	-0.36	0.34
ρ_{HL}	0.01	0.09	-0.94	0.38
ρ_{HH}	0.01	0.35	-0.48	0.20
A_p	38.61	1.16	36.27	0.28
γ_1	0.40	0.07	0.60	0.06
γ_2	0.14	0.12	0.19	0.09
A_z	39.80	5.64	42.43	3.29
ψ	0.78	0.02	0.80	0.01
σ_β	0.03	0.02	0.12	0.48

Notes: s.e. denotes standard errors. See Section 4.4.2 for a description of how these standard errors are computed.

[20] Underlying this statement is the following: We perform a likelihood ratio test comparing the model presented here with a restricted version of the model that imposes homogeneous home production parameters. We estimate both models, targeting the same moments (i.e. moments by type of couple).

Figure O.15: Model Fit: Model Moments (red) with Data Confidence Intervals (blue)





Notes: The red dots indicate the level of the model moments while the blue bars are their corresponding data 90%-confidence intervals, computed from a bootstrap sample. We rescaled moments $M6 - M9$ to be able to plot all moments in the same graph. In the moments descriptions, 'H' stands for 'High-Skilled' and 'L' for 'Low-Skilled'. The first type entry refers to the husband while the second one refers to the wife; see this appendix for details.

OH Home Production in GTUS vs. Home Production in GSOEP

Even though home production tasks in the GSOEP are less disaggregated, we find similar patterns as in the GTUS. The magnitude of the over-time change in aggregate home production correlation is similar across datasets, from 0.09 to 0.24 between 1991/92 and 2012/13 in the GTUS and from 0.19 to 0.32 between 1990-1996 and 2010-2016 in the GSOEP. Moreover, also in the GSOEP, partner's correlation in childcare hours is larger than in other home production activities (0.53 versus 0.07). This also holds when controlling for confounding factors, see Table O.19 below.

Table O.19: Complementarity in Home Production Hours: Childcare vs. Housework

	(1) Childcare Male Hours	(2) Housework Male Hours
Childcare Female Hours	0.207***	
	(0.024)	
Housework Female Hours		0.117***
		(0.021)
Demographic Controls	Yes	Yes
State and Year FE	Yes	Yes
Period	2010-2016	2010-2016
Observations	4,007	3,874
R^2	0.215	0.080

Notes: Standard errors clustered at the state level in parentheses. *** p<0.01, ** p<0.05, * p<0.1. We impose the same sample restrictions and controls as in Table O.6. We further restrict our attention to the sample of households with children under 18 years old. Source: GSOEP.

References

J. D. Angrist. Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics*, 47(2):243–266, 1991.

S. Bauernschuster and M. Schlotter. Public child care and mothers' labor supply—evidence from two quasi-experiments. *Journal of Public Economics*, 123:1–16, 2015.

R. Blundell, A. Duncan, and C. Meghir. Estimating labor supply responses using tax reforms. *Econometrica*, 66(4):827–861, 1998.

C. Boll and A. Lagemann. Public childcare and maternal employment—new evidence for Germany. *LABOUR*, 33(2):212–239, 2019.

M. Browning, P.-A. Chiappori, and Y. Weiss. *Economics of the Family*. Cambridge Books. Cambridge University Press, 2014.

S. K. Burns and J. P. Ziliak. Identifying the elasticity of taxable income. *The Economic Journal*, 127 (600):297–329, 2016.

Y. Chang and F. Schorfheide. Labor-supply shifts and economic fluctuations. *Journal of Monetary Economics*, 50(8):1751–1768, 2003.

P.-A. Chiappori and E. Gugl. Transferable utility and demand functions. *Theoretical Economics*, 15(4), 2020.

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

G. de Clippel, K. Eliaz, and K. Rozen. Competing for consumer inattention. *Journal of Political Economy*, 122(6):1203–1234, 2014.

L. Eika, M. Mogstad, and B. Zafar. Educational assortative mating and household income inequality. *Journal of Political Economy*, 127(6):2795–2835, 2019.

R. Fernández and A. Fogli. Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, 1(1):146–77, 2009.

A. Fogli and L. Veldkamp. Nature or nurture? Learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138, 2011.

N. M. Fortin. Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery. *Annals of Economics and Statistics*, (117/118):379–401, 2015.

J. Greenwood, N. Guner, and G. Vandenbroucke. Family economics writ large. *Journal of Economic Literature*, 55(4):1346–1434, December 2017.

German Time Use Survey (GTUS) Data for 1991/92 and 2012/13. German Federal Statistical Office. doi: 10.21242/63911.1992.00.00.4.1.0 and 10.21242/63911.2013.00.00.4.1.0

A. Hall, A. Seifer, and M. Tiemann. Bibb/baua employment survey of the working population on qualification and working conditions in germany 2012. Version 6.0. doi:10.7803/501.12.1.1.60, Research Data Center at BIBB (ed.); GESIS Cologne (data access); Bonn: Federal Institute for Vocational Education and Training., 2020.

S. Heise and T. Porzio. Spatial wage gaps in frictional labor markets. 2019.

K. Kalíšková. Tax and transfer policies and the female labor supply in the EU. *Empirical Economics*, 58(2):749–775, 2020.

E. Maurin and J. Moschion. The social multiplier and labor market participation of mothers. *American Economic Journal: Applied Economics*, 1(1):251–72, January 2009.

D. McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026, 1989.

J. Merz and M. Ehling. Time use – research, data and policy. Technical report, NOMOS Baden Baden, 1999.

P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

N. Mota, E. Patacchini, and S. Rosenthal. Neighborhood effects, peer classification, and the decision of women to work. Working Paper, 2016.

K.-U. Müller and K. Wrohlich. Does subsidized care for toddlers increase maternal labor supply? evidence from a large-scale expansion of early childcare. *Labour Economics*, 62:101776, 2020.

OECD. Starting strong IV. Early childhood education and care data country note. Germany. Technical report, OECD Publishing, Paris, 2016.

C. Olivetti, E. Patacchini, and Y. Zenou. Mothers, peers, and gender-role identity. *Journal of the European Economic Association*, 18(1):266–301, 2020.

A. Pakes and D. Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica*, 57 (5):1027–1057, 1989.

D. Rohrbach-Schmidt and A. Hall. Bibb/baua-erwerbstaetigenbefragung 2012. BIBB-FDZ Daten- und Methodenberichte 1/2013 Version 6.0., Bonn: BIBB. ISSN 2190-300X, 2020.

Socioeconomic Panel (SOEP) Data for Years 1984-2018, version 35, soep, doi:10.5684/soep.v35., 2019.

A. Stuckemeier and C. Kuehnen. German time use survey 2012/13. Technical report, Federal Statistical Office, Bonn, Germany, 2013.