Culture and marriage: Lessons from German reunification*

Marion Goussé† Nicolas Jacquemet‡ Jean-Marc Robin§

March 9, 2023

Abstract

Recent work by Lippman et al. (2020) examines the effect of German reunification in 1989 on gender and family norms, and show that female labor supply varies with husband-wife relative wage more in the West than in the East. In this paper, we try to understand what factors explain this fact. We first construct a cultural identity index from the attitude questionnaire in the GSOEP, with significant overlapping across East and West. We then build a search-matching/Nash bargaining model of marriage formation and divorce, and intrahousehold resource allocation, assuming same preferences in East and West, conditional on heterogeneous wages, education and cultural identity. We estimate 4 independent models in four 7-year periods (1992-1998, 1999-2005, 2006-2012, 2013-2019) using the GSOEP. The model fits the data very well. We then use the estimated model to measure how much of the observed differences between East and West (time uses and marriage sorting) are due to differences and changes in education, wages and cultural identity. Cultural identity is found to be the main factor before education and wages. It explains half of differences in time uses and marriage sorting on observables between East and West. Identity also interacts with education and wages (another 50% reduction).

Keywords. Cultural identity, household resource sharing, labor supply, marriage.

---

*We benefited from helpful seminar presentations at the Francqui conference (June 2022), at Yale, Sciences Po and PSE. We had grateful to Pierre Cahuc, Michele Fioretti and Helene Turon for helpful comments.

†CREST-ENSAI, 51 Rue Blaise Pascal, 35172 Bruz, France; marion.gousse@gmail.com;

‡Paris School of Economics and University Paris 1 Panthéon–Sorbonne; Nicolas.Jacquemet@univ-paris1.fr; Centre d'Economie de la Sorbonne, MSE, 106 Bd. de l'Hospital, 75013 Paris, France.

§Sciences Po, Paris; jeanmarc.robin@sciences-po.fr; Sciences-Po, Department of Economics, 28 rue des St Pères, 75007 Paris, France. I gratefully acknowledge the support from the European Research Council (grant reference ERC-2020-ADG-101018130).

1 Introduction

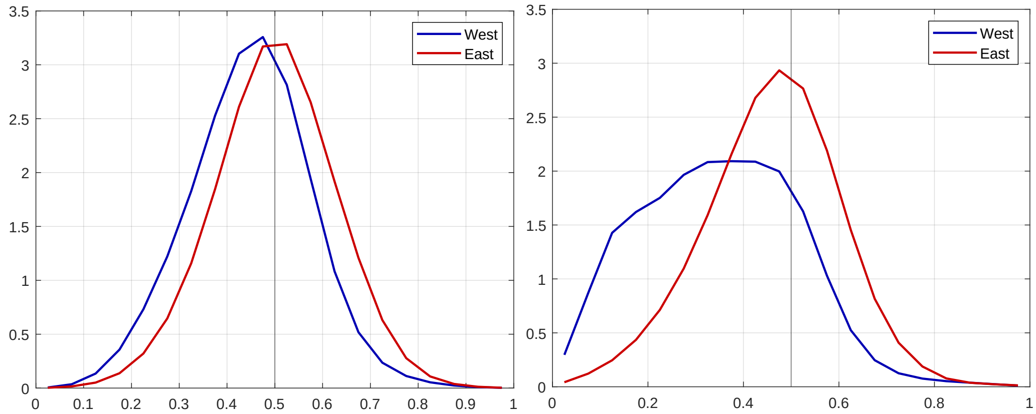
Following Becker's (1973, 1974) pioneering work on marriage, a long line of economic research has sought to link the dynamics of various gender inequalities in education, labor market participation, and marital structure to environmental changes affecting men's and women's comparative advantages in household tasks (see Greenwood, Guner, and Vandenbroucke, 2017, for a survey). For example, it is now well established that technological change, medical advances, greater availability of childcare, reduced discrimination, changes in divorce laws, etc. have all contributed to a substantial increase in female labor force participation (among other outcomes), which in turn has changed social norms relating to gender, marriage and employment (Fernández, 2013, Bau and Fernández, 2021).

However, social norms are slow to change. A emerging literature therefore seeks to identify other mechanisms that may explain the resilience of gender norms. For example, Cavapozzi, Francesconi, and Nicoletti (2021) infer from a detailed survey of UK mothers with dependent children that having peers with gender-egalitarian norms leads mothers to be more likely to have a paid job and to have a greater share of the total number of paid hours in their household. On the contrary, Bursztyn, Fujiwara, and Pallais (2017) found in a field experiment the mirror result that female students reported lower desired salaries and hours worked when they expected their classmates to see their preferences. In another survey, they find that women tend to avoid actions signaling personality traits, like ambition, that they believe are undesirable in the marriage market. A more indirect evidence of a social norm influencing time sharing is documented by Bertrand, Kamenica, and Pan (2015). They observe that the distribution of the share of income earned by the wife (among married couples in the United States) drops at 1/2, where the wife's income exceeds the husband's income. Moreover, in couples where the wife earns more than the husband, the wife spends more time on household work and the risk of divorce increases.

Building on Bertrand et al., Lippmann, Georgieff, and Senik (2020) use the 41-year division of Germany as a cultural experiment to support the idea that the male breadwinner norm is a product of Western institutions. More specifically, they compare the allocation of family resources in East and West Germany after 1989.[1] They find that West German couples resemble US couples as described by Bertrand et al. However, East German couples are different. Wages are more evenly distributed between wife and husband in the East than the West, and the distribution of the wife's share of income is much more symmetric and less concentrated to the left of 1/2 (see Figure 1).

These facts are interesting and indicative that labor market specialization differs between East and West Germany. However, we tend to be more circumspect about the 1/2 relative income norm. Following Bertrand et al., they regress female domestic time uses on a dummy variable — equal to one if the wife earns more than the husband — interacted with the region, while controlling for the region, spouses' incomes and other socio-demographics. They find in the West that female domestic hours increase if the wife earns more, but not in the East. However, some right-hand-side variables are endogenous. There are undoubtedly clear differences in the time uses and wages of married couples between East and West Germany, but to make sense of

[1] Alesina and Fuchs-Schündeln (2007) documented the cultural consequences of the communist regime.



(a) Wage share, $\frac{w_f}{w_f + w_m}$ (b) Income share, $\frac{w_f h_f}{w_f h_f + w_m h_m}$

Figure 1: Distributions of labor earnings and wage ratios

them, one must understand the co-determinations at work.

What are they? First, the wife's wage relative to her husband is endogenous because wages are in the set of characteristics that male and female singles see and respond to when dating. The distribution of spouse characteristics, including wages, education, and many other cultural traits such as religion or the willingness to have children or a passion for winter sports, and so on, is an equilibrium object. We therefore need to endogenize who marries whom. Second, we need a model for the allocation of time resources between spouses given their characteristics. Third, the model needs to distinguish in the mating decision what is related to resource sharing from what is related to other types of sharing — the same religion, common tastes in entertainment, a shared desire to raise children, etc.

Our modeling framework is that of Goussé, Jacquemet, and Robin (2017a), which is similar to the standard model in family macroeconomics (e.g. Greenwood, Guner, Kocharkov, and Santos, 2016).[2] The core is Shimer and Smith's (2000) search-matching model, with more heterogeneity than just wages, and to which we add an additional layer for decisions about time uses (leisure, wage work and domestic production). We use a bargaining model for resource allocation between spouses because it allows us to select a single allocation from the set of Pareto efficient allocations.[3] In addition, we introduce a bliss shock that explains why certain couples of the same type mate and why others do not, and it also allows to endogenize divorce.

There is a vast literature on homogamy in anthropology, sociology and demography. For computational reasons, we will not be able to include in our analysis all the socio-economic characteristics of men and women that have been found to be positively correlated across couples. For the budget constraint, we have to condition on market wages. Then, we control for education, since education is both a determinant of preferences and of human capital, and thus a potentially important input into household production. Finally, we need to account for preference heterogeneity in some way. In this paper, we rely on a questionnaire from the German

[2]Other recent applications of our framework include Goussé, Jacquemet, and Robin (2017b), Ciscato (2021) and Holzner and Schulz (2022).

[3]Bargaining explains why the Pareto weights (Chiappori, 1988) depend on distribution factors such as the relative wage, local sex ratios, etc. (e.g. Chiappori, Fortin, and Lacroix, 2002). As a matter of fact, life-cycle collective models with commitment often use bargaining to determine Pareto weights in the first period.

Socio-Economic Panel (GSOEP), which asks respondents what they consider important for satisfaction (a happy marriage, children, money, travel, etc.), what activities they enjoy (music, cinema, sports, etc.), and their religion.

We will aggregate the responses to these cultural questions into one single score of "West Germanity" by running a logistic regression of living in West Germany in 1989 on these attitudinal responses, separately for men and women. That is, we construct our cultural score by selecting the questions that best discriminate between the individuals in East and West Germany in 1989. Nevertheless, the logistic score distributions in the East and in the West overlap, so that some East Germans resemble West Germans and vice versa. We also show that dissident individuals are more likely to move across regions (in both directions), and that in culturally similar couples, regardless of the region in which they live, women work more.

We use the model to ask the following question. Assuming that exactly the same mechanisms are at work in the processes of mating and of resource allocation between spouses in both regions, how well can we explain differences in time uses by gender in East and West Germany, and their evolution, solely in terms of exogenous differences and variations in individual specific endowments in wages, education and preferences?

We draw data from the GSOEP on time uses, wages, education, cultural attitudes and marital status (married or cohabiting and single) between 1992 and 2019, burning a few of years after 1989 to allow the society to stabilize a bit. We pool East and West Germany and estimate a single model for all Germans. We allow for a linear trend in time uses, but all preference and household production parameters are otherwise constant across time and region.[4] The dynamics of male and female distributions of wages, education and West Germanity are exogenous. The model parameters are estimated so as to best fit the distribution of characteristics by marital status and gender, and the distribution of time uses by characteristics and gender.

Having estimated the model on four seven-year periods, we take the distributions of characteristics in East and West Germany as given and predict sorting and time uses by region. The fit across regions is good — spousal income shares included — in fact remarkably good given that we exclude any particular gender norms and ignore any differences across regions in the provision of public goods such as childcare facilities. Our main conclusion is therefore that East and West German couples differ not so much because of different norms on the division of labor, but because the marriage markets in the East and the West differ in the distributions of preferences, wages and education by gender, leading to different compositions of couples.

Finally, we proceed to a series of counterfactual analyses to determine which of wages, education and culture has the greatest influence on sorting. Specifically, we consider the distribution of individual types in West Germany and we change, say, the marginal distribution of education to make it like in East Germany. Then, we calculate by how much the distance between the distributions of characteristics and time uses of West and East Germany has been reduced. We use the Kullback-Leibler divergence to measure the distance between distributions. We find that culture explains 50% of sorting while education and wages explain a smaller 25%. Culture and education together, and culture and wages, explain 75%, whereas education and wages together

[4]Holzner and Schulz (2022) have even recently argued that most of the increase in female labor market participation of these last 20 years is related to the Hartz reforms of the early 2000s. These structural changes are factored in in our model through exogenous trends.

explain 60%. We also find that culture moves labor supply and education moves domestic time, and that only culture has an influence on the income ratio.

The layout of the paper is as follows. Section 2 presents the data and some descriptive evidence about wages, education and cultural differences between East and West Germany after 1989. Section 3 presents the model. Section 4 presents the empirical strategy for estimation. In Section 5 we show how the estimated model fits the regional data. Section 6 is devoted to counterfactual analyses. The last section concludes.

2 Preliminary evidence

2.1 Data

We use the GSOEP representative samples covering the 1992-2019 period.[5] We start in 1991, excluding a couple of years after 1989 to let the economy absorb some of the shock. Each year we select individuals aged 19-55, born in Germany (i.e. non migrant). Each individual is either single or living with a stable partner (marriage or cohabitation). We exclude married couples living separately, divorced couples living together, and same-sex couples. On average, in the sample of couples living in West Germany in any given year, 2.8% did not exist in the previous year (marriage rate) and 1.2% separated the year after (divorce rate). In the East, we estimate slightly lower rates: 2.2% and 1.1%.

In addition to the living area, gender and marital status, we observe gross labor earnings, working hours, housework (excluding child care), and education.[6] We keep only labor-market-active individuals, dropping self-employed. Hourly wages are calculated by dividing labor earnings (in constant euros of 2010) by the number of hours. Education is classified into three levels: (1) low education (general elementary school and basic vocational qualification), (2) medium education (intermediate general qualification, intermediate vocational education, general maturity certification, vocational maturity certification), and (3) high education (tertiary education).[7] The final number of non missing data (non missing wages and non zero hours) used for the empirical analysis is $43,686$ men and $44,652$ women from West Germany, and $20,667$ men and $20,649$ women from the East. Table 1 provides descriptive statistics for the estimation sample.

2.2 Education and wages

In the early 1990s, East and West Germany display stark differences in terms of education and wages, and the subsequent 30 years are characterized by remarkable changes. Figure 2 shows that at the beginning of our observation period East Germans are a lot more educated that Westerners, particularly women. This difference gradually disappeared between 1991 and 2019. Wages are higher in the West (more than 1.5 times higher). The absolute gender wage gap

[5] We keep the initial sample that started in 1984 in West Germany ([A]) and the additional sample drawn in East Germany in 1990 ([C]). We then include the refreshment samples. This ensures that the data is representative of the German population ([E] in 1998, [F] in 2000, [H] in 2006, [I] in 2009, [J] in 2000, [K] in 2012). Other additional samples of the GSOEP overweight a specific category of the population such as high-income households or large families with low income or single mother families. We exclude those.

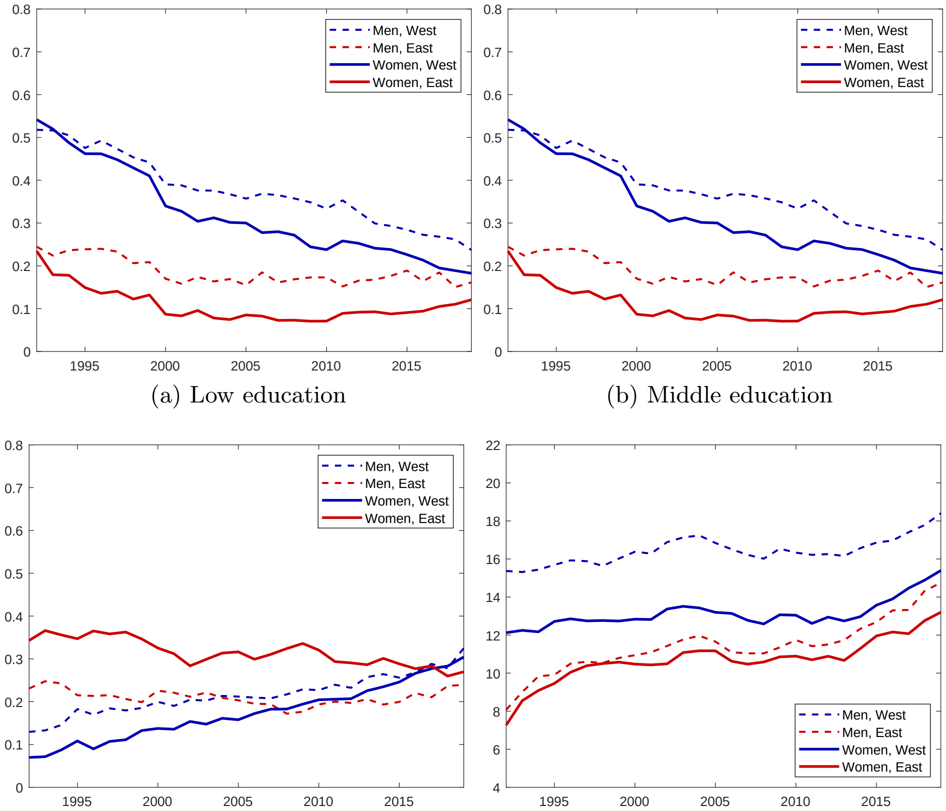
[6] We aggregate the information on housework realized on Saturdays, Sundays and any representative weekday.

[7] We trim the top 1% and bottom 1% for hours, wages and identity index by gender, year and area (East/West).

Table 1: Descriptive Statistics

	West	East	Overall	West	East	Overall
		Men			**Women**	
Housework	5.77	6.02	5.85	13.6	13.7	13.6
	(4.35)	(4.54)	(4.41)	(7.95)	(6.48)	(7.52)
Paid work	42.3	43.7	42.7	31.7	37.9	33.6
	(8.25)	(8.09)	(8.23)	(12.2)	(9.01)	(11.7)
Wage	16.4	11.1	14.8	13.1	10.4	12.3
	(7.27)	(5.46)	(7.20)	(6.00)	(5.11)	(5.87)
Cultural index	8.36	6.06	7.63	8.83	5.79	7.89
	(1.82)	(1.40)	(2.00)	(2.08)	(1.72)	(2.42)
Education	1.85	2.02	1.90	1.87	2.21	1.97
	(0.75)	(0.63)	(0.72)	(0.68)	(0.62)	(0.68)
Age	38.9	38.8	38.9	38.3	38.6	38.4
	(9.90)	(10.1)	(9.95)	(9.58)	(9.51)	(9.56)
N	31,465	14,507	45,972	31,436	14,090	45,526
		Married men			**Married women**	
Housework	5.58	5.92	5.69	16.2	15.3	15.9
	(4.27)	(4.53)	(4.36)	(8.25)	(6.42)	(7.70)
Paid work	43.3	44.6	43.7	28.3	37.8	31.5
	(7.29)	(7.65)	(7.44)	(12.3)	(8.72)	(12.1)
Wage	18.4	12.3	16.4	13.5	11.1	12.7
	(6.87)	(5.31)	(7.02)	(5.91)	(4.93)	(5.72)
Cultural index	8.24	5.90	7.47	8.83	5.66	7.79
	(1.82)	(1.34)	(2.00)	(2.05)	(1.69)	(2.45)
Education	1.88	2.14	1.96	1.89	2.30	2.02
	(0.77)	(0.59)	(0.72)	(0.67)	(0.58)	(0.67)
Age	42.5	42.7	42.6	40.3	40.5	40.4
	(7.88)	(7.81)	(7.85)	(7.96)	(7.85)	(7.92)
N	18,390	9,091	27,481	18,390	9,091	27,481
		Single men			**Single women**	
Housework	6.04	6.18	6.08	9.91	10.8	10.2
	(4.45)	(4.55)	(4.48)	(5.75)	(5.50)	(5.70)
Paid work	40.9	42.3	41.3	36.3	38.1	36.8
	(9.26)	(8.59)	(9.09)	(10.4)	(9.52)	(10.2)
Wage	13.7	9.19	12.4	12.5	9.19	11.6
	(6.92)	(5.17)	(6.77)	(6.07)	(5.19)	(6.02)
Cultural index	8.51	6.33	7.88	8.81	6.01	8.04
	(1.81)	(1.46)	(1.98)	(2.11)	(1.76)	(2.38)
Education	1.80	1.83	1.81	1.84	2.04	1.90
	(0.72)	(0.65)	(0.70)	(0.70)	(0.65)	(0.69)
Age	33.9	32.2	33.4	35.5	35.1	35.4
	(10.2)	(9.96)	(10.9)	(11.1)	(11.0)	(10.8)
N	13,075	5,416	18,491	13,046	4,999	18,045

Notes: *Housework* and *Hours worked* are measured as hours per week. *Wages* are measured in Euros. *Education* is equal to 1 for low education, 2 for medium and 3 for high education. Sample of 18-55 year old individuals born in Germany from the GSOEP, 1992-2019 period.



(a) Low education (b) Middle education

(c) High education (d) Mean wages (€/h)

Figure 2: Education and wage trends

is smaller in the East (10.7% on average, compared to 24.7% in the West). However, while the gender gap has narrowed in West Germany over the past decade, after many years of wage moderation (2000-2013), it has increased in the East.

2.3 Measuring preference heterogeneity

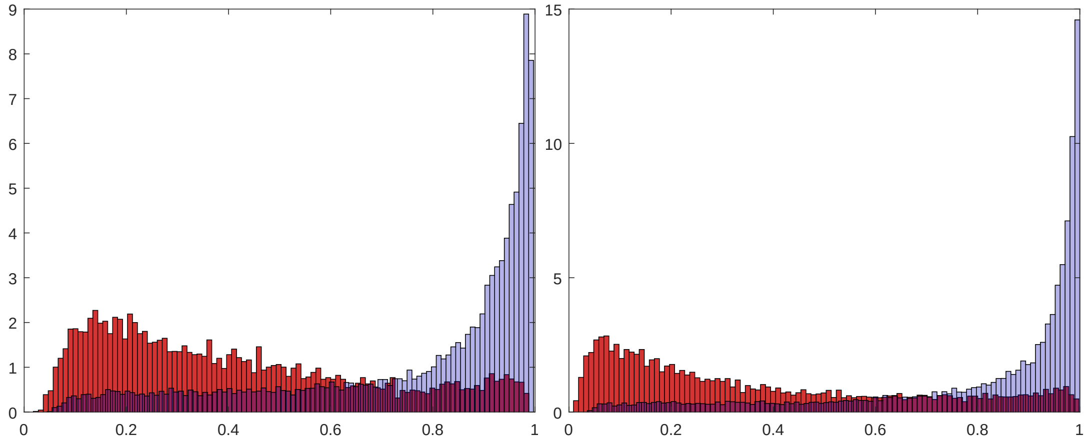
In the marriage market, singles meet and observe the characteristics of potential partners, and then decide whether or not to match. Education and wages are two main drivers of socio-economic status, which is one criterion for mating. Cultural affinity is another important mating criterion. Whether we both like opera or sports, whether we share similar views about the family, politics, religion, etc, is likely to matter for marriage. We could assume that there is unobserved heterogeneity in preferences, but this would tremendously complicate the identification and the estimation of the structural model of marriage and resource allocation. Instead, we make use of the GSOEP cultural questionnaire (available in 1995, 2004, 2008, 2012 and 2016) to construct a synthetic cultural index.

The questions used are listed in Table 2. There are questions about participation in recreational, artistic and sport activities, about religious practice and about life expectations (or attitudes). We construct a Western identity score by regressing the probability of living in the West in 1989 on activities, attitudes and religious practice, separately for men and women. More precisely, we estimate the two logistic regressions using the survey for the years 1991-1994. Then, we use the regression parameters to calculate identity scores for all the years for which the

Table 2: The western identity score

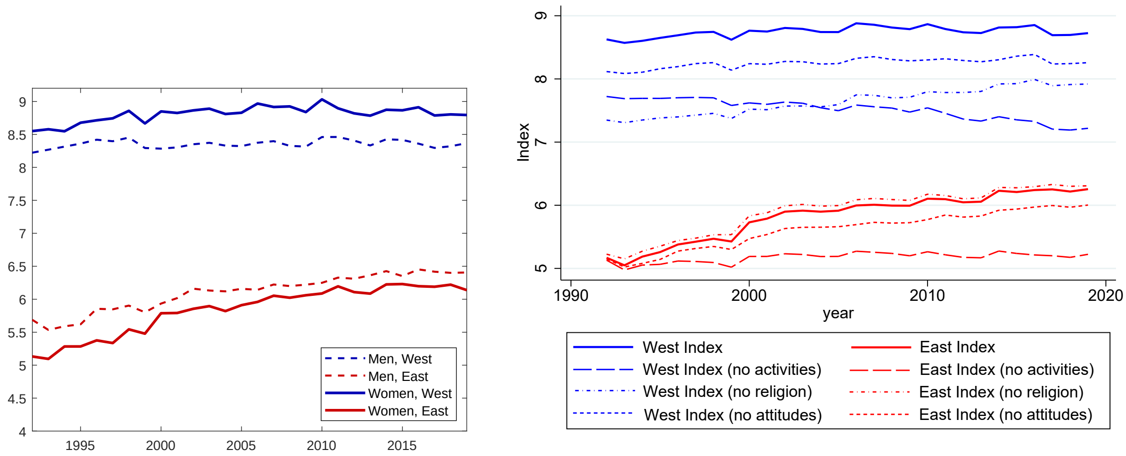
	Women	Men
Attitudes		
Important for satisfaction to ...		
... be able to afford something	0.079	-0.099**
	(0.052)	(0.048)
... help others	0.003	-0.103**
	(0.051)	(0.049)
... fulfill ones' potential	-0.051	-0.086**
	(0.046)	(0.042)
... success in the job	-0.812***	-0.327***
	(0.045)	(0.046)
... have an own house	0.090***	0.106***
	(0.032)	(0.031)
... have an happy marriage	0.086	0.307***
	(0.056)	(0.050)
... have children	-0.551***	-0.501***
	(0.039)	(0.036)
... be socially, politically active	-0.105**	0.137***
	(0.047)	(0.043)
... be able to travel	-0.188***	-0.075**
	(0.039)	(0.037)
Religion		
Religion/attend church		
No religion	Ref.	Ref.
catholic/weekly	2.748***	2.501***
	(0.137)	(0.139)
catholic/monthly	4.838***	3.886***
	(0.282)	(0.213)
catholic/less	4.392***	4.127***
	(0.144)	(0.149)
catholic/never	5.012***	4.050***
	(0.196)	(0.155)
protestant/weekly	1.516***	0.957***
	(0.227)	(0.215)
protestant/monthly	1.650***	1.492***
	(0.133)	(0.165)
protestant/less	2.460***	2.176***
	(0.077)	(0.076)
protestant/never	2.818***	2.662***
	(0.082)	(0.078)
other/weekly	2.560***	1.904***
	(0.293)	(0.305)
other/monthly	2.619***	1.518***
	(0.390)	(0.464)
other/less	0.329	-0.228
	(0.353)	(0.411)
other/never	-0.242	-1.229***
	(0.253)	(0.313)
Activities		
Attend opera, classic concert, theater		
Never	Ref	Ref.
Occasionally	0.388***	0.211***
	(0.068)	(0.062)
Every month	1.270***	0.881***
	(0.118)	(0.115)
Every week	2.193***	-0.155
	(0.469)	(0.356)
Activities (continued)		
Attend cinema, pop concert, disco		
Never	Ref	Ref.
Occasionally	-0.020	-0.304***
	(0.075)	(0.076)
Every month	-0.311***	-0.394***
	(0.104)	(0.096)
Every week	-1.431***	-1.489***
	(0.155)	(0.128)
Attend artistic activities		
Never	Ref	Ref.
Occasionally	-0.171***	-0.077
	(0.066)	(0.060)
Every month	-0.207**	0.291***
	(0.099)	(0.102)
Every week	-0.286***	0.219**
	(0.102)	(0.101)
Participation in sport		
Never	Ref	Ref.
Occasionally	0.593***	0.588***
	(0.070)	(0.068)
Every month	1.174***	1.239***
	(0.089)	(0.085)
Every week	1.713***	1.458***
	(0.103)	(0.084)
Attend social gathering/visit friends and family		
Never	Ref	Ref.
Occasionally	-1.229***	-0.743***
	(0.310)	(0.251)
Every month	-0.555*	-0.212
	(0.308)	(0.250)
Every week	0.343	0.347
	(0.312)	(0.254)
Perform volunteer work		
Never	Ref	Ref.
Occasionally	0.368***	-0.036
	(0.080)	(0.071)
Every month	0.654***	0.069
	(0.133)	(0.091)
Every week	1.007***	0.545***
	(0.219)	(0.126)
Participate In local politics		
Never	Ref	Ref.
Occasionally	-0.296**	-0.404***
	(0.121)	(0.096)
Every month	-0.555**	-0.439**
	(0.268)	(0.178)
Every week	0.121	0.138
	(0.502)	(0.251)
Constant	-4.678***	-2.513***
	(0.390)	(0.330)
Observations	11,783	11,512

Notes: Logistic regression of living in the West in 1989 on attitudes and religion over the years 1991-1994. Statistical significance: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.



(a) Men (b) Women

Figure 3: Distribution of predicted region of origin



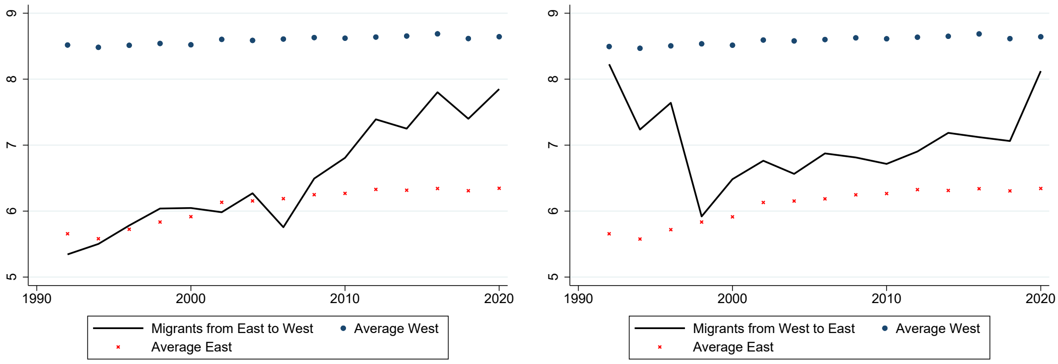
(a) Trends by region and gender (b) Components of female index

Figure 4: Western identity index

questions are posed between 1992 and 2019. By this way, all the variation of the index over the years for the same individual questioned several times, or between individuals, reflects changes in the answers, and not changes in the index weights.

The results are presented in Table 2, in which positive coefficients indicate traits more predictive of living in the West in 1989. West Germans value less having personal success, children, travels, and value more a happy marriage. They also prefer classical (music, theater, opera) to modern cultural activities (cinema, pop music, disco), and group activities (sport, social gathering, volunteer work) to being politically active. Lastly, religion is a lot more practiced in the West than in the East, as we could expect.

The cultural differences between the East and the West are pronounced, but the Eastern and Western distributions overlap significantly. Figure 3 displays the distributions of the probabilities of living in the West in 1989 predicted from the answers to the culture questions, separately by region and gender. While the distribution is concentrated near one for Westerners and near 0 for Easterners, many people in the West resemble East Germans and many in the East resemble



(a) East to West (b) West to East

Notes: The dotted lines are the average identity indexes in each region (as in Figure 2). The solid line is the average identity index of migrating individuals.

Figure 5: Identity and regional migration

West Germans. Specifically, 22.1% and 17.6% of West German men and women have a predicted probability of living in West Germany in 1989 below 0.5, and 27.1% and 26.5% East German men and women have a predicted probability above 0.5.

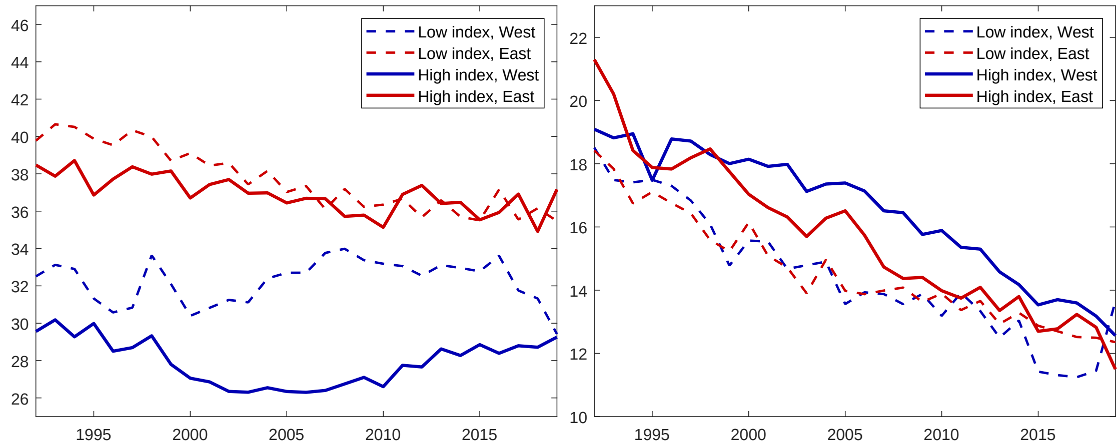
The average level of the index over time and region is presented in Figure 4a. The index discriminates well between individuals in the two blocs, and there are remarkably limited differences between men and women. Cultural differences are very stable over time (consistent with the evidence in Schmelz, 2021), despite an initial increase of the mean index in the East, indicating an initial convergence to Western standards, slowing down after 2005. Figure 4b shows how the three components (attitudes, activities and religion) affect these evolutions. If we remove activities from the score, then we remove most of the initial convergence of the East toward the West. Religion does not matter in the East, but it is a very significant component of cultural differentiation in the West.

2.4 Culture matters

The resulting cultural index has predictive power on regional mobility. Figure 5 shows the average identity index of individuals switching region across time. Over the years the index becomes increasingly predictive of a regional migration, both eastward and westward.

Then, we classify individuals depending on whether the predicted probability of living in the West in 1989 (given attitudinal responses) is above or below 0.5. Figure 6 displays the trends in labor supply and domestic hours for married women by cultural index (low/high) and region. A low index is associated with more hours of paid work and fewer hours of housework. This statistical link is stronger in the West than in the East.

Finally, we estimate what is left to the region for predicting married women's time uses once we control for both spouses' characteristics (education, wages and western identity/preferences). Table 3 first shows the overall effect of region with no control, then we add education (and age), then wages, then the cultural index. First, we see that the residual effect of the region on labor supply never ceases to be significant but is strongly reduced in magnitude by the introduction of



(a) Paid work hours

(b) Housework hours

Figure 6: Married women's weekly time uses by identity

Table 3: Residual effect of region on time uses controlling for characteristics

	Wife's paid work hours				Wife's housework hours			
female educ = Mid		2.3807	2.4966	2.6112		-2.0453	-1.8284	-1.8605
		(0.7967)	(0.7345)	(0.7135)		(0.5139)	(0.5016)	(0.4973)
female educ = High		2.1275	3.3156	3.1322		-1.6873	-1.7563	-1.6978
		(1.3108)	(1.2069)	(1.1578)		(1.0164)	(1.0148)	(0.9984)
male educ = Mid		2.8157	1.9918	1.9865		-2.9731	-2.3020	-2.2903
		(0.6296)	(0.6053)	(0.5895)		(0.4462)	(0.4369)	(0.4334)
male educ = High		7.5647	6.0604	6.2968		-5.3120	-4.0835	-4.1485
		(1.4727)	(1.4121)	(1.3725)		(1.0328)	(1.0062)	(1.0106)
log female wage			4.5668	4.6354			-3.0887	-3.1176
			(0.3331)	(0.3266)			(0.2128)	(0.2114)
log male wage			-4.5706	-4.0516			0.5058	0.3152
			(0.3659)	(0.3610)			(0.2309)	(0.2303)
log female identity				-4.8742				1.3567
				(0.5520)				(0.3465)
log male identity				-0.5951				1.0584
				(0.6361)				(0.3929)
East	9.5539	8.4069	7.6808	5.3641	-0.8449	0.3027	-0.4357	0.4952
	(0.3164)	(0.3384)	(0.3720)	(0.4365)	(0.2195)	(0.2184)	(0.2482)	(0.2831)
Observations	26,749	26,749	26,749	26,749	26,749	26,749	26,749	26,749
R-squared	0.1397	0.2076	0.2419	0.2554	0.0027	0.0983	0.1248	0.1297

Notes: The regressions on education, wages and cultural identity also contain age controls and interactions between male and female education.

individual characteristics in the regression. For domestic time, the region becomes non significant as we control for characteristics. Second, married women of Western cultural identity work less outside the home, and more inside, even after controlling for education, wages and region.

Next, the characteristics of both spouses are generally significant. For example, having a partner with medium or higher education increases female labor supply by 2 to 3 hours on average, and reduces female domestic hours by nearly 2 hours. Both wages affect female labor supply, while only the female wage matters for female domestic hours. The negative effect of the male wage on female labor supply is an expected income effect. The positive effect of the female wage on hours of paid work and the negative effect of the wage on domestic time are consistent with the interpretation that the substitution effect dominates the income effect. For cultural identity, this is the other way around. The cultural identities of both spouses influence women's domestic hours, while only female identity matters for female labor supply.

We conclude from this reduced form analysis that sorting matters for the married women's allocation of time between leisure, wage work and domestic tasks. The characteristics of husband and wife, chosen at marriage, jointly determine the allocation of time in the household. However, marriage sorting on characteristics is not random. We therefore need a model in which time uses are a function of male and female characteristics, and marriage results from gender specialization, as well as from other homogamic forces such as religion and education. This is is what we do in the next section.

3 The model

Our empirical approach closely follows the marriage market model with endogenous resource sharing between spouses developed in Goussé, Jacquemet, and Robin (2017a). We herein summarize the main building blocks of the model.

3.1 The intra-household decision model

The intra-household model is a dynamic version of the seminal Nash bargaining models of Manser and Brown (1980) and McElroy and Horney (1981).

3.1.1 Preferences for consumption and leisure, and home production

Individuals draw private utility from consumption (c), leisure (e) and a public good (q) produced in the household using personal time as sole input. Each individual is endowed with a fixed amount of time, arbitrarily set equal to one, to be shared between leisure (e), paid work (h) and housework (d).

Let $U(R, w, q)$ denote the maximal utility of consumption c and leisure e that a person can attain given income R and wage w, and for a given value q of a public good. We normalize the price of the consumption good to one. For simplicity, we only consider interior solutions with $c > 0$ and $1 - d > e > 0$, where housework d will be determined in a second stage. Market labor supply follows by difference: $h = 1 - d - e$. Note that by ruling out a corner solution at $e = 1 - d$ (no wage work) the indirect utility function is not a function of d given q.[8]

[8]Corner solutions break the linearity of the indirect utility function with respect to total income R. Utility

Specifically, we postulate a linear Gorman polar form

$$U(R,w,q) = q\frac{R - A(w)}{B(w)}, \quad A(w) = a_0 + a_1 w, \quad B(w) = w^b.$$

The denominator $B(w)$ is an individual-specific price index and $A(w)$ is a minimal total expenditure.[9] Then, consumption and leisure follow, by Roy's identity, as the linear expenditure system:

$$\begin{aligned} we &= a_1 w + b\left[R - A(w)\right], \\ c = R - we &= a_0 + (1-b)\left[R - A(w)\right]. \end{aligned}$$

We deliberately keep this part of the model as simple and linear as possible. In general, linearity makes models more difficult to identify, because factors are more easily confounded. Yet, we can prove that the model is identified and our estimation procedure closely follows the identification steps. Note that we also tried a version of the model with $A(w)$ quadratic in wage; but this extension did not improve the fit very much. Parameters a_0, a_1 and b may vary with gender, education and cultural identity (any individual characteristic but the wage).

3.1.2 Household production

The public good production function is

$$q = F^0(d) = \left(d - D^0\right)^{K^0}$$

for a single, and

$$q = zF^1(d_m, d_f) = z\left(d_m - D^1_m\right)^{K^1_m}\left(d_f - D^1_f\right)^{K^1_f}$$

for a couple, where m is the male index and f is the female index.

Singles and couples differ in two important aspects. First, couples can exploit the substitutability of male and female domestic times d_m and d_f in home production, and specialize in any way that is optimal given parameters (D^1_g, K^1_g) for gender $g = m, f$. This is how we model comparative advantages in home production. Second, total factor productivity z, that we call *marital bliss*, is a stochastic process that is drawn at the first encounter from some distribution G, and is then infrequently redrawn from the same distribution, independently of past realizations, with probability δ per unit of time. The location of G is a function of male and female characteristics. Hence, male and female education, for example, may generate marital complementarities above and beyond the effect of of education on task specialization.

Couples' home production also involves some intermediate consumption expenditure C, that we assume independent of wages and the marital bliss, but possibly a function of all other male

becomes non-transferable and the equilibrium is much more difficult to solve. We leave this extension to further study.

[9]The corresponding direct utility function is $u(c,e,q)$ is Stone-Geary:

$$u(c,e,q) = q\left(\frac{c - a_0}{1-b}\right)^{1-b}\left(\frac{e - a_1}{b}\right)^b.$$

and female characteristics. Specifically, for a single whose wage is w, the budget to be spent on consumption and leisure is $R = w(1-d)$. For couples, additional transfers t_m, t_f are collected, which can be positive or negative but satisfy the constraint $t_m + t_f = C$, yielding private incomes $R_g = w_g(1-d_g) - t_g$ by gender $g = m, f$. These transfers between partners help finance the public good (say children consumption) and redistribute income between husband and wife.

3.1.3 Resource sharing

Spouses bargain over resources.[10] Conditional on being married for at least the current period and given the current realization z of marital bliss, they choose particular domestic time inputs d_m, d_f and transfers t_m, t_f by Nash bargaining.

We assume no commitment (see Mazzocco 2004, 2007), that is, resources are reallocated whenever there is a change in the environment (here an infrequent bliss shock).[11] Under this assumption, the equilibrium individual values of marriage, denoted $V_g(z)$, $g = m, f$, follow as simple shares of total surplus:

$$B_g \left[V_g(z) - V_g^0\right] = \beta_g S(z),$$

where β_g is the bargaining parameter (with $\beta_m = 1 - \beta_f$) and V_g^0 denotes the values of being single (derived later). The equilibrium surplus $S(z)$ solves the integral equation:

$$(r+\delta)S(z) = qX - B_m r V_m^0 - B_f r V_f^0 + \delta \int \max\left\{S(z'), 0\right\} \, \mathrm{d}G(z'),$$

where $X := R_m - A_m + R_f - A_f$ denotes the household income net of necessary expenditures $A_g := A_g(w_g)$. Note that $qX = B_m U_m + B_f U_f$ measures the aggregate indirect utility of husband and wife in euros. The match surplus is the present value of future match utility flows net of the flow values of remaining single ($B_g r V_g^0$).

Efficient match formation (resp. separation) requires $S(z) > 0$ (resp. $S(z) \leq 0$). The bliss shock is at the heart of both marriage and divorce. If a couple is formed despite being not very well assorted, but just because they initially drew a large z, then it is likely that the next bliss draw will be too low for the match to continue. We view this as a modeling advantage over an alternative model where z is fixed and divorce shocks are exogenous.

Surplus sharing in turn implies income sharing. The Nash bargaining solution induces the following income sharing rule:

$$R_g - A_g = \rho_g(z) X,$$

where $\rho_g(z)$ transforms the Nash bargaining parameter β_g as follows

$$\rho_g(z) qX = B_g r V_g^0 + \beta_g \left[qX - B_m r V_m^0 - B_f r V_f^0\right]. \tag{1}$$

The Nash bargaining model and the Collective framework just differ by an intercept ($B_g r V_g^0$),

[10] On the role of bargaining in explaining task sharing on top of technological and wage changes, see Knowles (2013) who shows that a bargaining model can explain why married men's labor supply has remained high despite decreasing gender wage inequality.

[11] There is no savings and no income shocks. See Voena (2015), Chiappori, Costa Dias, and Meghir (2018) for models in which households resource sharing includes saving decisions.

at least with linear utilities. Collective models parameterize the income share as a function of distribution factors such as spousal wage ratios or local sex ratios (e.g. Chiappori, Fortin, and Lacroix, 2002). Equation (1) links a spouse's market power within the household to her outside option, the value of being single, which varies according to the relative distributions of wages and other individual characteristics in the marriage market (see bellow).

Finally, the Stone-Geary home production function delivers the following equilibrium values for time uses,

$$w_g d_g = w_g D_g^1 + K_g^1 X, \quad g = m, f,$$

yielding corresponding equilibrium values for net income,

$$X = \frac{w_m(1 - D_m^1) - A_m + w_f(1 - D_f^1) - A_f - C}{1 + K_m^1 + K_f^1},$$

and public good,

$$q = zF^1(d_m, d_f) = z \left(\frac{K_m^1 X}{w_m} \right)^{K_m^1} \left(\frac{K_f^1 X}{w_f} \right)^{K_f^1}.$$

3.2 The marriage market

All individual decisions in the preceding section were made conditional on the value of singlehood V_g^0, which cannot be expressed without describing how people meet and match. We now develop the matching model, in the vein of Shimer and Smith (2000).

Time is discrete and the economy is in a steady state — that is to say, all population distributions remain fixed over time. The economy is populated by L_m men and L_f women, characterized by different combinations of wages, cultural identity and education that we call a *type* — indexed by i for males and j for females. Denote $\ell_m(i), \ell_f(j)$ the density functions of male and female types, with $L_m = \int \ell_m$ and $L_f = \int \ell_f$. Let also $n_m(i), n_f(j)$ be the corresponding measures (numbers) of male and female singles of any type, with $N_m = \int n_m$ and $N_f = \int n_f$. Lastly, let $\mu(i,j)$ be the density of couples' exogenous characteristics, with $M = \int \mu$. These distributions are related by the following accounting restrictions,

$$\ell_m = n_m + \int \mu \, \mathrm{d}j, \quad \ell_f = n_f + \int \mu \, \mathrm{d}i.$$

Singles meet according to a random matching process. The flow of new dates of type (i,j) per unit of time is $\lambda n_m n_f$, where λ is a function of (N_m, N_f), specifically: $\lambda(N_m, N_f) = \xi(N_m N_f)^{-1/2}$.[12] Conditional on type, individuals are sampled uniformly. Upon dating, an idiosyncratic quantity of marital bliss z is drawn from the distribution G, given the couple's characteristics. Specifically, we assume that $\ln(z)$ follows a normal distribution with mean $\ln Z$ and variance σ^2. Hence,

$$G(z) = \Phi \left(\frac{\ln z - \ln Z(i,j)}{\sigma} \right).$$

[12]We assume a linearly homogeneous meeting function. If we multiply both n_m and n_f by the same constant k, then $\lambda n_m n_f$ is also multiplied by k. Single men meet single women at rate λN_f and single women meet single men at rate λN_m. Yet, we will call λ simply the meeting rate.

Parameter $Z(i,j)$ (or just Z) is a function of the couple's wages, identities and educations. It is the average match quality and incorporates all sources of marriage externalities that are not embodied in domestic production specializations.[13]

The probability of matching given dating, and given the couple's characteristics, is the probability of drawing a value z delivering positive surplus: $\alpha = G\left(S(z) > 0\right)$. During marriage, bliss shocks accrue infrequently, with probability δ per unit of time, as independent draws from G. A bliss shock induces divorce with probability $1 - \alpha = G\left(S(z) \leq 0\right)$.

In a steady-state equilibrium, the flows in and out of marriage must be exactly balanced:

$$\lambda n_m n_f \alpha = \delta(1-\alpha)\mu.$$

The inflow is the number of dates time the matching probability. The outflow is the number of current matches hit by a marital bliss shock yielding negative match surplus.

By marginalizing this equation, we obtain the equilibrium measures of singles of each type, conditional on the matching probability, as the fixed point of the system:

$$n_m = \frac{\ell_m}{1 + \frac{\lambda}{\delta}\int n_f \frac{\alpha}{1-\alpha}\,\mathrm{d}j}, \quad n_f = \frac{\ell_f}{1 + \frac{\lambda}{\delta}\int n_m \frac{\alpha}{1-\alpha}\,\mathrm{d}i}.$$

The equilibrium distribution of the characteristics of married couples then follows as

$$\mu = \frac{\lambda}{\delta}\frac{\alpha}{1-\alpha} n_m n_f.$$

3.3 The value of being single

Finally, we calculate the values of being single, V_g^0. First, male and female singles choose to devote a time d to domestic chores by solving the optimization problem:

$$U_g^0 = \max_d F_g^0(d)\frac{w_g(1-d) - A_g}{B_g},$$

yielding solution for domestic time:

$$w_g d_g^0 = w_g D_g^0 + K_g^0 X_g^0,$$

net income:

$$X_g^0 := w_g(1-d_g^0) - A_g = \frac{w_g(1-D_g^0) - A_g}{1 + K_g^0},$$

home production:

$$q_g^0 = F_g^0(d_g^0) = \left(\frac{K_g^0 X_g^0}{w_g}\right)^{K_g^0},$$

and private utility in euros: $B_g U_g^0 = q_g^0 X_g^0$.

[13] The matrix $[Z(i,j)]$ is similar to the affinity matrix of Dupuy and Galichon (2014), generalizing Choo and Siow (2006). Our structural search-matching model is however more flexible and richer than Dupuy and Galichon's as sorting depends on $Z(i,j)$ but also on household time use externalities in public good production. An attempt at incorporating labor supply in the Choo-Siow framework can be found in Choo, Seitz, and Siow (2008).

Let $\overline{S} = \int \max\{S(z), 0\} \, \mathrm{d}G(z)$ be the expected match surplus. The present value of being single for each gender is

$$B_m r V_m^0 = q_m^0 X_m^0 + \lambda \beta_m \int \overline{S} n_f \, \mathrm{d}j, \quad B_f r V_f^0 = q_f^0 X_f^0 + \lambda \beta_f \int \overline{S} n_m \, \mathrm{d}i.$$

These equilibrium definitions of the values of singlehood summarize the trade-off individuals need to solve. While remaining single, flow values $B_g U_g^0 = q_g^0 X_g^0$ keep been added, until a meeting occurs (at rate λ). Actual match formation requires a bliss shock z such that $S(z) > 0$ (implicit in the definition of $\overline{S}$). The relative distributions of singles' types (distributions n_g) in the marriage market are equilibrium objects, which are determined by exogenous population distributions and the particular matching process at work.

3.4 Equilibrium solution

Let us simplify the economy to a population of same-sex individuals who can be matched (married) or unmatched (single). Moreover, they only differ ex ante by their wage w. There is also no home production and transfers add up to zero ($C = 0$). The public good is therefore simply $q = z$ (the bliss variable) for couples and $q = 1$ for singles. Lastly, the bargaining parameter is $\beta_m = \beta_f = \frac{1}{2}$.

Three equations describe this simpler equilibrium:

- The surplus of a match (w, w', z):

$$\begin{aligned}(r+\delta) S(w, w', z) = zX - B(w) r V^0(w) - B(w') r V^0(w') \\ + \delta \int \max\left\{S(w, w', z'), 0\right\} \, \mathrm{d}G(z'), \quad (2)\end{aligned}$$

  where $X = w + w' - A(w) - A(w')$;

- The equilibrium distribution of wages for singles:

$$n(w) = \frac{\ell(w)}{1 + \frac{\lambda}{\delta} \int n(w') \frac{\alpha(w,w')}{1-\alpha(w,w')} \, \mathrm{d}w'} \quad (3)$$

  where $\alpha(w, w') = G\left(S(w, w', z) > 0\right)$;

- The values for singles:

$$B(w) r V^0(w) = w - A(w) + \frac{\lambda}{2} \int \overline{S}(w, w') n(w') \, \mathrm{d}w', \quad (4)$$

  where $\overline{S}(w, w') = \int \max\{S(w, w', z), 0\} \, \mathrm{d}G(z)$.

The surplus functions $S(w, w', z)$ and $\overline{S}(w, w')$ are symmetric in (w, w'). Plugging (4) in (2), we can simplify the surplus equation as

$$(r+\delta) S(w, w', z) = (z-1) X - \lambda \int \overline{S}(w, w') n(w') \, \mathrm{d}w' + \delta \overline{S}(w, w').$$

By integrating over z, it follows that

$$(r+\delta)\overline{S}(w,w') = \int \max\left\{(z-1)X - \lambda \int \overline{S}(w,w')n(w')\,\mathrm{d}w' + \delta\overline{S}(w,w'), 0\right\} \mathrm{d}G(z). \tag{5}$$

Moreover,

$$\alpha(w,w') = 1 - G\left(X + \lambda \int \overline{S}(w,w')n(w')\,\mathrm{d}w' - \delta\overline{S}(w,w')\right). \tag{6}$$

Hence, the equilibrium is characterized by two function: $\overline{S}(w,w')$ and $n(w)$, solving two integral equations: (5) and (3) with $\alpha(w,w')$ as in equation (6). The only input is the exogenous wage distribution $\ell(w)$.

In practice, $\ell(w)$ can be estimated using a kernel density estimator. We use a quadrature rule to approximate the integrals. The functional equations can thus be discretized at the quadrature nodes. We use the Clenshaw-Curtis quadrature and Chebychev nodes, because the quadrature weights can be easily and fast calculated using Fast Fourier Transform (see Trefethen (2008) and Waldvogel (2006) and the computation appendix of Goussé, Jacquemet, and Robin (2017a)).

4 Estimation strategy

The estimation of the model is done in three steps, summarized below (see Appendix A for a detailed description).

Step 1. Distribution of exogenous characteristics

We first pool years of data to increase the sample size. Namely, we consider the following four subsamples (denoted by τ): 1992-1998, 1999-2005, 2006-2012, 2013-2019. For each time period and each region, we first estimate the meeting rate parameter ξ (in $\lambda = \xi(N_m N_f)^{-1/2}$) as well as the bliss shock frequency δ by fitting new marriage flows and divorce flows. We also estimate distributions μ, n_m, n_f on a Chebychev grid of types (wages, education and identity) by kernel density estimation. Finally, we estimate (nonparametrically) the matching probability α from these distributions as

$$\alpha = \frac{\mu}{\mu + \frac{\lambda}{\delta} n_m n_f}.$$

Step 2. Estimation of preference and home production parameters

We pool all data (across years, regions, marital status and gender) and we estimate preference and home production parameters by nonlinear least squares. We allow for observed heterogeneity in preference parameters in the following way:

- a_0, a_1, D^1, D^0: gender, a linear time trend, education, identity;
- b: gender, education, identity;
- K^1, K^0: gender;
- C: education and identity of both spouses.

The time use equations depend linearly on the sharing rule $\rho_g(z)$. We can thus linearly aggregate the labor supplies of married men and women and remove this dependence. It happens that this aggregation does not prevents us from identifying preference and home production parameters from time use observations. Hence, the estimation of private preferences and home production functions is robust to equilibrium misspecification. The individual levels of married individuals' paid hours identify the bargaining parameter β and the bliss shock variance in the third stage of the estimation procedure.

Specifically, for given preference and home production parameters, we start by calculating the following residuals:

- For couples:

$$u_1 = \frac{1}{b_m} w_m(e_m - a_{1m}) + \frac{1}{b_f} w_f(e_f - a_{1f}) - X^1,$$
$$u_{2g} = w_g(d_g - D_g^1) - K_g^1 X^1, \quad g \in \{m, f\},$$

  with total household expenditure

$$X = \frac{w_m(1 - D_m^1 - a_{1m}) + w_f(1 - D_f^1 - a_{1f}) - a_{0m} - a_{0f} - C}{1 + K_m^1 + K_f^1}.$$

- For singles:

$$u_{3g} = \frac{1}{b_g} w_g(e_g^0 - a_{1g}) - X_g^0, \quad g \in \{m, f\},$$
$$u_{4g} = w_g(d_g^0 - D_g^0) - K_g^0 X_g^0, \quad g \in \{m, f\},$$

  with total expenditure

$$X_g^0 = \frac{w_g(1 - D_g^0 - a_{1g}) - a_{0g}}{1 + K_g^0}.$$

So, all residuals are expressed in euros and we introduce the income effect parameters b_m, b_f in a way that preserves gender symmetry.[14]

There is too few data for so many interactions and fine dependence to exogenous characteristics. In practice we keep all data, and we weight observations by a smooth kernel indicator of the distance of a particular observation to the particular combination of period, education and/or identity for which we want to estimate a parameter. In order to estimate parameters $\theta = (a, b, D, K, C)$ for a given value (i_0, j_0) on the grid of male and female wages, education and

[14]The proof that these seven residuals suffice to identify all preference and home production parameters is tedious but relatively straightforward. Identification can be verified empirically by checking the non singularity of the Jacobian.

identity, we minimize the least-square criterion:

$$\sum_{\text{male single } i} K(i-i_0)\left(u_{3i}^2+u_{4i}^2\right)+\sum_{\text{female single } j} K(j-j_0)\left(u_{3i}^2+u_{4i}^2\right)$$
$$+\sum_{\text{couple } (i,j)} K(i-i_0)K(j-j_0)\left(u_{1ij}^2+u_{2mij}^2+u_{2fij}^2\right).$$

Step 3. Estimation of match quality Z, bargaining coefficient β and bliss shock variance σ

This step is a lot more involved as it makes use of the structural model. Basically, we again use non linear least squares to estimate Z, β, σ with the following conditionality to exogenous variables:

- Z: wage, education and identity of both spouses,
- β: education and identity,
- σ: constant.

Three least square residuals are constructed so as to fit:

- The matching probability function α,
- Married male or female expected leisure,
- Married male or female leisure variance.

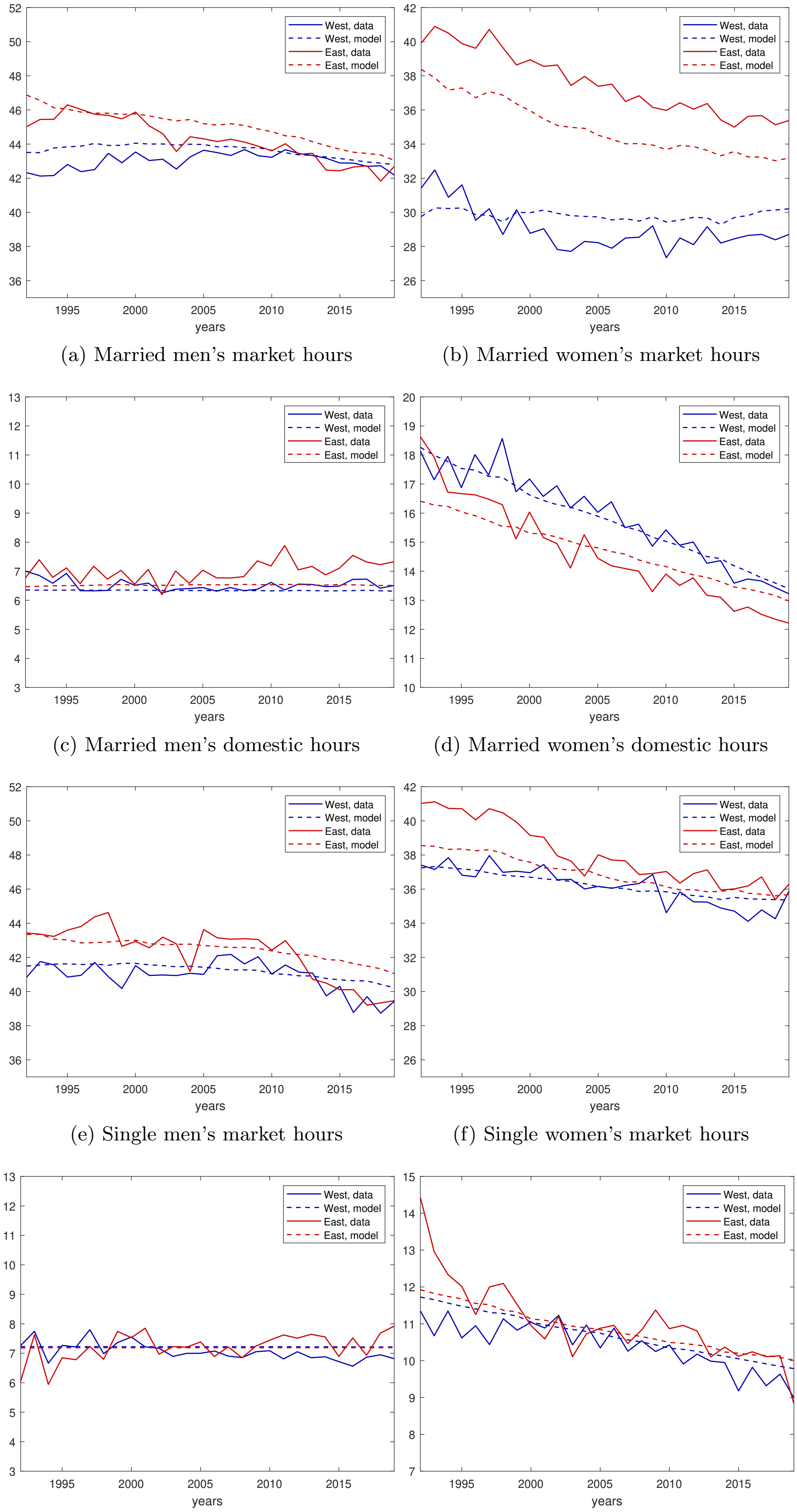
The matching probability identifies the average match quality Z. In contrast with Step 2, in which we fit the aggregate leisure or labor supply of couples, Step 3 fits the distribution of individual leisure to identify and estimate the Nash bargaining coefficient β and the dispersion of marital bliss shocks σ.

5 Fit of the model

The point estimates of the parameters are difficult to interpret. In this section, we rather check the ability of the estimated model to replicate both the East-West differences and the time trends of the main relevant outcomes.

We first show how we fit average time uses in the East and West samples (Figure 7). Although preferences do not depend specifically on the region, the way we model heterogeneity and time trends allows us to fit the main differences between German regions across time and gender well. However, the model underestimates female labor supply in the East (by about 2 hours out of 40) and overestimates it in the West. Notice also that the model reproduces the differences between married and unmarried women, although private preferences are independent of marital status (in contradistinction to public good production).

Then, we ask if the model reproduces the main features of couples' income and wage inequality (Figure 8). In doing so, we start interrogating the model's ability to fit sorting. We fit within-couple wage inequality, as measured by the female wage share $w_f/(w_m+w_f)$, very well.



(a) Married men's market hours

(b) Married women's market hours

(c) Married men's domestic hours

(d) Married women's domestic hours

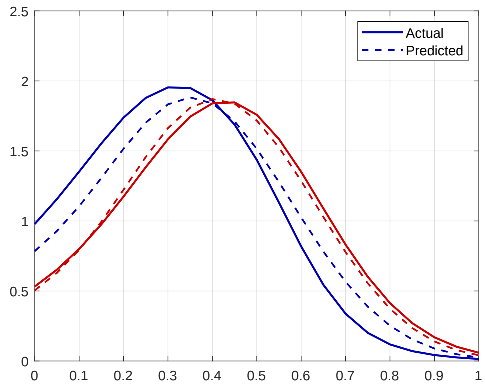
(e) Single men's market hours

(f) Single women's market hours

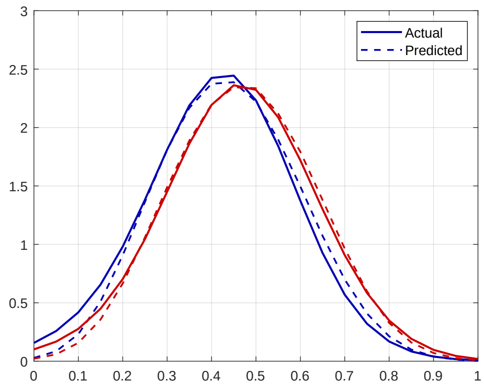
(g) Single men's domestic hours

(h) Single women's domestic hours

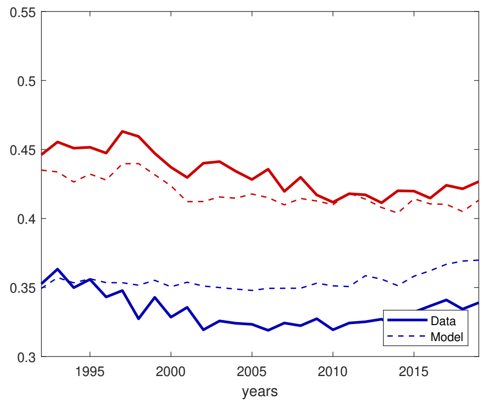
Figure 7: Fit of time uses



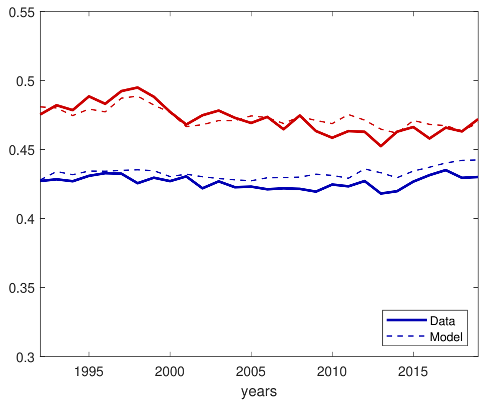
(a) Income share distribution (2013-19)



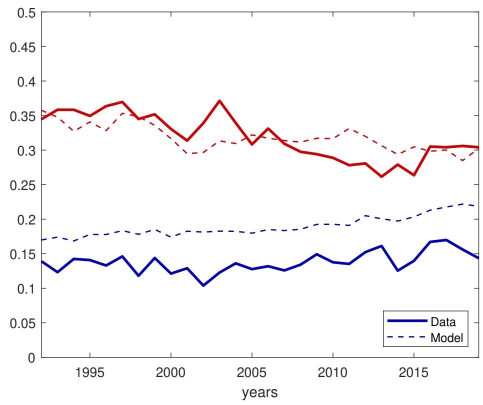
(b) Wage share distribution (2013-19)



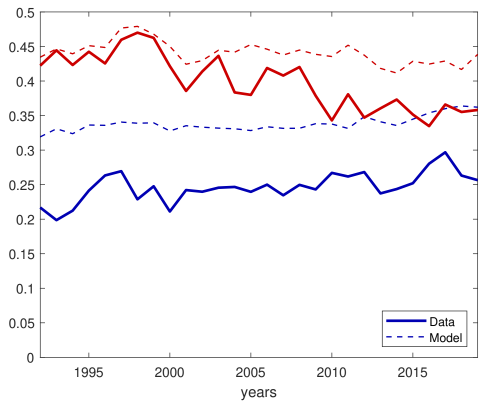
(c) Average income share over time



(d) Average wage share over time



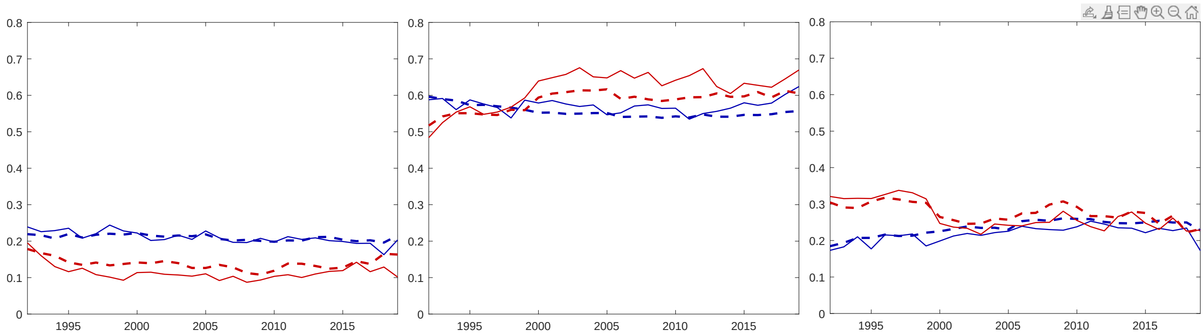
(e) Couples with income share > 0.5



(f) Couples with wage share > 0.5

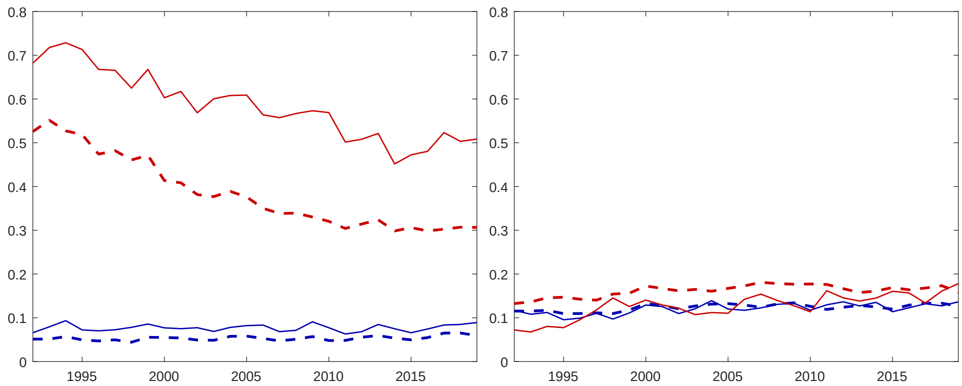
Notes: The income share is $\frac{w_f h_f}{w_f h_f + w_m h_m}$. The wage share is $\frac{w_f}{w_f + w_m}$.

Figure 8: Fit of within couple inequality

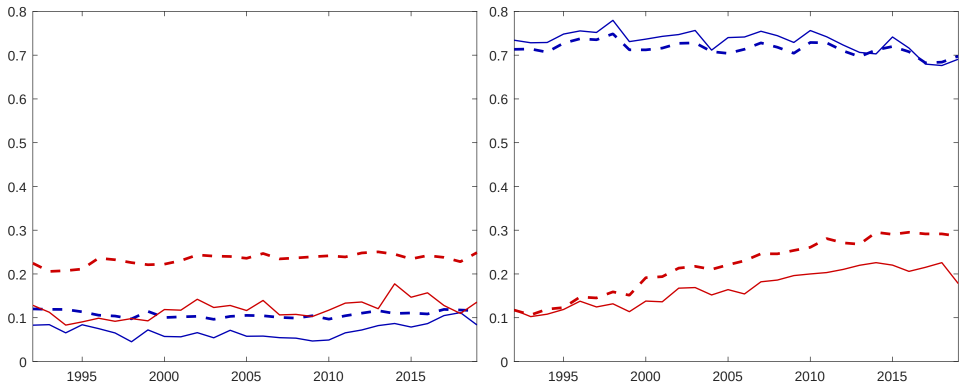


(a) Fraction female < male (b) Fraction female = male (c) Fraction female > male

Figure 9: Fit of sorting by education



(a) Fraction male low - female low (b) Fraction male low - female high



(c) Fraction male high - female low (d) Fraction male high - female high

Figure 10: Fit of sorting by identity

We do get that the East is a lot more egalitarian, with about 40% couples such that the wife is paid more by the hour than her husband (25% in the West, slowly rising). We also reproduce the property, common to both regions, that labor supply makes the distribution of income (wage times hours worked) more unequal than wages.

Lastly, we consider sorting on education and cultural identity. Figure 9 summarizes education sorting by the probability that the wife has less, equal, or more education than her husband. First, we note that these sorting statistics are considerably more stable and similar across regions than education levels (Figure 2). Second, the model replicates them very well; in particular the tendency for female spouses in the East to be more educated than their husbands. Figure 10 shows identity sorting roughly measured by the relative positions of male and female identity indices with respect to the 0.5 probability threshold of living in the East. The model does not quite succeed in reproducing the extent of cultural homogamy in East Germany. This is hardly surprising given that we summarize individual heterogeneity by just three dimensions (education, wages and the identity index).

We end this section by providing a global measure of the fit of the distribution of characteristics by marital status using the Kullback-Leibler divergence. For each individual, we observe a set of characteristics — i for males, j for females. Some individuals are matched in couples, and others are single. The vector of characteristics (i or j) is comprised of education — which is discrete — and of the wage and the identity index — which are continuous. The distribution of individual characteristics of men and women in the population (or the sample) is described by $\{n_m(i); n_f(j); 2\mu(i,j)\}$ where the number of couples $\mu(i,j)$ is multiplied by 2 to count the individuals in the household. Then,

$$\int n_m + \int n_f + 2\iint \mu = N_m + N_f + 2M = L_m + L_f = L$$

is the total population size. Let $\mathcal{D} = \left(\frac{n_m(i)}{L}, \frac{n_f(j)}{L}, 2\frac{\mu(i,j)}{L}\right)$ denote the population distribution normalized by its size.

We use these distributions to assess the ability of the model to fit sorting in two ways. First, we compare the actual distribution of types to its prediction by the model. Second, we build an upper bound for the divergence between actual and predicted distributions, where the distribution of types is predicted assuming a matching probability equal to its mean value:

$$n_m = \frac{\ell_m}{1 + \frac{\lambda}{\delta} N_f \frac{\overline{\alpha}}{1-\overline{\alpha}}}, \quad n_f = \frac{\ell_f}{1 + \frac{\lambda}{\delta} N_m \frac{\overline{\alpha}}{1-\overline{\alpha}}}, \quad \mu = \frac{\lambda}{\delta} \frac{\overline{\alpha}}{1-\overline{\alpha}} n_m n_f.$$

In both cases, we measure the divergence of a predicted distribution $\mathcal{D}'$ ($\mathcal{D}' = \widehat{\mathcal{D}}$, the model's prediction, or $\mathcal{D}' = \widehat{\mathcal{D}}_{\text{rm}}$, the random matching model) from the actual one $\mathcal{D}$, using the Kullback-Leibler divergence:

$$d_{KL}(\mathcal{D} \,||\, \mathcal{D}') = 2\iint \frac{\mu(i,j)}{L} \ln \frac{\mu(i,j)/L}{\mu'(i,j)/L'} \,\mathrm{d}i\,\mathrm{d}j + \int \frac{n_m(i)}{L} \ln \frac{n_m(i)/L}{n'_m(i)/L'} \,\mathrm{d}i + \int \frac{n_f(j)}{L} \ln \frac{n_f(j)/L}{n'_f(j)/L'} \,\mathrm{d}j.$$

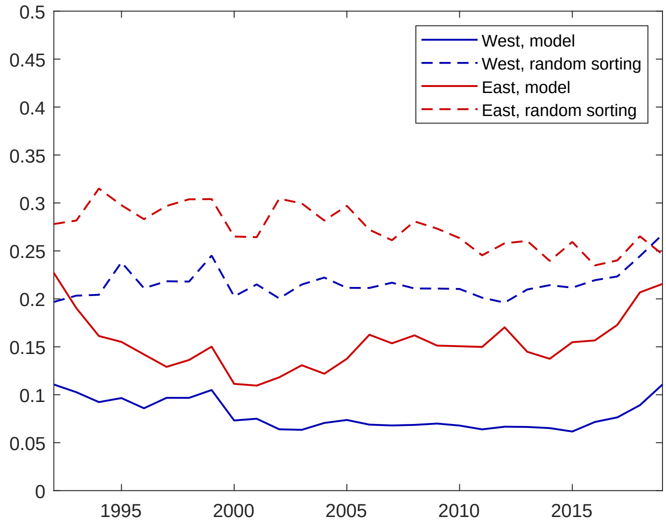


Figure 11: Model and random matching versus data

The results for both alternative distributions are presented in Figure (11) . We see that the fit is better for the West than for the East (lower divergence), which is not surprising given that there are six times more observations in the West sample than in the East sample. Hence the estimation procedure favors the western sample over the eastern one. The estimated model does also much better than random matching, which indicates that there is a significant amount of sorting.

6 Comparing distributions across regions

The estimated model manages to fit the distribution of wage and income shares (8) without conditioning on the region. The only exogenous variation across regions comes from the distributions of individual wages, education and cultural identity by gender. We now ask which of wages, education and culture matters more. To this end, we rely on simulation exercises in which the model is used to predict the distribution of characteristics of singles and couples, as well as their time uses, in a counterfactual world in which the distribution of some exogenous characteristics in the West is replaced by the corresponding one in the East (preserving the differences between men and women), while all other characteristics are set to their observed values in both regions. The distance to the actual distribution provides a measure of the marginal effect of the corresponding distribution of characteristics.[15]

6.1 Sorting

We start by measuring the observed differences in sorting, that is by marital status, between East and West Germany, and we ask how much of this difference is separately explained by education, wages and cultural identity.

Again we use the KL divergence $d_{KL}(\widehat{\mathcal{D}}_W \,||\, \widehat{\mathcal{D}}_{alt})$, where we compare the model prediction for West Germany $\widehat{\mathcal{D}}_W$ to various alternative distributions $\widehat{\mathcal{D}}_{alt}$. The benchmark comparison is

[15]This is a similar exercise as in DiNardo, Fortin, and Lemieux (1996).

Period	Data	Benchmark	Counterfactuals (West||counterfactual West)					
	(West||East)	(Predicted)	Culture	Educ	Wages	Index + wages	Culture + educ	Educ + wages
1992-1998	2.96	2.62	1.36	1.88	1.86	0.55	0.60	1.38
1999-2005	1.65	1.60	0.79	1.13	1.16	0.33	0.33	0.88
2006-2012	1.32	1.29	0.61	0.95	0.91	0.25	0.28	0.74
2013-2019	0.98	0.90	0.35	0.73	0.71	0.17	0.19	0.61
All		100%	39%	81%	79%	19%	0%	68%

Table 4: KL-divergence of West from East (benchmark and counterfactuals)

$\widehat{\mathcal{D}}_{alt} = \widehat{\mathcal{D}}_E$, the model prediction for East Germany. For example, the counterfactual distribution in which we change the marginal distribution of education and wages is

$$\begin{aligned}\widehat{\mathcal{D}}_{alt}(index, educ, w) &= \widehat{\mathcal{D}}_E(index \mid educ, w)\widehat{\mathcal{D}}_W(educ, w) \\ &= \widehat{\mathcal{D}}_E(index, educ, w)\frac{\widehat{\mathcal{D}}_W(educ, w)}{\widehat{\mathcal{D}}_E(educ, w)}.\end{aligned}$$

Note that for this to work, we need identical distribution supports in both regions.

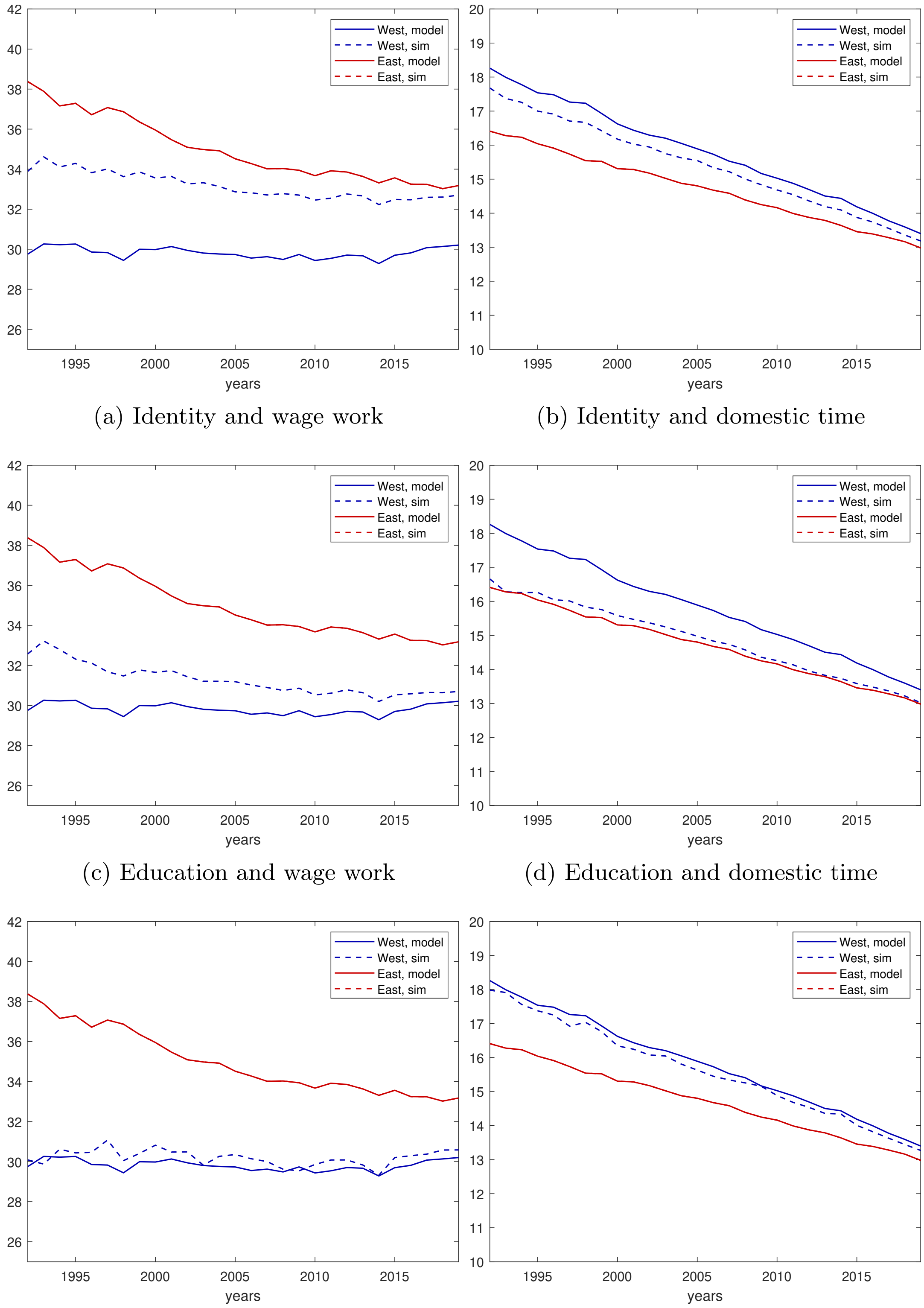
Table 4 displays the values of the KL-divergence of West from East (benchmark and counterfactuals) at different points in time. We see that the benchmark divergence gets smaller over time, indicating convergence. The table also shows the KL-divergence between East and West distributions over time for both the actual distributions (data) and the ones predicted by the model (benchmark). It is reassuring to check that both actual and predicted divergences coincide.

Changing the marginal distribution of cultural identity reduces the divergence by about 50%, whereas education and wages reduce it by only 25%. Then, changing culture and education, or culture and wages both reduce the divergence by another 50%, whereas education and wages together reduce the divergence much more marginally. We draw two conclusions. First culture (preferences) matters more than education and wages for explaining differences in sorting on characteristics between East and West Germany. Second, identity and education and identity and wages show complementarity. The total divergence reduction is approximatively the sum of each marginal reduction (50% + 25%). In contrast, education and wages seem to be a lot more substitutable, the total reduction being significantly less that the sum of the two marginal ones (37% instead of 24% + 23% = 47%).

6.2 Time uses

Having established strong sorting with respect to all characteristics, we now ask which labor outcome is determined by which characteristic. In Figure 12, we display the predicted trends of time uses in East and West Germany (solid red and blue lines) and the predicted trends replacing the marginal distribution of a characteristic in the West by its value in the East, separately for each characteristic (dotted lines). If a characteristic is a strong determinant of a time use, then we should see the dotted line moving close to the red, solid line.

Identity strongly determines labor supply. Education strongly determines domestic time.

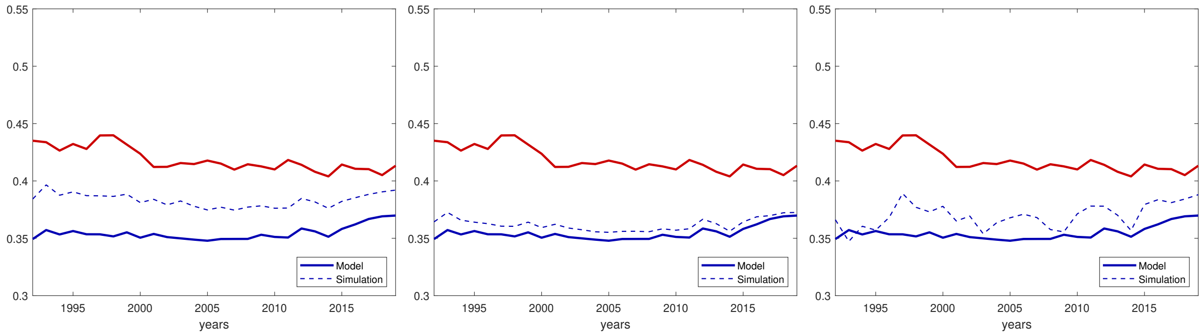


(a) Identity and wage work (b) Identity and domestic time

(c) Education and wage work (d) Education and domestic time

(e) Wages and wage work (f) Wages and domestic time

Figure 12: Marginal effects of socioeconomic characteristics on time uses



(a) Identity (b) Education (c) Wages

Figure 13: Marginal effects of socioeconomic characteristics on income share

Education determines labor supply, but a lot less than identity. Identity determines domestic time, but a lot less than education. Wages barely change time uses. This is surprising as wage levels are very different in the East and in the West. This does not seem to be a consequence of the linear demand system, as replacing the linear specification of $A(w)$ by a quadratic function did not change this estimation.

Lastly, what is the contribution of each characteristic to the income share? From what we just showed, we expect identity to move labor supply and therefore affect the income share. We also expect wages to mechanically change the income share as wages are more gender-equal in the East. Figure 13 shows that it is indeed the case.

7 Conclusion

In this paper, we develop a search-matching model of marriage formation and household task sharing to explain differences between East and West Germany after reunification. We measure the importance of education, culture and wages for marriage sorting and time uses by marital status. We find that a common trend (in meeting rates and preferences/home production) is sufficient to capture the 1992-2019 dynamics of sorting and time uses. We show that culture is the most important driver of marriage sorting by far. Culture also explains most of the differences in female labor supply. Education explains differences in female domestic hours. Wages explain nothing of time use differences. Overall, differences in relative income are explained by identity through its effect on labor supply.

Our general conclusion is that there is no reason to believe that the 40 years of communism in East Germany specifically changed gender norms in the family. Of course, individual cultural attitudes have been profoundly affected, including certain attitudes that affect the family. For example, fewer West Germans (especially women) think that it is important to have success in the job or to have children. However, what we seem to find is that relative income distributions differ between East and West Germany essentially because husbands and wives in the East and in the West have different cultural attitudes, not because same-type couples would behave differently. Moreover, marriage sorting differs not because affinity rules are different, but simply because the populations available for marriage are different.

We take it as a remarkable finding that the communist regime, to a first approximation, changed individuals but not the rules of marriage, nor the rules of intra-household resource sharing.

References

Alesina, A., and N. Fuchs-Schündeln (2007): "Goodbye Lenin (or Not?): The Effect of Communism on People's Preferences," *American Economic Review*, 97(4), 1507–1528.

Bau, N., and R. Fernández (2021): "Culture and the Family," NBER Working Papers 28918, National Bureau of Economic Research, Inc.

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.

——— (1974): "A Theory of Marriage: Part II," *Journal of Political Economy*, 82(2), S11–S26.

Bertrand, M., E. Kamenica, and J. Pan (2015): "Gender Identity and Relative Income within Households," *Quarterly Journal of Economics*, 130(2), 571–614.

Bursztyn, L., T. Fujiwara, and A. Pallais (2017): "'Acting Wife': Marriage Market Incentives and Labor Market Investments," *American Economic Review*, 107(11), 3288–3319.

Cavapozzi, D., M. Francesconi, and C. Nicoletti (2021): "The impact of gender role norms on mothers' labor supply," *Journal of Economic Behavior & Organization*, 186(C), 113–134.

Chiappori, P.-A. (1988): "Rational Household Labor Supply," *Econometrica*, 56(1), 63–90.

Chiappori, P.-A., M. Costa Dias, and C. Meghir (2018): "The Marriage Market, Labor Supply and Education Choice," *Journal of Political Economy*, 126, S26–S72.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002): "Marriage market, divorce legislation, and household labor supply," *Journal of Political Economy*, 110(1), 37–72.

Choo, E., S. Seitz, and A. Siow (2008): "Marriage matching, risk sharing and spousal labor supplies," *University of Toronto WP*, (332).

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114(1), 175–201.

Ciscato, E. (2021): "The changing wage distribution and the decline of marriage," Discussion paper, mimeo.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–1044.

Dupuy, A., and A. Galichon (2014): "Personality Traits and the Marriage Market," *Journal of Political Economy*, 122(6), 1271–1319.

Fernández, R. (2013): "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 103(1), 472–500.

Goussé, M., N. Jacquemet, and J.-M. Robin (2017a): "Marriage, labor supply, and home production," *Econometrica*, 85(6), 1873–1919.

Goussé, M., N. Jacquemet, and J.-M. Robin (2017b): "Household labour supply and the marriage market in the UK, 1991-2008," *Labour Economics*, 46, 131 – 149.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2016): "Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation," *American Economic Journal: Macroeconomics*, 8(1), 1–41.

Greenwood, J., N. Guner, and G. Vandenbroucke (2017): "Family Economics Writ Large," *Journal of Economic Literature*, 55(4), 1346–1434.

Holzner, C., and B. Schulz (2022): "Marriage and Divorce under Labor Market Uncertainty," Discussion paper, mimeo.

Knowles, J. A. (2013): "Why are Married Men Working So Much? An Aggregate Analysis of Intra-Household Bargaining and Labour Supply," *Review of Economic Studies*, 80(3), 1055–1085.

Lippmann, Q., A. Georgieff, and C. Senik (2020): "Undoing Gender with Institutions. Lessons from the German Division and Reunification," *Economic Journal*, 130(3291), 1445–1470.

Manser, M., and M. Brown (1980): "Marriage and household decision-making: A bargaining analysis," *International Economic Review*, 21(1), 31–44.

Mazzocco, M. (2004): "Saving, Risk Sharing, and Preferences for Risk," *American Economic Review*, 94(4), 1169–1182.

——— (2007): "Household Intertemporal Behaviour: A Collective Characterization and a Test of Commitment," *Review of Economic Studies*, 74(3), 857–895.

McElroy, M. B., and M. J. Horney (1981): "Nash-bargained household decisions: Toward a generalization of the theory of demand," *International Economic Review*, 22(2), 333–349.

Schmelz, K. (2021): "Enforcement May Crowd out Voluntary Support for COVID-19 Policies, Especially Where Trust in Government Is Weak and in a Liberal Society," *Proceedings of the National Academy of Sciences*, 118(1), e2016385118.

Shimer, R., and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68(2), 343–369.

Trefethen, L. N. (2008): "Is Gauss Quadrature Better than Clenshaw–Curtis?," *SIAM Review*, 50(1), 67–87.

Voena, A. (2015): "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?," *American Economic Review*, 105(8), 2295–2332.

Waldvogel, J. (2006): "Fast Construction of the Fejér and Clenshaw-Curtis Quadrature Rules," *BIT Numerical Mathematics*, 46, 195–202.

Appendix

A Estimation of the model

The data provides the following information for each observation unit:

- For couples:
  - the region (East or West Germany);
  - the ID of the husband i, the ID of the wife j and the year t;
  - leisure hours e^1_{it} and e^1_{jt};
  - hours of housework d^1_{it} and d^1_{jt};
  - wages w_{it} and w_{jt};
  - exogenous characteristics x_{it} and x_{jt}, comprising in our application education and the cultural identity index.
- For singles:
  - the region (East or West Germany);
  - the ID of the individual i, the year t and the gender g;
  - the endogenous and exogenous variables $e^0_{it}, d^0_{it}, w_{it}, x_{it}$.

A.1 Type distributions

Step 1 closely follows the estimation of Kernel densities described in Goussé, Jacquemet, and Robin (2017a), to which we refer the interested reader. The two next steps differ.

A.2 Preference and household production

A.2.1 Residuals

For any couple's observation (i, j, t) and any value of the parameters, consider the residuals:

$$\begin{aligned}
u_{1ijt} &= \frac{1}{b_m} w_{it} \left[e^1_{it} - a_{1m}\right] + \frac{1}{b_f} w_{jt} \left[e^1_f - a_{1f}\right] - \left(1 - k^1_m - k^1_m\right) Y^1_{ijt} \\
u_{2mit} &= w_{it} \left[d^1_{it} - D^1_m\right] - k^1_m Y^1_{ijt} \\
u_{2fjt} &= w_{jt} \left[d^1_{jt} - D^1_f\right] - k^1_f Y^1_{ijt}
\end{aligned}$$

where $k^1_m = \frac{K^1_m}{1+K^1_m+K^1_f}$, $k^1_f = \frac{K^1_f}{1+K^1_m+K^1_f}$ and

$$Y^1_{ijt} = w_{it} \left[1 - D^1_m - a_{1m}\right] + w_{jt} \left[1 - D^1_f - a_{1f}\right] - a_{0m} - a_{0f} - C.$$

For any single's observation (i, t, g):

$$\begin{aligned}
u_{3gt} &= \frac{1}{b_g} w_{it} \left[e^0_{it} - a_{1g}\right] - \left(1 - k^0_g\right) Y^0_{it} \\
u_{4gt} &= w_{it} \left[d^0_{it} - D^0_g\right] - k^0_g X^0_{it}
\end{aligned}$$

where $k^0_g = \frac{1}{1+K^0_g}$ and

$$Y^0_{it} = w_{it} \left[1 - D^0_g - a_{1g}\right] - a_{0g}.$$

A.2.2 Parameters and smoothed sum of squared errors

We want to estimate preference and home production parameters that vary with time and exogenous characteristics in a flexible way. For any pre-specified time grid τ and a grid of exogenous characteristics (x_g), which may differ by gender g, we estimate the following parameters subject to the following specific variations:

- $a_{0g}(\tau, x_g), a_{1g}(\tau, x_g), D_g^1(\tau, x_g), D_g^0(\tau, x_g)$ by gender, time, education and identity,
- $1/b_g(x_g)$ by gender, education and identity,
- $C(x_m, x_f)$ by spousal educations and identities,
- k_g^1, k_g^0 by gender.

We seek to minimize the sum of squared errors

$$SSE = \sum_{\tau, x_m, x_f} \sum_{couple\, i,j,t} \omega_{ijt}^1(\tau, x_m, x_f) \left[(u_{1ijt})^2 + (u_{2mit})^2 + (u_{2fjt})^2 \right] \\ + \sum_g \sum_{\tau, x_g} \sum_{single\, i,t,g} \omega_{itg}^0(\tau, x_g) \left[(u_{3git})^2 + (u_{4gjt})^2 \right],$$

where residuals are calculated at the parameter values determined by the particular grid point in the first summations and where

$$\omega_{ijt}^1(\tau, x_m, x_f) = \phi_h(t - \tau)\phi_h(x_{it} - x_m)\phi_h(x_{jt} - x_f)$$
$$\omega_{itg}^0(\tau, x_g) = \phi_h(t - \tau)\phi_h(x_{it} - x_g)$$

are smooth, fuzzy classification indicators. In the simulation, we will assign parameters to observations according to

$$\max_{\tau, x_m, x_f} \omega_{ijt}^1(\tau, x_m, x_f), \max_{\tau, x_g} \omega_{itg}^0(\tau, x_g),$$

choosing the (τ, x_m, x_f) closest to (t, x_{it}, x_{jt}) for couples and (τ, x_g) closest to (t, x_{it}) for singles of gender g.

A.2.3 Concentration

Note that the problem is linear conditional on k_g^1, k_g^0 and $b_g(x)$. The other parameters can thus be easily estimated by OLS conditional on k_g^1, k_g^0 and $b_g(x)$. For this, create a (sparse) matrix M that has each row indexed by

- couples: $(n, i, j, t, \tau, x_m, x_f)$ where $n = 1, 2, 3$ indexes the residual,
- singles: $(n, i, t, g, \tau, x_g, n_u)$ where $n = 1, 2$ indexes the residual,

(pooling both regions) and where the first column contains the weights ω (say), the second column is the dependent variable, and the other columns are the regressors associated to the parameters:

$$a_{0g}(\tau, x_g), a_{1g}(\tau, x_g), D_g^1(\tau, x_g), D_g^0(\tau, x_g), C(x_m, x_f), g = m, f.$$

Defining this matrix as sparse will save memory and speed up the OLS calculation. These conditional estimates can then be plugged into the SSE which minimization is now a lot more manageable.

A.3 Matching quality, bargaining coefficient and bliss shock

A.3.1 Matching probability

The match surplus solves

$$(r+\delta)S(z) = zF^1X^1 - B_m r V_m^0 - B_f r V_f^0 + \delta\overline{S},$$

where $\overline{S} = \int \max\{S(z), 0\}\, \mathrm{d}G(z)$ and the equilibrium F^1 and X^1 were derived in Section 3.1.3. The matching probability follows as

$$\begin{aligned}
\alpha &= \Pr\left\{ z > \frac{B_m r V_m^0 + B_f r V_f^0 - \delta\overline{S}}{F^1X^1} \right\} \\
&= 1 - G\left(\frac{B_m r V_m^0 + B_f r V_f^0 - \delta\overline{S}}{F^1X^1} \right) \\
&= 1 - \Phi\left[\frac{1}{\sigma} \ln\left(\frac{B_m r V_m^0 + B_f r V_f^0 - \delta\overline{S}}{ZF^1X^1} \right) \right], \qquad (7)
\end{aligned}$$

and the average surplus is

$$\begin{aligned}
\overline{S} &= \int \max\{S(z), 0\}\, \mathrm{d}G(z) \\
&= \frac{1}{r+\delta} ZF^1X^1\, \mathcal{G}\left(\frac{B_m r V_m^0 + B_f r V_f^0 - \delta\overline{S}}{ZF^1X^1} \right) \qquad (8)
\end{aligned}$$

with

$$\mathcal{G}(s) = \int (z-s)^+\, \mathrm{d}\Phi\left(\frac{\ln z}{\sigma}\right) = -s\Phi\left(-\frac{\ln s}{\sigma}\right) + e^{\frac{\sigma^2}{2}}\Phi\left(-\frac{\ln s}{\sigma} + \sigma\right).$$

Note that $\mathcal{G}'(s) = -\Phi\left(-\frac{\ln s}{\sigma}\right) = -1 + \Phi\left(\frac{\ln s}{\sigma}\right)$ and that

$$\frac{\mathrm{d}\overline{S}}{\mathrm{d}\left(B_m r V_m^0 + B_f r V_f^0\right)} = -\frac{\alpha}{r + (1-\alpha)\delta} < 0.$$

Hence, equations (7) and (8) define a one-to-one relationship between α, $B_m r V_m^0 + B_f r V_f^0$ and $\overline{S}$.

Inverting equation (7) yields

$$\frac{B_m r V_m^0 + B_f r V_f^0 - \delta\overline{S}}{ZF^1X^1} = \exp\left[\sigma\Phi^{-1}(1-\alpha)\right],$$

where

$$F^1 = F^1(d_m^1, d_f^1) = \left(\frac{K_m^1 X^1}{w_m}\right)^{K_m^1} \left(\frac{K_f^1 X^1}{w_f}\right)^{K_f^1}.$$

Plugging this expression into equation (8) then implies that

$$\frac{\delta\overline{S}}{ZF^1X^1} = \frac{\delta}{r+\delta}\mathcal{G}\left[\exp\left[\sigma\Phi^{-1}(1-\alpha)\right]\right],$$

and

$$\frac{B_m r V_m^0 + B_f r V_f^0}{ZF^1X^1} = \frac{\delta}{r+\delta}\mathcal{G}\left[\exp\left[\sigma\Phi^{-1}(1-\alpha)\right]\right] + \exp\left[\sigma\Phi^{-1}(1-\alpha)\right].$$

Define

$$\theta=\frac{\frac{\delta}{r+\delta}\mathcal{G}\left[\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right]\right]}{\frac{\delta}{r+\delta}\mathcal{G}\left[\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right]\right]+\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right]}.$$

Given β and σ, θ can be calculated. Then, average match surplus follows as

$$\delta\overline{S}=\left(B_m rV_m^0+B_f rV_f^0\right)\theta.$$

Denote $v_g^0=B_g rV_g^0$, $g=m,f$, to simplify notations. This is a function of the person's characteristics (w_g,x_g). We can rewrite the equations for the value functions of singles as

$$v_m^0=B_mU_m^0+\frac{\lambda}{\delta}\beta\int\left(v_m^0+v_f^0\right)\theta n_f\,\mathrm{d}(w_f,x_f), \tag{9}$$

$$v_f^0=B_fU_f^0+\frac{\lambda}{\delta}(1-\beta)\int\left(v_m^0+v_f^0\right)\theta n_m\,\mathrm{d}(w_m,x_m), \tag{10}$$

where

$$B_gU_g^0=\left(\frac{K_g^0}{w_g}\right)^{K_g^0}\left(X_g^0\right)^{1+K_g^0}.$$

This is a linear system that can easily be solved, given β and σ, on a grid of wages and covariates. So, for each period τ, each (w_m,x_m,w_f,x_f), and each region, we deduce the values of $v_m^0+v_f^0$ from the estimated λ,n_m,n_f,α and from the parameters estimated in step 2.

We can also easily deduce match quality Z from the equation

$$\frac{v_m^0+v_f^0}{ZF^1X^1}=\frac{\delta}{r+\delta}\mathcal{G}_\sigma\left[\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right]\right]+\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right]. \tag{11}$$

In practice, however, we aim to fit the marriage probabilities across German regions and periods. So we define the following residuals: for all (τ,w_m,x_m,w_f,x_f,r),

$$u_6=\alpha-\Phi\left[\frac{1}{\sigma}\ln\left(\frac{v_m^0+v_f^0-\delta\overline{S}}{ZF^1X^1}\right)\right].$$

Given β and σ, we can thus estimate Z over the grid for (w_m,x_m,w_f,x_f) by solving

$$\sum_{\tau,\mathcal{R}}\phi(\cdot)\left[\alpha-\Phi\left[\frac{1}{\sigma}\ln\left(v_m^0+v_f^0-\delta\overline{S}\right)-\frac{1}{\sigma}\ln\left(ZF^1X^1\right)\right]\right]=0$$

repeatedly for all values of α,θ. This is like a PROBIT model.

A.3.2 Bargaining coefficient β, and marital bliss dispersion σ

Finally, returning to the sharing rule,

$$\begin{aligned}\rho=\frac{R_m-A_m}{X^1}&=\frac{w_m(e_m^1-a_{1m})}{b_mX^1}\\&=\beta+\frac{v_m^0-\beta(v_m^0+v_f^0)}{ZF^1X^1}\frac{1}{\varepsilon},\end{aligned}$$

where $\varepsilon=z/Z$ is a truncated log-normal distribution with lower bound (from $S(z)>0$) equal to

$$\frac{v_m^0+v_m^0-\delta\overline{S}}{ZF^1X^1}=\exp\left[\sigma\Phi^{-1}\left(1-\alpha\right)\right].$$

Hence,

$$\mathbb{E}\left[ZF^1X^1\frac{\frac{w_m(e_m^1-a_{1m})}{b_mX^1}-\beta}{v_m^0-\beta(v_m^0+v_f^0)}\right]=\mathbb{E}\left(\frac{1}{\varepsilon}|\frac{\ln\varepsilon}{\sigma}>\Phi^{-1}(1-\alpha)\right)=e^{\frac{\sigma^2}{2}}\frac{\Phi\left(\Phi^{-1}(\alpha)-\sigma\right)}{\alpha},$$

and

$$\mathbb{E}\left[\left(ZF^1X^1\frac{\frac{w_m(e_m^1-a_{1m})}{b_mX^1}-\beta}{v_m^0-\beta(v_m^0+v_f^0)}\right)^2\right]=\mathbb{E}\left(\frac{1}{\varepsilon^2}|\frac{\ln\varepsilon}{\sigma}>\Phi^{-1}(1-\alpha)\right)=e^{2\sigma^2}\frac{\Phi\left(\Phi^{-1}(\alpha)-2\sigma\right)}{\alpha}.$$

These two equations identify β and σ from the first two moments of the sharing rule.

Consider the following additional residuals: for each couple (i,j,t) and each parameter value on the grid (τ,x_m,x_f), for (w_m,w_f) closest to (w_{it},w_{jt}),

$$u_{7ijt}=w_{it}(e_{it}^1-a_{1m})-b_mX_{ijt}^1\beta-b_m\frac{v_m^0-\beta(v_m^0+v_f^0)}{ZF^1}e^{\frac{\sigma^2}{2}}\frac{\Phi\left(\Phi^{-1}(\alpha)-\sigma\right)}{\alpha},$$

and

$$u_{8ijt}=\left[w_{it}(e_{it}^1-a_{1m})-b_mX_{ijt}^1\beta\right]^2-\left[b_m\frac{v_m^0-\beta(v_m^0+v_f^0)}{ZF^1}\right]^2e^{2\sigma^2}\frac{\Phi\left(\Phi^{-1}(\alpha)-2\sigma\right)}{\alpha}.$$

We minimize the sum of these residuals over married couples to estimate Z, β and σ:

$$\sum_{\tau,w_m,x_m,w_f,x_f,r}\mu u_6^2+\sum_{\tau,x_m,x_f}\sum_{couple\,i,j,t}\omega_{ijt}^1(\tau,x_m,x_f)\left[u_{7ijt}^2+u_{8ijt}^2\right].$$