CESifo Working Papers

www.cesifo.org/wp

Taxation and Labor Supply of Married Couples across Countries: A Macroeconomic Analysis

Alexander Bick
Nicola Fuchs-Schündeln

CESifo Working Paper No. 6325
Category 6: Fiscal Policy, Macroeconomics and Growth
January 2017

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 6325*

Taxation and Labor Supply of Married Couples across Countries: A Macroeconomic Analysis

Abstract

We document contemporaneous differences in the aggregate labor supply of married couples across 17 European countries and the US. Based on a model of joint household decision making, we quantify the contribution of international differences in non-linear labor income taxes and consumption taxes to the international differences in hours worked in the data. Through the lens of the model, taxes, together with wages and the educational composition, account for a significant part of the small differences in married men's and the large differences in married women's hours worked in the data. Taking the full non-linearities of labor income tax codes, including the tax treatment of married couples, into account is crucial for generating the low cross-country correlation between married men's and women's hours worked in the data, and for explaining the variation of married women's hours worked across European countries.

JEL-Codes: E600, H200, H310, J120, J220.

Keywords: taxation, two-earner households, hours worked.

Alexander Bick
Arizona State University
Department of Economics
USA - Tempe, AZ 85287-9801
alexander.bick@asu.edu

Nicola Fuchs-Schündeln
Goethe University Frankfurt
Department of Macroeconomics &
Development / House of Finance
Germany – 60323 Frankfurt
fuchs@wiwi.uni-frankfurt.de

January 18, 2017
We thank Domenico Ferraro, Paul Klein, Dirk Krueger, José-Víctor Ríos-Rull, Richard Rogerson, Kjetil Storesletten, the editor and four anonymous referees, as well as numerous seminar and conference participants for helpful comments and suggestions. Bettina Brüggemann and Hannah Paule-Paludkiewicz provided truly outstanding research assistance with the data, and Enida Bajgoric, Pavlin Tomov, and Desiree Winges with the tax codes. The authors gratefully acknowledge financial support from the Cluster of Excellence "Formation of Normative Orders" at Goethe University and the European Research Council under Starting Grant No. 262116. Alexander Bick further thanks the Research Department of the Federal Reserve Bank of Minneapolis for the hospitality, and the Fritz Thyssen Foundation for the financial support of that stay. All errors are ours.

1 Introduction

The labor supply of married men and women differs substantially across countries and by gender. While married men in the core age group 25 to 54 work 1970 hours per person annually in the US, their annual hours in Norway amount only to 1620 hours, i.e. 350 fewer hours. The picture is even more heterogeneous for married women. Married women in Portugal work 1270 hours annually, but married women in Italy only 750 hours, resulting in a difference of 520 hours. The standard deviation of married men's hours worked across 17 European countries and the US amounts to 110 hours, but the one for married women to 180 hours.[1] On average, married men in the 17 European sample countries work 11% fewer hours than married men in the US, and married women 18% fewer hours than their US counterparts. Moreover, the cross-country correlation of average hours worked of married men and married women is essentially zero, while countries with low hours worked by married men are also countries with low hours worked by singles, independent of the gender (the cross-country correlation between hours worked of married men and single men [women] is 0.79 [0.64]). Explaining these large differences in the labor supply behavior of married men and married women with linear (average marginal) taxes – a prominent and successful predictor in explaining aggregate hours worked differences in the literature pioneered by Prescott (2004) – seems challenging. Even if women have a higher labor supply elasticity than men, the relative country ordering of hours worked should be similar for both sexes, possibly with larger differences for women, which is not consistent with the data. Yet, international differences in the tax treatment of married couples might have the potential to explain these facts.

In this paper, we build a simple model of joint household decision making that incorporates international differences in consumption and non-linear labor income taxes, including the tax treatment of married couples. The quantitative framework is based on the model of joint labor supply of married couples developed in Kaygusuz (2010) and Guner et al. (2012a, 2012b). As typical for cross-country studies in macroeconomics, we calibrate the model to match the labor supply behavior in a benchmark country, namely the US. We then predict labor supply behavior in 17 European countries, holding preferences fixed but using the country-specific economic environment. The latter comprises non-linear labor income taxes and consumption taxes, as well as gender and educational wage premia, and the educational distribution plus the degree of assortative matching into couples. For the non-linear labor income taxes, we implement the country-specific statutory labor income tax codes in detail, specifically also the tax treatment of married couples. The calibrated model replicates the international differences in married couple's labor supply fairly well: it accounts for 58% of the US-Europe gap in male hours worked, and for 88% of the US-Europe gap in female hours worked; the cross-country correlation between hours worked in the model and the data is 0.43 for both men and women; and the predicted correlation of hours worked of married men and married women in the model is low.

[1] The 17 European countries are Czech Republic, Hungary, and Poland from Eastern Europe; Denmark, Sweden, and Norway from Scandinavia; Austria, Belgium, France, Germany, Ireland, Netherlands, and United Kingdom from Western Europe; and Greece, Italy, Portugal, and Spain from Southern Europe.

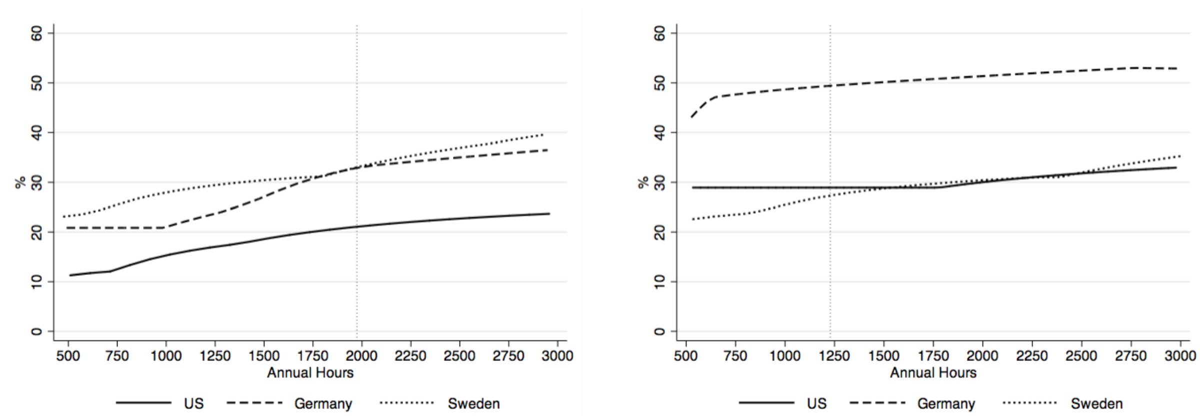
The key to the success of the model in explaining the within-Europe variation in married women's hours worked, and in breaking the correlation between married men's and women's labor supply, is the explicit modeling of non-linearities in the labor income tax codes. These non-linearities comprise both the tax treatment of married couples and the progressivity of the tax system. The tax treatment of married couples across countries ranges from separate to different degrees of joint taxation. In a system of separate taxation, each spouse's marginal tax rate increases only in the own income, i.e. each spouse is taxed like a single individual is taxed. In contrast, in systems of joint taxation one spouse's marginal tax rate increases not only in the *own* income, but also in the *spousal* income. To understand the workings of joint taxation, consider the case of Germany. The incomes of husband and wife are summed up and divided by two, and the tax burden on this hypothetical income is calculated using the tax function of singles. The total household tax burden is then twice this amount. Due to the progressivity of the German tax code, this implies that the secondary income earner (whose contribution to the household income is less than half, i.e. typically the wife) faces a higher marginal tax rate than under separate taxation, while the opposite is true for the primary earner (typically the husband).[2] Joint taxation therefore leads to a higher model-predicted hours gap between spouses than separate taxation. Hence, while average labor income tax rates drive hours of both spouses in the same direction, feeding the international differences in the tax treatment of married couples into our quantitative model reduces the cross-country correlation of hours worked of married men and women.

The key importance of the tax treatment of married couples can best be illustrated with a concrete example. Take the cases of the US, a country with **low** average tax rates and *joint* taxation, Germany, a country with **high** average tax rates and *joint* taxation, and Sweden, a country with **high** average tax rates but with *separate* taxation. Figure 1 shows in the left panel the average labor income tax rate of a household in which the wife does not work, plotted against the hours worked by the husband. This average labor income tax rate is slightly larger in Sweden than in Germany, and substantially lower in the US. Evaluated at the mean hours of US married men (indicated by a vertical line at 1970 hours) it amounts to 21.4% in the US, compared to 33.4% in Germany, and 33.5% in Sweden. The right panel shows the average marginal tax rate that a wife faces if she goes from not working to working a specific number of hours, varied on the horizontal axis, keeping the husband's hours fixed at 1970 hours. This average marginal tax rate is substantially higher in Germany, but similar in Sweden and the US. At the average hours of US married women, indicated with a vertical line at 1235 hours, it amounts to 29.1% and 28.0% in the US and Sweden, respectively, but is drastically higher in Germany at 49.6%. Comparing the left and the right tax schedule,

[2]Under the German system of joint taxation, the overall tax burden of the household is always at most as large as under separate taxation, i.e. there generally exists a marriage bonus. While the US system works differently, it also features joint taxation, i.e. the marginal tax rate of one spouse is increasing in the income of the other spouse. In the US, the incomes of both spouses are summed up, and the income thresholds defining each tax bracket are multiples of the thresholds for singles. In 2008, e.g., these thresholds are multiplied by two only for the two lower tax brackets, and by 1.67, 1.22, and 1 for the three subsequent thresholds. Since higher tax brackets feature higher marginal income tax rates, married households in the US may face a marriage bonus or a marriage penalty. For 2008, the lowest level of household income from which onwards a marriage penalty can occur is $131,450 (which is slightly above the 90th percentile of household income in the 2007 Survey of Consumer Finances, see Díaz-Giménez et al., 2011). This happens however only when both spouses' earnings are identical. For households with only one earner there is always a marriage bonus. For households with a household income above $131,450 and unequal earnings across spouses, the level from which onwards a marriage penalty occurs is increasing in the ratio of earnings of the primary to the secondary earner.

Figure 1: Labor Income Tax Codes in the US, Germany, and Sweden

(a) Avg. Income Tax Rate if Only the Husband Works (b) Avg. Marginal Income Tax Rate on Wives' Earnings



Note: The tax rates are calculated assuming that husband and wife earn the country-specific mean wages of married men and married women. Note that in each country the female wage is lower than the male wage. Hours worked of wives are fixed at 0 in the left panel, and hours worked of husbands are fixed at 1970 in the right panel. The average marginal tax rate in the right panel is calculated as the additional tax payments of the household divided by the wife's earnings if the wife goes from working 0 hours to working the respective hours on the x-axis. Vertical dashed lines represent mean hours worked per married man (left panel) and married woman (right panel) in the US.

one can see that tax rates are slightly lower for secondary income earners in Sweden, a country with separate taxation, but substantially higher for secondary income earners than for primary income earners in the US and Germany, countries with joint taxation of married couples. In the data, married men in Germany and Sweden work roughly the same hours, around 15% fewer hours than in the US. Married women in Sweden, in turn, work only slightly fewer hours than US married women (4%), while German married women work 34% fewer hours. The model generates this variation because it features both differences in the average tax rate and in the tax structure, with the latter reflecting the progressivity of the tax code as well as the tax treatment of couples.

This logic carries over to the full sample of countries. The non-linearities in the labor income tax codes generate variation in married women's labor supply within Europe that is in line with the data, and further predict on average lower hours in Europe than in the US for married men. However, since the US features a system of joint taxation, while many European countries tend more towards separate taxation systems, the tax treatment also reduces work incentives for married women in the US. Hence, imposing the European labor income tax code on US married women does not always predict lower hours worked, despite higher average labor income tax rates in Europe. The substantially higher consumption taxes in Europe by contrast unambiguously predict lower hours in Europe. Thus, to explain the on average higher hours worked of US than European married women, differences in consumption taxes are crucial. For married men, the effects of the tax structure (i.e. the combination of the tax treatment of couples and progressivity of the tax code) go in the same direction as the effects of the average tax rate, both predicting lower hours in Europe than in the US. Furthermore, high shares of low educated individuals, who exhibit the lowest hours worked in each country, and high educational wage premia in Southern Europe help explain low female hours worked there.

This paper is organized as follows: The next section relates our paper to the literature. Section 3 shortly discusses our data and sample selection criteria, and then shows some facts on the labor supply of married couples across countries. Section 4 introduces the model, as well as its parametrization and calibration. Section 5 presents the results, and investigates the relative role of the various model inputs, specifically of the non-linear labor income tax schedule. Section 6 then turns the attention to the two margins of labor supply for married women. The last section concludes.

2 Related Literature

A series of papers (Prescott, 2004; Rogerson, 2006; Rogerson, 2008; Ohanian et al., 2008; Rogerson, 2009) have shown that differences in consumption and average income tax rates, combined into one linear tax rate on income, can largely explain differences in the time series of *aggregate* hours worked across European countries and the US after World War II, with the exception of Scandinavia. These papers abstract from modeling gender. We distinguish explicitly between consumption and labor income taxes, and show that it is crucial to take the full non-linearity of labor income tax schedules into account to explain hours worked by married men and women. Two features that we abstract from in our analysis are capital income taxes and retirement incentives through social security programs. McDaniel (2011) shows in a dynamic model that labor income and consumption taxes are much more important than capital income taxes and productivity growth in explaining the different developments of total hours over time across countries. Moreover, Wallenius (2013) finds that differences in retirement systems have almost no effect on labor supply behavior before retirement, i.e. in the age group we focus on, although they play a crucial role in explaining the large international differences in the timing of retirement (see also Erosa et al., 2012; and Alonso-Ortiz, 2014).[3] Guvenen et al. (2014) investigate the role of cross-country differences of non-linear labor income tax codes for singles in driving post- and pre-tax inequality. They show that differences in progressivity can explain a sizable part of the differences in the US-Europe hours gap for men, but do not investigate female hours worked.

High hours worked in Scandinavia despite high consumption and labor income taxes there have been raised as a puzzle in the literature. Ragan (2013), and in a similar fashion Ngai and Pissarides (2011), suggest government subsidies in sectors that serve as substitutes to home production, especially child care, as the main explanation for this apparent puzzle, a point that had already been raised theoretically by Rogerson (2007). They therefore explicitly model home production in addition to market work, and analyze the effects of international differences in government subsidies.[4] By contrast, we can replicate Scandinavian

[3] By starting at the age 25, we also abstract from international differences in educational systems and youth unemployment rates.

[4] Duernecker and Herrendorf (2016) also show the importance of home production in explaining international aggregate hours worked differences, but focus on differential productivity improvements in the home production sector, rather than government subsidies. Bridgman et al. (2016) report hours of home production for a large set of countries.

hours well by taking the non-linearity of the labor income tax code into account.[5] Our approach thus offers a complementary explanation to Ragan (2013) and Ngai and Pissarides (2011) for high hours worked by Scandinavian married women.

The paper most closely related to ours is Chakraborty et al. (2015). They build a comprehensive life-cycle model with idiosyncratic income risk to investigate the cross-country variation in hours worked of married and single men and women. Besides taxes, they concentrate on exogenous marriage and divorce probabilities as driving forces of labor supply differences. Chakraborty et al. (2015) estimate polynomial tax functions for single and married individuals based on data from the OECD taxing wages modules. The tax functions of married individuals are a weighted average of two tax polynomials assuming fully separate or fully joint taxation. They find that marriage and divorce rates generate variation in female hours that is in line with the data, but conclude that taxes cannot explain any of the variation in female hours worked across countries. In contrast to this study, their paper focuses on single and married households aged 20 to 64, which makes it difficult to compare the results directly.

Besides the literature on international labor supply differences, our paper connects to the large literature documenting the increase in labor supply of married women in the US over the last decades. Based on the results in Albanesi and Olivetti (2009), Jones et al. (2014), and Olivetti (2006), who trace part of this increase back to changes in the gender wage gap and the return to experience for women, we account for differences in gender and educational wage gaps between countries.[6] Attanasio et al. (2008) point to the important role of the decline in child care costs. In Section 6, we do not find supporting evidence that international differences in child care costs are related to labor supply differences of married women, with the caveat that it is difficult to obtain appropriate and comparable measures of individual child care costs across countries. Further important driving factors of the increase in the labor supply of US married women analyzed in the literature are technological improvements in the household sector (Greenwood and Seshadri, 2002; Greenwood et al., 2005), improvements in maternal health and the introduction of infant formula (Albanesi and Olivetti, 2015), and structural change (Buera et al., 2014; Ngai and Petrongolo, 2014; and Rendall, 2015).

Within this literature, Rendall (2015) analyzes the effect of taxation on the increase in female labor supply in the US over time, and links both phenomena to structural change by focusing on the size of the service sector. Crossley and Jeon (2007) study in a difference-in-differences approach a Canadian tax reform in 1988 which reduced the "jointness" of the labor income tax system, while Eissa (1995, 1996) analyzes in a similar approach the effects of significant decreases in the top marginal income tax rate in the US in the 1980s. These three studies conclude that the relevant tax reforms increased the labor supply of wives of

[5] Even for the Scandinavian country for which we obtain the worst fit, namely Denmark, our prediction for hours worked by married couples is off by only 12%, whereas the corresponding number in Ragan (2013) is 51% in her benchmark calibration, and 41% in the specification with government subsidies. Note, however, that our data and predictions refer to married men and women in the core age group, while Ragan's sample comprises all men and women aged 15 to 64. We cannot directly compare our results to Ngai and Pissarides (2011), as they do not predict total hours, but only relative shares in different sectors.

[6] Knowles (2013) shows further implications of changes in the gender wage gap through changes in the bargaining power within the household, a feature we omit in our unitary household model.

high-earning husbands significantly. Kaygusuz (2010) analyzes the effects of the same US tax reforms on the labor supply of married women with a quantitative model. Bar and Leukhina (2009) carry out a similar analysis including a wider range of tax reforms. Guner et al. (2012a, 2012b) evaluate counterfactual policy reforms in an elaborate quantitative life-cycle model. They find that going from joint to individual taxation would increase the labor supply of married women in the US substantially. In Bick and Fuchs-Schündeln (2017), we analyze a similar counterfactual tax reform in European countries in addition to the US.

3 Hours Worked of Married Couples

3.1 Data and Sample Selection

We use three different micro data sets, namely the European Labor Force Survey (ELFS), the Current Population Survey (CPS) from the US, and the German Microcensus, to construct internationally comparable hours worked for married men and women aged 25 to 54. The main challenge lies in the variation in reference weeks in the surveys across countries and within countries over time, which we overcome by adjusting for vacation weeks from external data sources. A detailed description of the data work can be found in Bick et al. (2016).

We include only married individuals into the sample. There are a few countries which differentiate between marriage and a civil union. In this case, the ELFS makes it explicitly clear that every respondent who is treated for tax purposes as "married" should indicate married as the civil status. Moreover, we focus on couples in which both husband and wife are aged 25 to 54. Since we are mainly interested in the role of taxation in explaining international differences in hours worked of married couples, we focus on the core age group and avoid discussing international differences in the educational systems, degrees of youth unemployment, and early retirement programs.

We concentrate on the sample period 2001 to 2008. We use more than one year of data and do not further analyze the time series in order to avoid that cross-country differences might be driven by uncorrelated business cycles. The start year is determined by the availability of the OECD Taxing Wages modules, which are a key input into our quantitative analysis.

3.2 Hours Worked of Married Couples: Some Facts

Table 1 shows some statistics on hours worked per person by gender and marital status over the 18 sample countries and averaging over the years 2001 to 2008. Hours worked per person are the product of the employment rate and hours worked per employed; we discuss the breakdown into both margins for married women in Section 6. On average, married men aged 25 to 54 work around 730 hours per person more than married women in the same age group. Single women (including divorcées) work 160 hours more than married women, and single men 280 hours less than married men. While married women are thus clearly the group with the lowest hours worked, they exhibit the largest cross-country standard deviation in mean hours worked per person: the standard deviation of hours worked of married women is more than 60% higher than

Table 1: Cross-Country Statistics on Annual Hours Worked by Gender and Marital Status (Ages 25-54)

	Men		**Women**	
	Married	**Single**	**Married**	**Single**
Mean	1761.3	1484.6	1027.9	1189.4
Standard Deviation	104.4	110.6	179.8	100.7
Coefficient of Variation	0.059	0.075	0.175	0.085
Var(log hours)	0.003	0.006	0.033	0.007
Correlation with Married Men	1.00	0.79	0.06	0.64

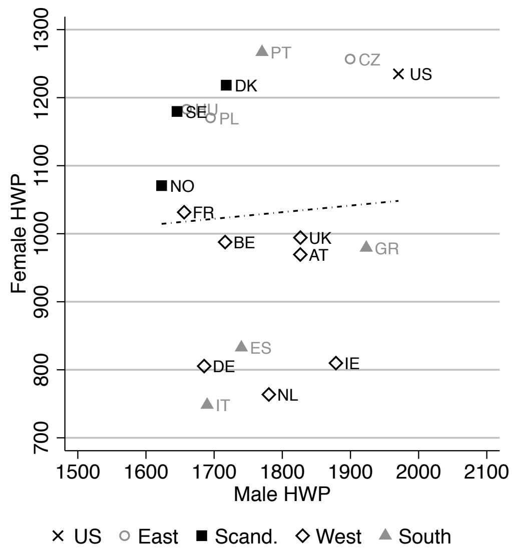
the ones of the other three demographic groups, while the coefficient of variation is even more than twice as large, and the variance of log hours is up to an order of magnitude bigger. Married women contribute on average 20% to total hours worked, but account for 41% of the variance of total hours worked. Moreover, the cross-country correlation of hours worked of married men with hours worked of single men or single women is 0.79 and 0.64, respectively, while the correlation with hours worked of married women amounts only to 0.06. Thus, there is clearly something special about married women, and investigating the sources of the different behavior of married men and married women is of great interest if one wants to understand international differences in hours worked among core aged individuals.

Since from now on we focus on married couples, the issue of selection into marriage arises. While we do not model this selection, we report in Figure B.1 in the Online Appendix the fraction of women in the core age group who are married. It amounts on average to 64%, with a standard deviation of 7.5%. Any potential selection bias could go in either direction, but we find it reassuring that the cross-country correlation of the fraction of married women in our core age group with married women's hours worked is virtually zero. Similarly, the tax treatment of couples, i.e. whether a country employs a system of joint or separate taxation, is not correlated with the marriage rate.[7] Last, Chade and Ventura (2002) and Chade and Ventura (2005) show in a quantitative equilibrium model of the marriage market for the US that the marriage rate would barely change if the US replaced the current system of joint taxation with one of separate taxation.

Figure 2 shows average hours worked of married men (on the x-axis) and married women (on the y-axis) aged 25 to 54 over the period 2001 to 2008 for all eighteen countries in our sample. The European countries are grouped into four regions, namely Eastern Europe (circles: Czech Republic, Hungary, Poland), Scandinavia (squares: Denmark, Norway, Sweden), Western Europe (diamonds: Austria, Belgium, Germany, France, Ireland, Netherlands, United Kingdom), and Southern Europe (triangles: Spain, Greece, Italy, Portugal), while the US is marked by an x. The figure illustrates the almost zero correlation between hours worked of married men and married women reported in Table 1. The absence of any correlation between

[7] We compute cross-country correlations between the marriage rate and two (imperfect) measures of the degree of jointness of taxation that we discuss later in this paper: the correlation between the marriage rate and the difference in the tax rates shown in panel (b) of Figure 3 is 0.07, and the correlation between the marriage rate and the size of the tax structure effect shown in Table 4 is 0.09.

Figure 2: Average Annual Hours Worked Per Person of Married Women and Men (Ages 25-54)



hours worked of married men and women is very surprising, since one would expect institutional features that can potentially generate differences in hours worked across countries, e.g. taxes, to work similarly on both men and women, at least qualitatively. Figure 2 also reflects the finding that the variation across countries is higher for the hours of married women than for the hours of married men. Hours worked of married men are highest in the US, followed closely by Greece, the Czech Republic, and Ireland. At the lower end of the sample are Norway, Sweden, France, and Hungary. Norwegian married men work 350 hours less than, or only 82% of, US married men.

For married women, there exists a regional pattern of hours worked per person, which are high in the US, Eastern Europe, and Scandinavia, and substantially lower in Western and Southern Europe. Portugal is a notable exception among the Southern European countries and actually features the highest hours worked of married women among all countries in our sample. Western Europe is somewhat divided with Germany, Ireland, and the Netherlands having relatively low hours worked, comparable to Italy and Spain, whereas France, Belgium, the UK, and Austria have higher hours worked, but still below the level of Scandinavian and Eastern European countries. The lowest hours worked arise in Italy with 750 hours, 485 hours less than, or only 61% of, US married women.

4 Model

4.1 A Model of Joint Household Labor Supply

We build a static model of married couples' hours decisions to investigate in how far cross-country differences in consumption and labor income taxes contribute to the cross-country differences in male and female labor supply presented in Figure 2. The model framework is based on Kaygusuz (2010) and features a maximization problem of a two person household which optimally determines male and female labor supply.

There is a continuum of married households of mass one. Each household member exhibits one of three possible education levels, denoted by $x \in \{low, medium, high\}$ for women and by $z \in \{low, medium, high\}$ for men, which determine the offered wages $w_f(x)$ and $w_m(z)$. We denote the fraction of households of type x,z by $\mu(x,z)$ with

$$\sum_x \sum_z \mu(x,z) = 1. \tag{1}$$

Households draw a utility cost of joint work q from a distribution $\zeta(q|z)$ which depends on the husband's education level. This cost is only incurred if the wife participates in the labor market, and thus introduces an explicit extensive margin choice for women. We abstract from modeling fixed costs of work for men. As a consequence, they always optimally choose to provide positive hours. We follow this approach because male participation rates in the data in our sample countries are above 90% and display only little variation across countries. The draw q can be interpreted as a utility loss due to joint work of two household members originating from, for example, inconvenience of scheduling joint work, home production and leisure activities, or spending less family time with children, see Kaygusuz (2010). It captures residual heterogeneity across households - conditional on the husband's education level - regarding the participation choice. For each household x,z, there exists a threshold level $\bar{q}(x,z)$ from which onwards the utility cost of working is so high that the woman chooses not to work, i.e. $h_f = 0$.

Households face two types of taxes, namely a linear consumption tax at rate τ_c and a non-linear labor income tax τ_l, which depends on the gross incomes of husband and wife, as well as the number of children in the household k through tax credits and/or child benefits. The maximization problem of a type $\{x,z\}$ household is given by

$$\max_{h_m,h_f} \left\{ \ln c - \alpha h_m^{1+\frac{1}{\phi}} - \alpha h_f^{1+\frac{1}{\phi}} - q\mathbf{I}_{h_f>0} \right\} \tag{2}$$

$$\text{s.t. } c = \frac{y_{hh} - \tau_l}{(1+\tau_c)} + T \tag{3}$$

$$y_{hh} = w_m(z)h_m + w_f(x)h_f \tag{4}$$

$$\tau_l = \tau_l(w_m(z)h_m, w_f(x)h_f, k) \tag{5}$$

where $\mathbf{I}_{h_f>0}$ takes the value one if the wife is working and zero otherwise, c represents household consumption, and y_{hh} represents gross household income.

T represents a lump-sum transfer from the government, which redistributes a share $\lambda \in [0,1]$ of all government revenues:

$$T = \frac{\lambda}{1+\tau_c}\sum_x\sum_z \mu(x,z)\left[\int_{-\infty}^{\infty}\tau_l(w_m(z)h_m^*(q),w_f(x)h_f^*(q),k)\zeta(q|z)dq \right. \\ \left. +\tau_c\int_{-\infty}^{\infty}\left(w_m(z)h_m^*(q)+w_f(x)h_f^*(q)\right)\zeta(q|z)dq\right], \tag{6}$$

where * denotes the optimal hours choice given the draw of q.[8] When deciding on how much to work, households do not internalize that their choices impact the equilibrium transfer, but rather take the transfer as given.

As usual in the literature explaining aggregate hours worked differences between Europe and the US, consumption and labor supply are assumed to be separable, and utility from consumption is logarithmic. Therefore, cross-country differences in mean wages are irrelevant, and only cross-country differences in the gender and education premia matter for labor supply decisions, see Online Appendix Section C.1. α captures the relative weight on the disutility of work, and ϕ determines the curvature of this disutility. Both parameters are the same for men and for women.

One might be worried that part of the international differences in married women's hours worked across countries are driven by differential effects of children on mothers' hours, different life-cycle or cohort effects across countries, or differences in the supply side rather than the demand side, all factors that our model does not incorporate. In Online Appendix Sections B.4 and B.5, we present extensive evidence on hours worked by presence of children, as well as hours worked for different age groups. While one can identify country-specific idiosyncracies, these are not large enough to drive the international differences in aggregate hours worked of married couples. Moreover, in Section C.4 in the Online Appendix, we show results from our model focusing exclusively on women without children, and the overall model fit is qualitatively and quantitatively very similar. Last, in Online Appendix Section B.6, we show that the hours worked differences are robust to excluding unemployed individuals from the sample.

4.2 Some Theoretical Results

In order to highlight the key implications of differences in the tax treatment of married couples, we provide theoretical insights from a closed-form analysis. In order to do so, we abstract from heterogeneity in education, and assume the simplified labor income tax schedule introduced by Benabou (2002). After-tax income is given by $y^{net} = (1-\delta)(wh)^{1-\zeta}$, where ζ measures the degree of tax progressivity, and δ is a measure of the average tax rate. In the absence of any progressivity ($\zeta = 0$) the average tax rate is exactly given by δ. In the case of full progressivity ($\zeta = 1$), in turn, after tax income is independent of the gross income.

[8] Equation (6) is derived as follows. For ease of exposition, assume there would be just one household consisting of a single member. Total government revenues R are the sum of the revenues from the labor income tax and from the consumption tax, i.e. $R = \tau_l + \tau_c(c-T)$. T is subtracted in this calculation since the transfer is not subject to the consumption tax. Replacing c from the budget constraint ($c = \frac{1}{1+\tau_c}(y_{hh}-\tau_l)+T$), yields $R = \frac{1}{1+\tau_c}[\tau_l + \tau_c y_{hh}]$. The transfer T is then the fraction λ of government revenues R.

We compare labor supply under a system of separate taxation (s) to labor supply under a system of joint taxation (j) built on the "German model". Simplifying, a couple's tax burden in Germany is calculated as twice the tax burden of the *average* household income, such that net household income and the marginal return to working an extra hour are given by

$$y_{hh}^{net,j} = 2(1-\delta)\left(\frac{w_m h_m + w_f h_f}{2}\right)^{1-\zeta} \Rightarrow \frac{\partial y_{hh}^{net,j}}{\partial h_g} = (1-\delta)(1-\zeta)\left(\frac{w_m h_m + w_f h_f}{2}\right)^{-\zeta} w_g \ \forall\, g = m,f. \quad (7)$$

Under separate taxation, household net income and the marginal return to working an extra hour are

$$y_{hh}^{net,s} = (1-\delta)(w_m h_m)^{1-\zeta} + (1-\delta)(w_f h_f)^{1-\zeta} \Rightarrow \frac{\partial y_{hh}^{net,s}}{\partial h_g} = (1-\delta)(1-\zeta)(w_g h_g)^{-\zeta} w_g \ \forall\, g = m,f. \quad (8)$$

Note that the marginal return to working an extra hour depends – in addition to the own wage – on the *household* income in the joint taxation regime, but on the *individual* income in the separate taxation regime. Hence, the marginal tax rate (defined as $1 - \frac{\partial y_{hh}^{net,j}}{\partial h_g}/w_g$) is always the *same* for both spouses under joint taxation, but *lower (higher)* for the lower (higher) earning spouse under separate taxation for $\zeta \in (0,1)$. Put differently, under joint taxation the marginal tax rate of the primary earner (i.e. in most cases the husband) is *lower* than the one he would face under separate taxation, and the marginal tax rate of the secondary earner (i.e. in most cases the wife) is *higher* than the one she would face under separate taxation.[9] Progressivity is a necessary condition for this to be true, and the effect is larger the higher the progressivity of the tax system. As a result, the optimal hours difference between both spouses is always larger under joint taxation than under separate taxation, and this difference between the two tax treatment regimes is increasing in tax progressivity. This can be directly seen from combining the first order conditions (assuming an interior solution) for each spouse and under each system of taxation:

$$\frac{h_m^{*j}}{h_f^{*j}} = \left(\frac{h_m^{*s}}{h_f^{*s}}\right)^{\frac{1+\zeta\phi}{1-\zeta}}, \quad (9)$$

because $\frac{1+\zeta\phi}{1-\zeta}$ is larger than one and increasing in ζ.[10]

4.3 Model Inputs

As inputs into the model, we use country-specific information on consumption tax rates τ_c and non-linear labor income taxes τ_l, and additionally the educational composition and matching into couples $\mu(x,z)$, male

[9] This results in a marriage bonus under joint taxation. Formally, the progressivity of the tax code implies that the tax burden (net income) is a convex (concave) function of income if $\zeta \in (0,1)$. Hence, by Jensen's inequality, the sum of the tax burden for two different incomes is larger than twice the tax burden for the average of those two incomes.

[10] Online Appendix Section C.1 derives Equation (9). Relying on the same model, we also show there that the average labor income tax rate and the consumption tax rate only matter for the optimal hours choice in the presence of positive transfers, while tax progressivity matters even in the absence of positive transfers. Qualitatively, the consumption tax rate and the average labor income tax rate have the same effect on hours worked.

hourly wages by education $w_m(z)$, and female hourly wages by education $w_f(x)$. Last, we need to calibrate the two preference parameters in the utility function, α and ϕ, and the parameters governing the fixed cost distributions. Table B.5 in Online Appendix Section B.7 lists all data sources for all countries.

4.3.1 Consumption Taxes

Consumption tax rates for our sample countries are provided by McDaniel (2012), who calculates consumption tax rates from NIPA data.[11] The advantage of these tax rates over simple value added tax rates is that they also capture excise taxes, exemptions from the value added tax, etc. They are shown in panel (a) of Figure 3. Differences in consumption tax rates between Europe and the US are large, amounting to 14 percentage points on average. The Czech Republic has the lowest consumption tax rate in Europe at only 14.8%, and Scandinavia the highest rates, where only Norway's is below 30%.

4.3.2 Non-Linear Labor Income Taxes

The labor income tax codes come from the OECD Taxing Wages modules. These are very similar to the NBER TaxSim module for the US, but cover all OECD countries from 2001 onwards. The OECD Taxing Wages module implements the statutory labor income tax code, including employees' social security contributions and cash benefits, by marital status and number of children in the household. It calculates a household's net income for any combination of husband's and wife's earnings. Standard deductions (i.e. basic allowances, allowances for children, deduction of social security contributions) are included in these calculations, whereas individual non-standard deductions (e.g. mortgage payment deductions, deductions of child care expenses, deductions for expenses on household helpers) are not.[12] Cash benefits include those obtained for children. We compute net household earnings for a grid of wives' annual earnings with 201 grid points, ranging from 0 earnings to three times the average annual earnings in the country, and for an earnings grid with 101 grid points for husbands, ranging from 0 earnings to four times the average annual earnings in the country.[13] We then linearly interpolate in two dimensions to assign a net annual household income to each possible annual hours choice of husband and wife. One additional input into the tax codes are the number of children. From the micro data, we calculate the percentage of married couples with 0, 1, 2, 3, or 4+ children conditional on the educational match, and then take the weighted average over these tax burdens for any pair of hours choices. Figure 1 gives an illustrative impression of the resulting tax schedules for the US, Germany, and Sweden, holding the earnings of the other spouse fixed at one specific level. When

[11] McDaniel (2012) does not provide consumption tax rates for Hungary, which we take from the OECD.

[12] State and local income taxes are included, assuming that the individual lives in a "typical area" in terms of income taxation. For the US, Michigan and Detroit are used.

[13] For women, we thus put in as many steps as the OECD Taxing Wages module allows. To give a specific example, for the US for the year 2005 the difference between two annual earning levels for men amounts to 2297 US-Dollars and for women to 689 US-Dollars. Note that even though in some countries the top tax bracket applies to incomes larger than four times the average annual earnings, the wage that we assign to highly educated men and women never exceeds this threshold even for high hours choices. There is some discretion in setting the distance between these grid points. Visual inspections of the differences in tax burdens between grid points let us conclude that our grid choice is sufficiently precise.

Figure 3: Model Inputs

Note: τ_c are consumption tax rates as calculated by McDaniel (2012). $\tau_l(0)$ is the country-specific average tax rate evaluated at the average US annual hours worked by married men, assuming the husband is earning the country-specific mean married male wage and the wife does not work. $\tau_l'(h_f^{US})$ is the average marginal tax rate if the woman goes from not working to working the mean hours of US married women and earns the country-specific mean married female wage, i.e. $[\tau_l(w_m h_m^{US}, w_f h_f^{US}) - \tau_l(w_m h_m^{US}, 0)]/[w_f h_f^{US}]$. The educational shares omit the group of medium educated women. The education premia are the wages of high educated women divided by the wages of low educated women.

we use the tax schedules as model inputs, we vary of course the incomes of both spouses at the same time.

While it is impossible to summarize the complex non-linear labor income tax systems in a few numbers, panel (b) of Figure 3 presents two illustrative measures that reflect key aspects of the labor income tax schedule: one is the country-specific average tax rate evaluated at US mean hours worked of married men, assuming that the husband is earning the country-specific mean married male wage and that the wife does not work, and thus gives one of many possible measures of an average tax rate. The other one is the average (marginal) tax rate paid by the household on the *additional* income earned if the woman goes from not working to working the mean hours of US married women and earns the country-specific mean married female wage, thereby capturing one possible measure of the interaction between progressivity and jointness of taxation.[14] We use for both men and women the corresponding US hours to show the average/marginal tax rates faced at the mean country-specific wages for the same hours choices across all countries. Let us stress again that these two tax rates shown here are only indicative values, and that we exploit the full non-linearities of the labor income tax codes in our quantitative analysis.

The US average tax rate of a single-earner married couple as depicted in the black bars of panel (b) of Figure 3 amounts to 21%, whereas the corresponding Danish married couple faces an average tax rate of 40%, and the Spanish couple a tax rate of only 17%. The measure of the average (marginal) tax rate of the secondary income earner, by contrast, amounts to 29% in the US, peaking at close to 50% in Germany, a country with high progressivity and joint taxation of married couples. Compared to the large Europe-US differences in consumption tax rates, the Europe-US differences in either of the two labor income tax rates are smaller for the vast majority of European countries, and sometimes even negative. Female US hours are lower than male US hours, and mean wages are always lower for women than for men. Therefore, in countries with a progressive tax code and strictly separate taxation of married couples, the average marginal tax rate of the wife if she starts working should be smaller than the average tax rate of the single-earner household. Thus, one can take the difference between the average marginal tax rate of the married woman and the average tax rate of the single-earner household as a suggestive measure of the effects of jointness of taxation interacted with the degree of progressivity within a country. According to this measure, the US, France, Belgium, Denmark, and Germany feature strong elements of joint taxation combined with high levels of progressivity. For Greece, Hungary, the UK, Austria, and Sweden, by contrast, the difference is strongly negative, indicating rather separate systems of taxation.

4.3.3 Educational Composition and Matching into Couples

We take the percentage of husbands and wives per education group, as well as their matching into couples, directly from the data, relying on the three education groups low, medium, and high.[15] The percentages of women with low and high education, omitting the group of medium educated women, are shown in panel (c) of Figure 3. There are substantial differences in the educational composition: the US features the lowest share of low educated women, amounting to 8%, and the highest share of high educated ones, amounting to 46%. The share of low educated women is largest in the Southern European countries, ranging from 34% in Greece to 73% in Portugal. It is also relatively high in the Western European countries. The fraction of

[14] We call the average tax rate if the woman does not work $\tau_l(0)$, and define the average marginal tax rate if the woman starts working as $\tau_l'(h_f^{US}) = [\tau_l(w_m h_m^{US}, w_f h_f^{US}) - \tau_l(w_m h_m^{US}, 0)]/[w_f h_f^{US}]$. All tax rates in panel b of Figure 3 are reported for couples without children. Children decrease $\tau_l(0)$ via tax credits etc. (the effect is rather similar across countries), but hardly affect $\tau_l'(h_f^{US})$.

[15] Low education is defined as primary and lower secondary education (ISCED categories 0 to 2), medium education as upper secondary and non-tertiary post-secondary education (ISCED categories 3 and 4), and high education as any tertiary education (ISCED categories 5 and above). In the US, low education is defined by having completed at most 11th grade of high school; medium education by having completed the 12th grade of high school, having a high school diploma, or attended some college; and high education by having at least a college degree. See Online Appendix B.2 for details how we proceed for Scandinavia, where the data do not contain household identifiers.

highly educated married women is relatively high in Scandinavia and Western Europe with the exception of Austria, while the Eastern European countries and Austria and Germany feature the highest share of medium educated women. Online Appendix Table B.11 reports the same shares for married men, which are very similar. Moreover, it shows a simple correlation coefficient of the matching into couples between the three education groups. The degree of assortative matching is relatively homogeneous across countries, with assortative matching being naturally more prevalent in countries in which a large share of the population has the same educational level.

4.3.4 Gender and Education Wage Gaps

To calculate hourly wages, we divide earnings by hours. Unfortunately, the ELFS does not provide earnings data, and the German Microcensus only net earnings. Therefore, we use the EU Statistics of Income and Living Conditions (EU-SILC) to compute mean wages for each gender-education cell. This European household data set captures income and usual hours, but features a sample size an order of magnitude smaller than the ELFS. For married women, the issue of self-selection into employment arises. If high ability women of each education group are more likely to join the labor force, then observed mean wages overestimate the mean of offered wages, see e.g. Olivetti and Petrongolo (2008). We therefore apply a simple two-stage Heckman procedure to impute wages of non-working women. The exclusion restrictions are that the income of the husband as well as the presence of children do not directly influence the wage of a woman, see e.g. Mulligan and Rubinstein (2008).

Panel (d) of Figure 3 shows the corresponding mean gender wage gap in each country, as well as the education premium (defined as the ratio of wages for high and low educated people) for women.[16] The average gender wage gap varies from 0.66 in Spain to 0.85 in Denmark and Belgium. Its variation across countries is thus small compared to the variation in education premia, which range from 1.1 in Norway to 3.6 in Poland.[17] The education premia are generally low in the Scandinavian countries, and high in Southern Europe.[18]

4.4 Redistribution of Government Revenues

The government redistributes a fraction $\lambda \in [0,1]$ of all government revenues back to the households in a lump-sum fashion. In the benchmark calibration, we follow Rogerson (2008), Ohanian et al. (2008), and Ragan (2013), and assume full redistribution of government revenues and thus set $\lambda = 1$. In Online Appendix Section C.4, we show results from two alternative specifications with either no redistribution of

[16]Corresponding education premia for men are shown in Table B.11 in the Online Appendix and are largely similar.

[17]In Poland, both the shares of high and low educated women are relatively small, indicating potentially larger selection than in other countries.

[18]We conduct two robustness checks with respect to wages, which are discussed in Online Appendix Section C.4. In the first robustness check, no Heckman-correction is applied on female wages and only observed wages are used. In the second robustness check, we allow for wage heterogeneity within each education group. Our country-specific estimate of heterogeneity is the standard deviation of the residuals from a regression of log male wages on year and education dummies.

government revenues (i.e setting $\lambda = 0$), or from setting λ equal to 1 minus twice the share of military expenditures from all government expenditures, similar to the specification used by Prescott (2004).

4.5 Calibration of Preference Parameters

As Kaygusuz (2010), we set the labor supply elasticity $\phi = 0.5$, which is consistent with the estimates surveyed in Blundell and MaCurdy (1999), Domeij and Flodén (2006), and Keane (2011).[19] The weight on the disutility of work (α) is calibrated to match average hours worked per person by men (recall that we do not model an explicit intensive margin for them) and hours worked per employed woman, using the US as the benchmark country.

Again following Kaygusuz (2010) and Guner et al. (2012a, 2012b), the utility cost parameter is distributed according to a flexible gamma distribution, with the shape parameter k_z and scale parameter θ_z being conditional on the husband's type:

$$q \sim \zeta(q|z) \equiv q^{k_z - 1} \frac{\exp(-q/\theta_z)}{\Gamma(k_z)\theta_z^{k_z}}, \tag{10}$$

where $\Gamma(\cdot)$ is the Gamma function. For each husband's education level z, we select the parameters k_z and θ_z to match as closely as possible the female labor force participation rates by their wives' own education levels $x \in \{low, medium, high\}$. For given preference parameters α and ϕ, and conditional on being married to a type z husband, the three different education levels x and implied wages generate three different threshold levels $\bar{q}(x,z)$ at which a woman of type x is indifferent between working and not working. Assume for simplicity that all type z husbands work the same amount of hours. Women with more education, i.e. a higher wage, will have a higher threshold q, and therefore a higher labor force participation rate, for any given distribution of q. This pattern is also prevalent in the data, i.e. conditional on the husband's education, the female labor force participation rate is increasing in the woman's own education. The parameters k_z and θ_z are then selected to ensure that the mass of women below these thresholds corresponds to the empirically observed female participation rates by female education conditional on the husband's education.

Table 2 shows the calibrated parameters, as well as the targeted data and model moments and their difference. Since we have more moments than parameters, we are not matching any moment perfectly. Hours worked per employed US married woman are 2.6% lower in the model than in the data, and hours worked per married man are 1.7% higher than in the data. The average female employment rates by the husband's education are matched almost perfectly. Employment rates of low and high educated women are however slightly higher in the model than in the data, and those of medium educated women are slightly lower.

[19] Robustness checks with respect to this parameter are shown in Online Appendix Section C.4.

Table 2: Data Targets from US and Calibrated Preference Parameters

	Parameters	Data	Model	$\Delta_{\text{Model-Data}}$
Hours Worked:	$\alpha = 0.469$			
HWP_m		1970	2003	33
HWE_f		1746	1701	-45
Female Employment Rates by Husband's and Own Education (in %)				
Low educ. husband:	$k_{low} = 1.064$, $\theta_{low} = 0.263$			
Low educ. woman		42.4	44.2	1.8
Medium educ. woman		63.3	60.3	−3.0
High educ. woman		76.1	77.6	1.5
Medium educ. husband:	$k_{med} = 1.073$, $\theta_{med} = 0.158$			
Low educ. woman		48.4	50.1	1.7
Medium educ. woman		71.2	68.6	−2.6
High educ. woman		83.2	85.8	2.6
High educ. husband:	$k_{high} = 0.521$, $\theta_{high} = 0.336$			
Low educ. woman		49.6	51.3	1.7
Medium educ. woman		66.2	63.0	−3.2
High educ. woman		73.7	75.2	1.5

4.6 Model Fit

Even though we only target average hours by gender, the model also replicates the gradients by education fairly well, see Table C.1 in the Online Appendix. The model is only somewhat off for the low educated: hours per employed low educated woman are underpredicted by 9%, and hours per person of low educated men are overpredicted by 18%. This is due to the relatively high non-employment rates among low educated men suppressing their hours worked per person, which the model does not capture.

We do a further validity check of the model by analyzing its performance in matching the time-series of hours worked of married couples in the US. To do that, we generate the US-specific model inputs back to the year 1979 and plug them into the model, keeping the preference parameters fixed. This exercise is in the spirit of Kaygusuz (2010), who analyzes only the time period from 1980 to 1990, which we extend through 2008. The model correctly predicts hardly any change in hours worked of married men over the period of three decades. For married women, the model captures both the increase in the employment rate and in hours worked per employed from 1979 to 2008 almost perfectly, with some deviations of the employment rates in model and data in the 1990s. Details are shown in Online Appendix Section C.3. Thus, this simple model can capture the time-series development of hours in the US very well.

5 Results

Keeping the preference parameters fixed across countries, we use country-specific labor income tax systems and consumption tax rates, plus wages and the educational composition (i.e. the educational distribution by gender and the degree of assortative matching), in order to obtain predicted hours worked of married couples across countries. We first present the cross-country predictions of hours worked per married person, and then break them down by gender. In a decomposition analysis, we evaluate the relative importance of consumption and labor income taxes in explaining the cross-country variations of married couples' hours worked. To understand the effects of international differences in labor income taxation better, we further decompose them into differences in average tax rates and differences in the tax structure, i.e. the tax treatment of married couples and the progressivity of the tax code. Last, we also analyze the relative importance of country-specific educational compositions and wages for the results, and their interactions with the labor income tax code. Online Appendix Section C.4 shows robustness checks of all results with respect to the presence of children in the household, the curvature of the disutility of hours worked ϕ, the share of government revenues redistributed to the households λ, and the estimated wages.

5.1 Hours Worked per Person in Model and Data

Table 3 shows in the first column the percent difference in married couples' hours worked per person between Europe and the US in the data, and in the second column the model predicted percent difference. The third column then divides these two numbers to obtain the fraction of the data difference that is explained by the model. These columns just focus on the *average* Europe-US gap in hours worked; country-specific results presented in levels of hours can be seen in Figure 4. Column 4 of Table 3 presents the correlation between hours in the data and the model, and column 5 the slope of a regression line of model predicted hours on data hours, as depicted in Figure 4. In both the regression and the calculation of the correlation, we exclude the US, which is the baseline country for the calibration.[20] Thus, these two columns do not refer to the Europe-US hours gap, but instead focus solely on the explanatory power of the model for the variation of hours worked *within* Europe. The first row of the table refers to married couples (summing hours of husband and wife), the second row to married men, and the third row to married women. The last two columns show the correlation of hours worked of married men and married women in the data (column 6) and in the model (column 7).

The model correctly predicts uniformly lower hours in Europe than in the US. Through the lens of the model, differences in consumption and labor income taxes, together with wages and the educational composition, can account on average for 72% of the hours worked difference of married couples between

[20]Thereby, the correlation is the same whether applied to the level of European hours as presented in Figure 4, or the deviation from US hours as presented in Table 3. Note that correlation and regression slope contain independent valuable information. The correlation captures how well the ordering of countries in the model reflects the one in the data, while the slope captures additionally how the standard deviation created by the model compares to the standard deviation in the data. Formally, $\beta = corr(h_{data}, h_{model})\frac{\sigma_{model}}{\sigma_{data}}$. A perfect model fit would result in both a correlation and slope of 1. However, it would theoretically be possible to obtain model results in which the correlation is 1 but the slope very low, or the slope is 1 but the correlation very low.

Table 3: Hours Worked per Married Person in Data and Model

	Eu.-US Hours Gap (all in %)			Variation within Europe		Corr. (HWP_m, HWP_f)	
	Data	Model	Fraction Explained	Correlation	Regr. Slope	Data	Model
Total	−13.7	−9.9	72.2	0.46	0.41	0.06	0.33
Men	−11.2	−6.5	57.7	0.43	0.47	–	–
Women	−17.7	−15.7	88.4	0.43	0.27	–	–

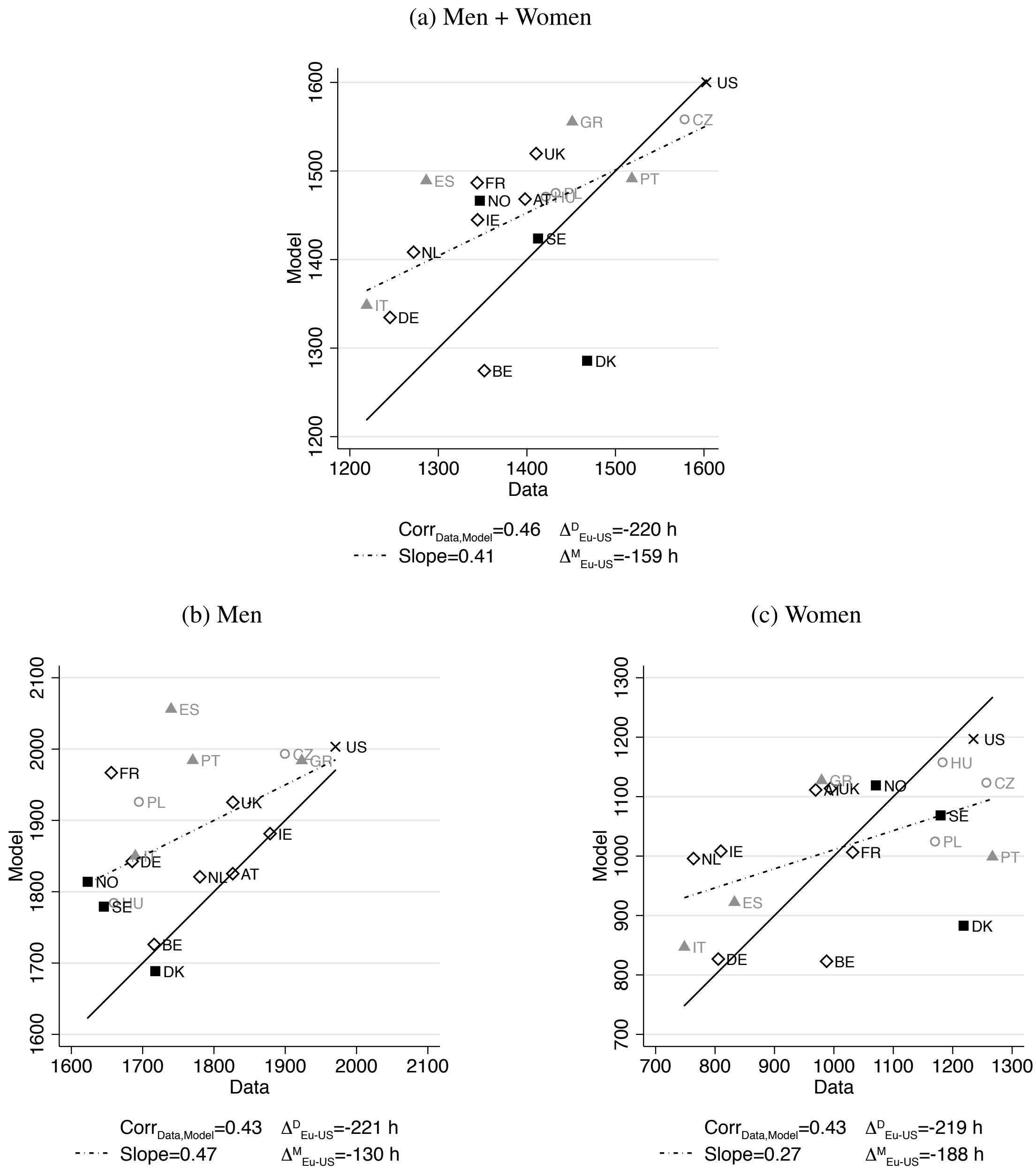
Note: Column 1 shows the Europe-US hours differences in % in the data, and column 2 the corresponding difference in the model. Column 3 shows the fraction of the data explained by the model. Column 4 indicates the cross-country correlation of model and data hours, and column 5 the slope of a regression line of model predicted hours on data hours, both times excluding the US. Column 6 states the cross-country correlation of hours worked of married men and married women in the data, and column 7 in the model.

Europe and the US. Regarding the cross-country variation within Europe, the correlation between model and data hours is 0.46, and the regression slope amounts to 0.41. As can be seen in Panel (a) of Figure 4, the model fit is universally quite decent. For half of the sample countries, the difference between hours in model and data is smaller than 100 hours annually. The poorest fits arise for Spain and Denmark, where hours in the model are 200 hours, or 16%, larger than in the data for Spain, and 180 hours, or 12%, smaller than in the data for Denmark. These results are driven by a poor fit for Spanish men and Danish women, as we will discuss later. Moreover, the model is largely able to replicate the low correlation of male and female hours worked per person in the data. As the last two columns of Table 3 show, the correlation amounts to 0.06 in the data, and to 0.33 in the model.

The next two rows of Table 3 break the previous results down by gender. Differences in our model features explain slightly more than half of the average Europe-US gap in married men's hours worked. The explanatory power of the within-Europe variation in married men's hours worked is very similar to the one for total hours worked. As panel (b) of Figure 4 shows, the fit is worst for France and Spain, where predicted hours worked of married men are 310 hours larger than hours in the data. The poor overall fit for Spain is thus driven by the poor fit for married men.

We now turn to hours worked of married women in the last row of Table 3. Through the lens of the model, differences in consumption and labor income taxes, together with wages and the educational composition, can account for 88% of the Europe-US hours gap in married women's hours worked. Not only the average Europe-US hours gap, but also the variation within Europe is explained fairly well: the correlation between model and data hours amounts to 0.43. However, the regression slope is smaller than for men with a slope of 0.27. As panel (c) of Figure 4 shows, the model fit is decent for the majority of countries. Importantly, on average the model correctly predicts the small hours differences to the US for Eastern European and Scandinavian countries, and the larger hours differences to the US for Western and Southern European countries. Two important outliers are however Denmark and Portugal. For Denmark, the model does not replicate the relatively high hours worked in the data, but predicts significantly lower hours than in the other Scandinavian countries, resulting in a difference of 335 hours worked between model and data hours.

Figure 4: Country-Specific Results: Hours Worked per Married Person



Note: $\text{Corr}_{\text{Data,Model}}$ refers to the cross-country correlation of model and data hours, and Slope to the slope of a regression line of model predicted hours on data hours as depicted in the figure, both times excluding the US. $\Delta^{\text{D}}_{\text{Eu-US}}$ gives the average hours difference in levels between Europe and the US in the data, and $\Delta^{\text{M}}_{\text{Eu-US}}$ in the model.

This drives the poor overall fit for hours of Danish married couples in panel (a). For Portugal, the model predicts relatively low hours worked, while in fact Portuguese married women work more hours than US married women, resulting in a difference of 270 hours between model and data hours. Denmark is an outlier relative to the other Scandinavian countries when it comes to the model predictions, and Portugal is an outlier relative to the other Southern European countries when it comes to the data. The drivers of the poor fit for Portugal and Denmark are discussed in the next subsection. Omitting Portugal and Denmark from the analysis increases the slope of the regression line to 0.44 and the correlation between model and data to 0.63.

5.2 Decomposition Analysis

To understand the relative importance of labor income and consumption taxes in explaining the cross-country differences in hours worked of married men and women, and the economic mechanisms driving the results, we simulate the model changing the input factors to country-specific levels separately and not all together. In Table 4, we present results for married men and married women setting only one feature of the economic environment country-specific and leaving all others as in the US. The table presents the average married male and female Europe-US hours gaps in the model (omitting the one in the data, which is reported in Table 3). With regard to the model fit of the within-European variation, we focus on married women only, and show the corresponding results for married men in Appendix Table C.5.[21]

The first row of Table 4 replicates the results from Table 3, in which the model is solved for each country with the full country-specific environment. The next six rows set only one single element in Equations (1) to (6) specific to the respective country, namely consumption tax rates (τ_c) in row 2 and non-linear labor income taxes (τ_l) in row 3, while keeping all others at the US level. Rows 4 and 5 further decompose the non-linear labor income tax schedules into the tax structure and the average tax rate as we describe below. Finally, wages (w) are set country-specific in row 6, and the demographic composition and matching into couples ($\mu(x,z)$) in row 7. In all these experiments, we adjust the transfers such that the government always maintains a balanced budget. When applying the country-specific labor income tax schedule, we account for the fact that tax systems are defined relative to the wage level in a country: for each combination of husband-wife hours choices, we calculate the country-specific tax rate using the US gender-specific education premia and the country-specific mean wage, and then apply this tax rate to the US gross earnings implied by the same husband-wife hours choices. Similarly, when applying the gender-education-specific wages from a European country, we leave the mean wage at the US level such that the US labor income tax schedule can be meaningfully applied. Details of these procedures are in Appendix A.

Consumption Tax Rates

Row 2 starts the decomposition with applying the country-specific consumption tax rates. Not surprisingly, given the universally higher consumption tax rates in Europe than in the US reported in panel (a) of Figure 3, consumption tax rates can explain a significant part of the lower hours of European men and women, namely 26% of the male hours gap and even 48% of the female hours gap. The disincentive effect of European

[21] Graphs showing country-specific results for women if each time only one feature is set country-specific, while the rest remain at the US level, are presented in Online Appendix Figure C.3.

Table 4: Decomposing the Effects of Different Model Inputs on Labor Supply

	Men		Women					
	Eu.-US Hours Gap		Eu.-US Hours Gap		Variation within Eu.		Corr. (HWP$_m$, HWP$_f$)	
Experiment	Model	% Expl.	Model	% Expl.	Corr.	Regr. Slope	Data	Model
All Country-Specific	−6.5	57.7	−15.7	88.4	0.43	0.27	0.06	0.33
τ_c	−2.9	26.2	−8.4	47.5	−0.24	−0.05	–	0.99
τ_l	−4.9	43.5	−2.0	11.3	0.17	0.13	–	0.56
Tax Structure	*-4.1*	*36.6*	*0.5*	*-3.0*	*0.27*	*0.13*	–	*0.32*
Avg. Tax Rate	*-0.8*	*6.9*	*-2.5*	*14.3*	*-0.01*	*-0.01*	–	*1.00*
w	−0.5	4.2	−0.7	3.8	0.06	0.02	–	0.13
$\mu(x,z)$	2.2	−19.2	−6.6	37.3	0.36	0.12	–	−0.79

Note: Column 1 shows the predicted Europe-US hours differences in % in the model for married men, column 2 the fraction of the data explained by the model for married men. Columns 3 and 4 show the same two statistics for married women. Column 5 gives the cross-country correlation between model and data predicted hours for married women, and column 6 the slope of a regression line of model predicted hours on data hours for married women, both times excluding the US. Column 7 states the cross-country correlation of hours worked of married men and married women in the data, and column 8 in the model. Row 1 shows the full model results if all input factors are set country-specific. For the decompositions in rows 2 to 7, exactly one model input is set country-specific, and the rest are left unchanged at their US values.

consumption taxes is smaller for men, predicting 3% lower hours worked than in the US, than for women, predicting 8% lower hours worked. This is due to the higher implied female elasticity of labor supply, which arises because women face lower wages and are affected both along the extensive and intensive margin.[22] The cross-country correlation between male and female hours if only the consumption tax is set country-specific amounts to 0.99, as can be seen in the last column of Table 4: qualitatively, consumption taxes affect male and female hours in exactly the same way.

While consumption taxes are the overall most important driver of the Europe-US hours difference for married women, they do not help explaining the variation of married women's hours worked *within* Europe. In fact, the regression slope is almost zero, and the correlation between hours in model and data negative when only consumption taxes are set country-specific (see columns 5 and 6). Consumption tax rates are largest in the Scandinavian countries, which however feature comparatively high hours worked of married women, such that the resulting correlation between hours in model and data is negative.[23]

[22]Note that, even if we would model an extensive margin for men, we would expect only minimal effects along this margin due to their already very high participation rates.

[23]For the same reason, the correlation between hours in model and data is in fact positive for married men if only the consumption tax rates are set country-specific (see Appendix Table C.5). Scandinavian countries feature the lowest hours of European married men.

Non-Linear Labor Income Taxes

While for women consumption taxes are on average the main factor driving the lower European hours, labor income taxes are the main factor explaining low European male hours: labor income taxes alone predict 4.9 percent lower hours of married men in Europe than in the US, thereby explaining 44 percent of the male Europe-US hours gap. Their disincentive effect on male hours worked is thus larger than the disincentive effect coming from consumption taxes, which explain only 26 percent of the Europe-US hours gap. For married women, however, the picture now looks very different. Labor income taxes in fact predict only 2 percent lower hours worked for European married women than US ones, thus explaining only 11 percent of the hours gap, compared to 48 percent explained by consumption taxes. Yet, they do matter for explaining the variation of hours worked of married women *within* Europe: the correlation between model and data amounts to 0.17, and the regression slope is 0.13 if only labor income taxes are set country-specific. Another striking difference to consumption taxes, which induce a near perfect correlation between hours worked of men and women, is the lower cross-country correlation of 0.56 between hours worked of married men and women if only the European labor income tax systems are introduced in the US economy (see last column).

What can explain that labor income taxes have very different effects on the labor supply of married men and married women? We derive more insights into this by further decomposing the non-linear labor income tax schedule into two components: first, the tax structure, and second the average tax rate. The tax structure captures the progressivity of the tax schedule along with the tax treatment of married couples, which together define how the marginal tax rate of each spouse changes with the own and the spousal income. We distinguish between the tax structure and the average tax rate by conducting the following experiment: to capture the effect of the tax structure alone, we calculate the taxes implied by the country-specific tax code as in Equation (A.2), but then levy an additional linear tax or subsidy on the household such that government revenues are left unchanged at the US level.[24] Put differently, one may think of this experiment as a reform which implements a different tax structure but is required to be revenue neutral. The effect of the country-specific average tax rate is then indirectly inferred by the difference in hours worked between setting the entire labor income tax schedule country-specific, or shifting it up or down to match the US government revenues.

The fourth and fifth rows of Table 4 show the resulting decomposition of the labor income tax effect into the tax structure and the average tax rate. For married men, the tax structure predicts 4.1 percent lower hours worked in Europe than in the US, but the average tax rate only 0.8 percent lower hours. Thus, the average labor income tax rates contribute only minimally to explaining the lower hours of married men in Europe than in the US, but the European tax structure alone explains 37 percent of the hours difference to the US. For married women, on the other hand, the average tax rate predicts 2.5 percent lower hours worked in Europe than in the US, but the tax structure 0.5 percent *higher* hours worked. While on average across all European countries, the effect of average tax rates is thus qualitatively similar for married men and women,

[24] The household's income tax liability (Equation 5) hence becomes $\tau_l = \tau_l(w_m(z)h_m, w_f(x)h_f, k) + \theta y_{hh}$, with $\theta > 0$ being an additional linear tax and $\theta < 0$ being a subsidy.

the effect of the tax structure goes in the opposite direction for both sexes. This also becomes evident in the predicted correlation between male and female hours, which amounts to 1 if the average tax rate is set country-specific, and to only 0.32 if the tax structure is set country-specific.

To explain the opposite effect of imposing the European tax structures on hours worked of married men and married women, we refer to the predictions of the theoretical model in Subsection 4.2. Moving from joint towards separate taxation decreases the optimal hours gap between spouses, i.e. increases hours worked of married women and decreases hours worked of married men. Varying the system of taxation across countries thus induces a negative cross-country correlation between married men's and women's hours worked. Specifically, the degree of joint taxation is more pronounced in the US than in most European countries. This is indicated in panel (b) of Figure 3: the difference between both suggestive tax rates is for all countries – except Denmark, Belgium, and Germany – smaller than in the US. Based on the tax treament of married couples alone, the US thus has a higher optimal hours gap than a European country with separate taxation. On average, this leads to lower predicted hours of married men and higher ones of married women if the tax structure is set to the European level.

In addition to the tax treatment of married couples, the tax structure also incorporates progressivity, whose effect on hours interacts with the tax treatment. The tax codes of European countries on average feature higher levels of progressivity than the US one, as shown in an illustrative example in Figure C.1 in the Online Appendix. The effect of progressivity on hours worked generally depends on the position of an individual's income in the earnings distribution of the country. In a system of joint taxation, increased progressivity leads to lower hours of both married men and married women, as long as married couples are relatively high income earners. Under separate taxation, increased progressivity predicts on average lower hours worked for married men, who tend to be high income earners, while the predicted effect on married women's hours worked is somewhat unclear, depending on their relative position in the earnings scale of the respective country. Therefore, also the predicted cross-country correlation between hours worked of married men and women based on international variation in progressivity could be either positive or negative.[25] For men, the two effects of the tax structure – coming through the tax treatment and through progessivity – go on average in the same direction if European tax structures are imposed on US married men. This leads to the large explanatory power of the tax structure for the male Europe-US hours gap.

While the tax structure on average does not contribute at all to explaining the lower hours worked of married women in Europe than in the US, it is crucial for explaining the variation of married women's hours worked *within* Europe. As columns 5 and 6 show, the correlation between hours of married women in model and data is 0.27 when only the tax structure is set country-specific, and the slope of the regression line is 0.13. The tax treatment, together with the progressivity of the tax system, thus leads to variation in predicted hours worked across the European countries that is in line with the data. Apart from the educational composition, which we discuss below, the tax structure has the largest explanatory power for the variation of married women's hours worked within Europe.

[25] Note that we cannot report the effects and implied correlations from tax treatment and progressivity separately, since our framework does not allow us to isolate each effect individually.

Turning to the average labor income tax rates, their effects on hours worked are qualitatively the same for men and women, with again larger effects for women due to the higher implied female elasticity, similar to the effects of consumption tax rates. Quantitatively, the disincentive effects of consumption taxes are larger than those of the average labor income tax rates for both men and women, which is in line with the larger consumption tax rate differences between Europe and the US shown in panel (a) compared to panel (b) of Figure 3. Higher average labor income tax rates predict 2.5 percent lower hours in Europe than in the US for married women, corresponding to 14 percent of the total hours gap. Yet, differences in the average tax rate do not aid in explaining the variation of married women's hours worked across European countries.[26]

Summarizing, to explain the variation of married women's hours worked across Europe, it is crucial to take differences in the tax structure into account.[27] However, since the US features a system of joint taxation, while many European countries feature tax systems closer to separate taxation, the tax structure cannot explain the on average lower hours worked of European than US married women. For married men, by contrast, the tax structure predicts significantly lower hours worked in Europe than in the US due to higher levels of progressivity and tax systems closer to separate taxation in Europe. Average tax rates predict on average slightly lower hours worked in Europe than in the US for both married men and women.

Wages

The effect of the country-specific gender and education wage premia on hours worked in Europe is on average relatively small, as can be seen in the sixth row of Table 4. Wages explain 4% of the average Europe-US hours gap for both men and women, and have only small explanatory power for the variation of married women's hours worked within Europe. However, introducing country-specific wage gaps leads to a low correlation of hours worked of married men and women, as one would expect at least for gender wage gaps. The wage premia interact of course with the tax systems in the respective countries. Thus, one could potentially obtain different effects of the tax structure if the wage premia are set first to the European level. We discuss this interaction effect below in conjunction with Figure 5.

Educational Composition

Turning to the educational composition and matching into couples in row 7, one can see that imposing European data induces a small increase in hours worked for married men of 2.2%, but a significant decrease of hours worked of married women of 6.6%, thus explaining 37% of the Europe-US hours gap for married women. The educational composition is after consumption tax rates the second most important model input explaining on average lower hours of married women in Europe than in the US.[28] The large effect of

[26]For married men, by contrast, the average tax rate is a more important driver of within-Europe variation in hours worked than the tax structure; see Appendix Table C.5.

[27]Online Appendix Section C.5 compares the model performance to a model relying on linear tax rates as calculated by Prescott (2004), McDaniel (2011) and Ragan (2013) as inputs, and confirms that a model with linear taxes cannot explain hours worked of married women in Europe well.

[28]Bick et al. (2016) conduct a simple decomposition exercise that decomposes the total Europe-US hours difference into differences of the components of labor supply (employment rates, weeks worked, and weekly hours worked in a work week) and the

imposing the European educational composition on women's hours worked arises because the gradient of employment rates by education is substantial in the US, while the gradient of hours worked per employed by education is relatively small. Thus, imposing a different educational composition has much larger effects on married women, whose employment rates differ by education in the model, than on married men. Given that the share of high-educated individuals is highest in the US, and the share of low-educated is lowest, as can be seen in panel (c) of Figure 3 for married women, imposing the European educational composition significantly decreases predicted hours worked of married women there. Since men married to non-working women work longer hours, having more of these men drives up average predicted male hours worked in Europe, despite the fact that the shares of low- and medium-educated married men are also higher in Europe than in the US (see Appendix Table B.11).

Moreover, the educational composition contributes significantly to explaining the variation of hours worked of married women within Europe: the correlation between married women's hours in model and data is 0.36 if only the educational composition is set country-specific. The high share of low educated married women in Southern Europe predicts significantly lower hours there, which is in line with the data. Last, the country-specific educational composition leads to a strong negative cross-country correlation of -0.79 between male and female hours. Yet, one has to keep in mind that it causes only a small variation in male hours, and thus cannot contribute significantly to explaining the low correlation of male and female hours worked in the data overall. As for wages, the educational composition and especially the matching into couples interact with the tax structure of a country, which we discuss now.

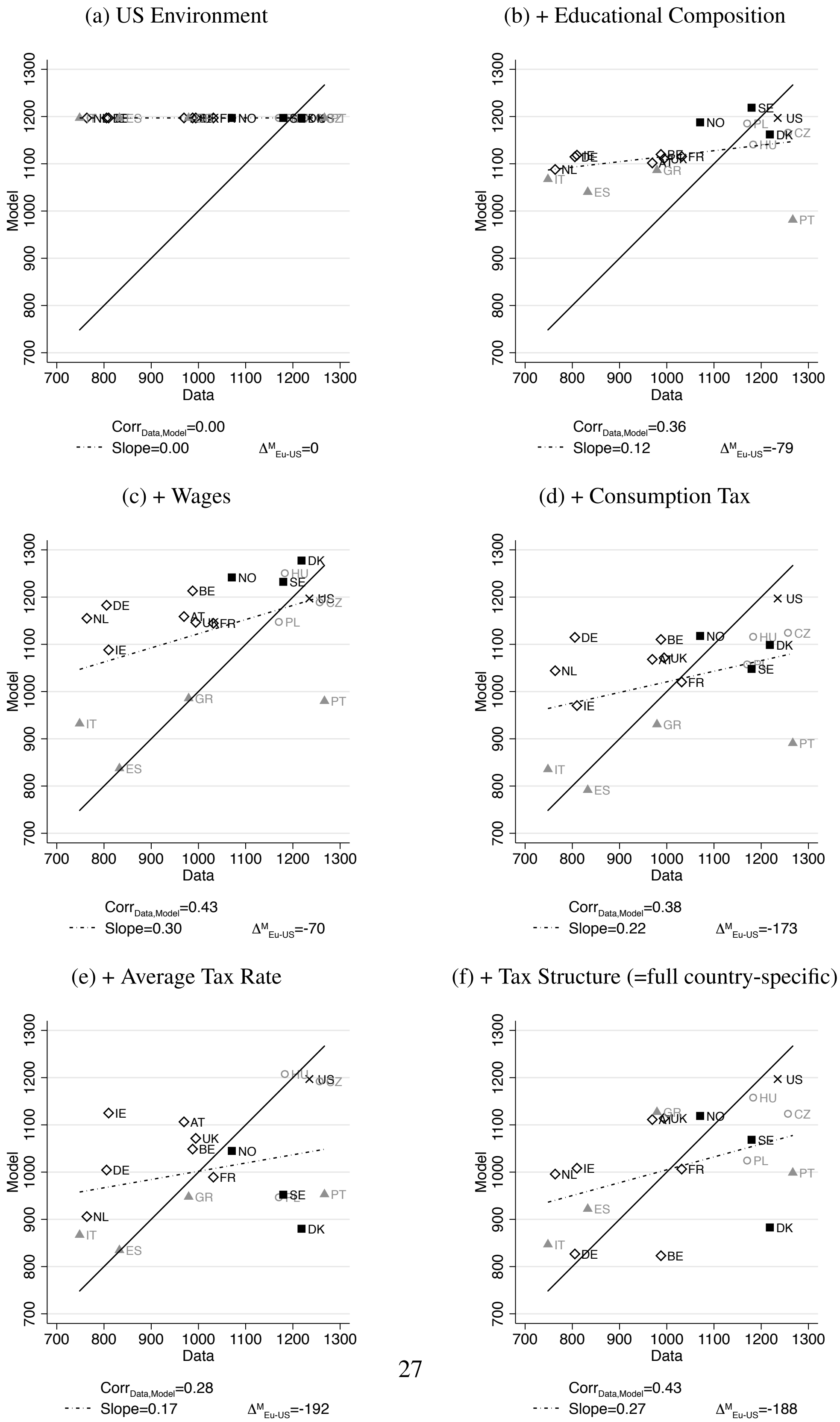
Interaction Effects and Country-Specific Results

In the decomposition exercise carried out so far, we left all model features at the US level and set only one specific feature to the country-specific level. To understand in how far the labor income tax system, and especially the tax structure, interacts with wages and the educational composition, we now conduct a cumulative decomposition exercise in which we set the model features to the country-specific level in a cumulative fashion. We start out by setting the educational composition and wages country-specific, and put the labor income tax structure last. We focus in this analysis on hours worked of married women. Figure 5 presents the results, and also allows analyzing individual countries. It starts in panel (a) with setting all model inputs at the US level, predicting the same hours as in the US for all countries. It then adds the different country-specific inputs sequentially into the model.

Panel (b) introduces the country-specific educational composition into the model. The high correlation between model and data hours if the educational composition is set country-specific from the last row of Table 4 is visualized here. The educational composition tilts the model predicted hours in the right direction: given the higher share of low- and medium-educated married women, it reduces predicted hours in Southern and Western Europe, and to a much smaller degree in Eastern Europe. Hours in Scandinavian countries are

demographic composition. They find that one third to one half of the Europe-US hours difference of all individuals aged 15 to 64 can be attributed to differences in the educational composition.

Figure 5: Decomposition – Cumulative Country-Specific Results for Hours Worked of Married Women



Note: Corr$_{\text{Data,Model}}$ refers to the cross-country correlation of model and data hours, and Slope to the slope of a regression line of model predicted hours on data hours as depicted in the figure, both times excluding the US. $\Delta^{\text{M}}_{\text{Eu-US}}$ gives the average hours difference in levels between Europe and the US in the model. In each panel of the figure, one additional model feature is set to the country-specific level, namely the one indicated in the subheading.

on average unchanged. Thus, especially for Southern Europe, and to some degree for Western Europe, the educational composition is a significant factor explaining lower hours worked than in the US. Yet, at the same time the educational composition is the main reason for the poor fit of hours worked of Portuguese married women. As panel (c) of Figure 3 shows, the share of low educated women is by far the highest in Portugal, amounting to over 70%. Thus, the educational composition alone predicts 216 fewer hours in Portugal than in the US, while in fact Portuguese married women work more hours than US married women.

Panel (c) adds country-specific wages to the educational composition.[29] Adding the country-specific education-gender-wage premia leaves the average hours difference with the US almost unchanged, but improves especially the regression slope, which rises to 0.30. As one can see in the figure, this is driven by the negative effect of wages on predicted hours in the three Southern European countries Italy, Greece, and Spain. This creates more variation in the model predicted hours, which given a slightly increased correlation results in a higher regression slope. As panel (d) of Figure 3 shows, education premia tend to be high in Southern European countries by international comparison. After imposing the high share of low-educated individuals there in panel (b), adding the education premia therefore decreases predicted hours in Southern Europe strongly. This highlights a strong interactive effect between the wage structure and the educational composition, since imposing only European wage premia results in a regression slope of essentially 0, as Table 4 shows.

Panel (d) then adds on top of this the country-specific consumption tax rates. Consumption taxes shift predicted hours in Europe substantially further downward, by on average 103 hours. They are thus the most important driver of low average hours of European married women compared to US married women. However, this downward shift occurs in an almost parallel fashion, with only slightly larger effects for the Scandinavian countries. Thus, consumption taxes do not explain any additional variation *within* Europe: the slope of the regression line of model hours on data hours falls, as does the correlation between hours in model and data.

Panel (e) now uses the average labor income tax rate as an additional country-specific model input. Average European hours fall by 19 hours compared to panel (d). Average tax rates thus overall predict slightly lower hours in Europe than in the US. However, the predicted hours in Ireland increase substantially, and somewhat in a few other European countries, largely those with low average tax rates depicted in panel (b) of Figure 3. Most importantly, hours in the Scandinavian countries and specifically in Denmark decrease substantially. The high average tax rate is thus the main driver of the low predicted hours by Danish married women. The slope of the regression line and the correlation both again decrease once the European average tax rates are added as model inputs. Thus, adding the average tax rates only modestly helps in explaining the average US-European hours difference, but not in explaining the variation of hours of married women within Europe.

Panel (f) finally adds the tax structure of the labor income tax code and presents the full country-specific results, already shown in Panel (c) of Figure 4. Predicted hours in Europe slightly increase by 4 hours

[29] Given log utility of consumption, differences in mean wages do not affect choices at all, and only the education-gender-wage premia matter, see Online Appendix Section C.1.

compared to panel (e), i.e. on average the effect of the European tax structure goes in the wrong direction, as seen before. However, the tax structure substantially helps improve the fit within Europe. The slope increases by 0.10 (from 0.17 to 0.27), and the correlation by 0.15 (from 0.28 to 0.43). While these increases are substantial, they are somewhat smaller than the ones in the individual decomposition (0.13 and 0.27, respectively), in which wages and the educational composition are left at the US-level. Thus, there exist interaction effects between wage premia and the educational composition on the one hand and the non-linear labor income tax code on the other hand. Still, our results of the importance of the tax structure for explaining the variation of married women's hours worked in Europe are robust to different orderings of the decomposition.[30]

Imposing the country-specific tax structure improves the fit substantially for Germany, Sweden, Poland, and Ireland. Specifically, for Germany the tax structure decreases hours substantially, while for Sweden the opposite happens. As we discuss in our motivating example in the introduction, Germany is a country with joint taxation, while Sweden features separate taxation. For Germany, the difference between the illustrative average marginal tax rate for married women and the average tax rate for single earner households depicted in panel (b) of Figure 3 is larger than in the US, indicating even stronger elements of joint taxation than in the US.[31] A few countries move however further away from the hours in the data at this step. Specifically, one can see the positive effect of the tax structure on Greece and Spain, which makes it harder to predict low hours in these two countries.

6 The Two Margins of Labor Supply

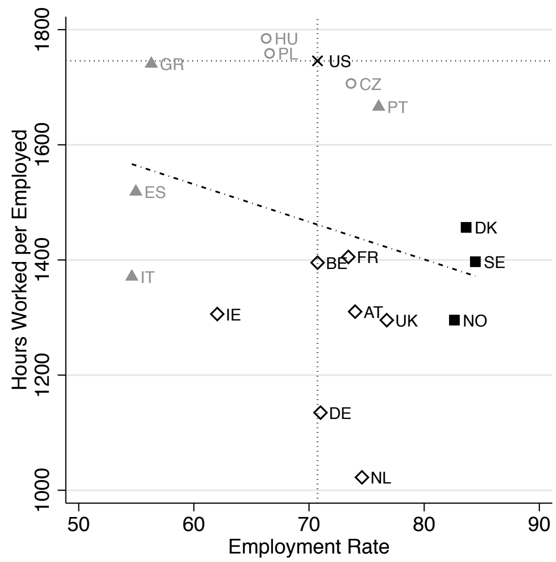
6.1 Extensive and Intensive Margin in Data and Model

So far, we only discussed data and results for hours worked per person. For married women, however, we model an extensive and an intensive margin, and we can thus compare the results for both margins to the data. Figure 6 decomposes hours worked per married woman in the data into the extensive margin, i.e. the employment rate, on the horizontal axis, and the intensive margin, i.e. hours worked per employed, on the vertical axis. Figure 2 showed that Scandinavia and Eastern Europe both exhibit high hours worked of married women. Figure 6, by contrast, documents stark differences between both regions for the extensive and intensive margins: relative to the US, Scandinavia features high employment rates but low hours worked per employed, while the decomposition into both margins for Eastern Europe resembles the one in the US. Similarly, Southern Europe (with the exception of Portugal) and Western Europe both exhibit low hours

[30] The slope increases by 0.17 (0.08) and the correlation by 0.10 (0.12) if only the educational composition (only wages) are set country-specific before the country-specific tax structures are imposed.

[31] The comparison of the model results to the illustrative tax measures in panel (b) of Figure 3 however also makes it clear that the full tax code is much more complex than the two illustrative tax rates show. E.g. adding the Irish tax structure predicts a substantial decrease in hours worked, despite a lower illustrative average marginal tax rate there than in the US which is even slightly smaller than the average tax rate of the single-earner household, see panel (b) of Figure 3. The Irish tax system features very strong progressivity, such that both tax rates and the difference between them rise rapidly with individual income (see also Figure C.1 in the Online Appendix for one illustrative measure of progressivity).

Figure 6: Average Employment Rates and Hours Worked per Employed of Married Women (Ages 25-54)



worked per married woman, but again are clearly distinguishable in the decomposition into the extensive and intensive margin. Relative to the US, Western Europe features high employment rates and low hours worked per employed, while Southern Europe has low employment rates, with hours worked per employed being lower than in the US, but higher than in Western Europe. Portugal resembles the Eastern European countries not only in its high level of hours worked of married women, but also in its decomposition into both margins. The resulting cross-country correlation of hours worked per employed and employment rates of married women in the data is negative with a value of -0.29. Low hours worked per employed married woman in the data are largely driven by a high share of part-time working women, and vice versa. The cross-country correlation between hours worked per employed woman and the share of women working less than 30 hours is -0.9. We get back to the issue of part-time work in Section 6.2.[32]

Table 5 shows the performance of the model in matching extensive and intensive margin differences relative to the US. The first column shows employment rate differences relative to the US (in percentage points) in the data and the model, while the second column shows hours worked per employed differences (in %). On average, the model overpredicts the employment rate differences to the US, and underpredicts the differences in hours worked per employed.[33] As column 3 shows, the model features a positive correlation of 0.32 between married women's employment rates and hours worked per employed, while in the data the correlation is -0.29. It is not entirely surprising that taxes, similarly to wages and the educational composition, have difficulties explaining the vastly different decompositions into the intensive and extensive margin

[32]Figure B.7 in the Online Appendix shows the decomposition into both margins for married men. There is little variation in employment rates of married men, which lie always above 90%. The only exceptions are Hungary and Poland, where lower employment rates are driven by older married men and are probably a phenomenon of the transition from Socialism to Capitalism.

[33]Online Appendix Figure C.4 shows the country-specific results for employment rates and hours worked per employed.

Table 5: Married Women's Employment Rate (ER) and Hours Worked per Employed (HWE)

	Eu.-US Gap		
	ER	HWE	Corr. (ER, HWE)
Data	0.0	−17.2	−0.29
Model	−7.5	−5.8	0.32

Note: For the Europe-US gap in employment rates, we show percentage point differences, and for hours worked per employed % differences.

in Scandinavia and Western Europe on the one hand, and Eastern and Southern Europe and the US on the other hand visible in Figure 6. The model factors affect both margins in the same direction, and thus the predictions relative to the US are qualitatively similar across both margins, resulting in a positive correlation between both margins.

6.2 Wedges and the Margins of Labor Supply

What can explain the negative correlation between employment rates and hours worked per employed by married women in the data? Many institutional, policy, and regulatory factors other than those featured in our model affect hours worked, and especially the decisions whether to work or not, and how many hours to work conditional on working. In order to get some insights into which factors could generate a negative correlation, we follow the approach pioneered by Chari et al. (2007) and Restuccia and Rogerson (2008), and introduce two country-specific wedges into the model, $\hat{\tau}_E$ and $\hat{\tau}_{h_f}$, in order to match the country-specific female employment rates and hours worked per employed. The wedges pick up any differences between the model's predictions and data that are not driven by taxes, wages, and the educational composition. They could potentially capture very different factors, such as restrictions on hours or employment from the labor demand side, child care, or maternal leave. Below we correlate these non-modeled factors with our wedges and explore which of them may be particularly important for understanding the discrepancy between the model and the data in the decomposition into both margins.

The two wedges affect the household's decision problem in the following form. Instead of Equation (2), households maximize the following objective function, in which the wedges are added at the end:

$$\max_{h_m,h_f}\left\{\ln c-\alpha h_m^{1+\frac{1}{\phi}}-\alpha h_f^{1+\frac{1}{\phi}}-\mathbf{1}_{\{h_f>0\}}q+\mathbf{1}_{\{h_f>0\}}\hat{\tau}_E+\hat{\tau}_{h_f}\cdot h_f\right\} \tag{11}$$

The wedge $\hat{\tau}_E$ works as an implicit subsidy or tax on female employment. A positive $\hat{\tau}_E$ increases the utility of working for women, and therefore at least partly offsets the fixed costs of working q associated with positive female hours worked. It only affects the extensive margin choice of married women. $\hat{\tau}_{h_f}$, on the other hand, represents an implicit subsidy or tax on female hours worked. It directly affects the intensive margin choice by entering the first order condition for female hours conditional on working: a negative $\hat{\tau}_{h_f}$ increases the disutility of work for every hour of work. Thereby, it also has an indirect effect on the extensive margin choice, by changing the utility of working a certain amount of hours vis-à-vis the utility of not working. We calibrate both wedges to match the country-specific employment rates and hours worked per employed married woman.

By construction, for countries for which the model predicts too high (low) hours or employment rates relative to the data, the respective wedge is negative (positive). The larger the difference between the predictions of the model without wedges and the data, the larger is the absolute value of the calibrated wedge, see Figure C.5 in the Online Appendix. Moreover, since the model predicts a positive correlation between employment rates and hours worked per employed rather than the observed negative correlation, the wedges are negatively correlated with a value of -0.71.

Correlation of Wedges with Country-Specific Factors

To understand which factors could drive the two margins in opposite directions, we follow the idea in Ohanian et al. (2008) and correlate the wedges with country-specific variables that are potentially important for women's labor supply choices and the decomposition into both margins. Table C.6 in the Online Appendix shows the country-specific wedges, as well as the country-specific values of all variables that we correlate with the wedges.

The variable that features one of the strongest correlations with both wedges is an OECD country ranking of part-time generosity. A higher part-time generosity rank implies more extensive rights of the worker to demand part-time work, fewer grounds for refusal from part of the employer, the right to go back to full-time work, as well as part-time legislation being in place for a longer time period. As Table 6 shows, the OECD part-time generosity rank exhibits a positive correlation of 0.44 with the hours wedge, and a negative correlation of -0.43 with the employment wedge. Part-time generosity also intuitively explains why both margins could be negatively correlated in the data. Suppose that there exists a group of women who find it optimal to work but to supply only relatively few hours. If part-time jobs are not available in a country, these women might prefer not working over working full-time, resulting in a low average employment rate and high average hours worked per employed. In a country with more generous rights to work part-time, these women will enter the work force at the desired low hours, thereby increasing average employment rates and decreasing average hours worked per employed. To give a concrete example, the US, our benchmark country, has the lowest OECD part-time generosity rank, while Sweden has the highest rank and has one of the highest employment and hours wedges (in absolute values). Thus, the differential regulation of part-time work across countries is a promising feature that could potentially explain the negative correlation of hours worked per employed and employment rates that we see in the data.

Table 6 also shows correlations of both wedges with other potential driving forces, namely the number of vacation days and public holidays, maternity leave policies (weeks of maternity leave, and pay replacement rates during the leave), child care variables (net child care cost, public child care expenditure over GDP, percent of preschool children in informal care, and hours of informal care conditional on receiving informal care), and divorce rates. The number of vacation days and public holidays exhibit a positive correlation with

Table 6: Correlation of Wedges with Different Factors

Policies	$\hat{\tau}_h$	$\hat{\tau}_e$
Part-time generosity rank	0.44	−0.43
Annual leave + public holidays	−0.28	0.42
Maternity leave: paid weeks	0.26	−0.20
Avg. pay during maternity leave	0.12	−0.02
Net child care costs (% of avg. earnings)	−0.33	0.05
Public child care expenditure (% of GDP)	−0.20	0.64
% of preschool child. in informal care	0.02	−0.41
Avg. hours of informal care for preschool child.	0.40	−0.38
Divorce rate (per 1000 persons per year)	0.15	0.20

Note: This table shows correlation coefficients between the two wedges and different factors. Part-time generosity is the reverse of the OECD part-time generosity rank, i.e. a higher rank implies higher generosity. Average pay during paid maternity leave is the proportion of gross earnings replaced by the benefit over the length of the paid leave entitlement for a person with average earnings. Net child care costs are child care costs minus benefits for two children aged 2 and 3 in full-time child care as % of the household income for a couple in which both earn the average wage. Divorce rates are the number of divorced marriages per year per 1000 inhabitants. All data come from the OECD with the exception of annual leave and public holidays, which are taken from Bick et al. (2016).

the employment wedge and a negative one with the hours wedge, indicating that more generous vacation time might induce more women to start working at lower annual hours worked. Thus, it works similarly to part-time generosity, but the correlation with the hours wedge is less strong. The correlations of both wedges with leave policies are generally low. Surprisingly, the correlation of child care costs with the employment wedge is small, and with the hours wedge of the opposite sign than one would expect. This finding is especially surprising in light of the substantial evidence in the literature that child care matters for female labor supply.[34] It could stem from the difficulties of constructing a comparable and appropriate measure of child care costs for all countries. In fact, the correlation of the employment wedge with public child care expenditure over GDP is high, and both correlations go in the expected direction. It might be the case that the measure of public child care expenditure is of better quality than the measure of net child care cost to the household. Additionally, public child care expenditure might capture substantial variation in quality or availability of slots across countries, which are not captured by a measure of net child care cost. Moreover, informal child care arrangements also play an important role in practice. According to the OECD, on average across our sample countries 28% of preschool children receive some form of informal child care, with the average hours conditional on receiving this form of child care being 15. Informal child care is strongly negatively correlated with public expenditures on child care.[35] This suggests that

[34]Relying on quantitative models calibrated to one single country, namely Germany or the US, respectively, Domeij and Klein (2013) and Guner et al. (2016) find that subsidizing child care is an effective policy if the government wants to increase female labor supply. Guner et al. (2016) additionally show that child care subsidies increase female employment rates, but decrease female hours worked. Last, Attanasio et al. (2008) find that the reduction in child care costs in the US over time helps in explaining the increase in labor supply of mothers.

[35]This correlation with public expenditures on child care is -0.65 for the fraction of children receiving informal care, and -0.41

private arrangements serve at least to some degree as substitutes for government child care policies, see e.g. Bick (2016) for the case of Germany. Last, we correlate the wedges with divorce rates, as suggested by Chakraborty et al. (2015), but correlations are low. Measuring and modeling part-time regulation or child care policies across countries is far from trivial, and we abstain from doing it in this paper.

Decomposition With and Without Wedges

One might be worried that the effects of taxes on hours worked would be different if both margins of labor supply were matched perfectly across countries, i.e. in the model with wedges. To address these concerns, we repeat the decomposition exercise in the model with country-specific wedges. In this model, we take the US environment with country-specific wedges as the benchmark, and then set either the consumption tax rate or the labor income tax rate country-specific and compare hours in the two cases. The incentive effects of the two taxes and the other model components are quite similar in the models with or without wedges, both qualitatively and quantitatively, see Online Appendix Table C.7.

7 Conclusion

Relying on three micro data sets, we document average hours worked of married couples for a sample of 17 European countries and the US over the time period 2001 to 2008. We find that hours worked vary significantly across countries, and substantially more for married women than for married men. Moreover, the cross-country correlation between hours worked of married men and married women is very low.

We investigate in how far international differences in consumption taxes and labor income tax systems can quantitatively account for the international differences in hours worked by married couples. We do this in the context of a static model of joint labor supply, calibrated to the US and holding preferences constant across countries. In addition to labor income and consumption taxes, we let gender and education wage premia, as well as the educational composition and matching into couples, vary across countries. These model features can account for 58% of the US-Europe gap in male hours worked, and 88% of the US-Europe gap in female hours worked, and also replicate the cross-country variation within Europe well: the cross-country correlation between hours worked in the model and the data is 0.43 for both men and women. Last, the model generates a low correlation of hours worked of married men and women.

Differences in consumption taxes are the main driving force of the *average* hours worked difference between Europe and the US for married women. The key to the success of the model in explaining the substantial *within*-Europe variation in married women's hours worked, and in breaking the cross-country correlation of married men's and women's labor supply, is the explicit modeling of non-linearities in the labor income tax code. The tax treatment of married couples, which ranges across countries from separate to different degrees of joint taxation, interacts with the progressivity of the tax system in affecting labor supply decisions of both spouses differently across countries. Differences in the educational composition

for average hours in informal care.

across countries further help in explaining low female hours worked in Southern and Western Europe. For married men, labor income taxes are the main driver of low European hours worked, aided by consumption taxes.

While the results leave scope for other factors explaining hours worked of married women (e.g., a home production sector in competition with the service sector of the economy, income risk, or a life-cycle component, which could be especially helpful to capture both margins of labor supply), labor income and consumption taxes, together with wages and the educational composition, have large explanatory power for international differences in hours worked of married women through the lens of the model. It is however crucial to model non-linear labor income tax systems in order to replicate the behavior of married women. The origins of the different decomposition of married women's hours worked into the extensive and the intensive margin across countries remain as an open question for future research.

References

Albanesi, S. and C. Olivetti (2009). Home Production, Market Production, and the Gender Wage Gap: Incentives and Expectations. *Review of Economic Dynamics 12*(1), 80–107.

Albanesi, S. and C. Olivetti (2015). Gender Roles and Medical Progress. *Journal of Political Economy* (forthcoming).

Alonso-Ortiz, J. (2014). Social Security and Retirement Across the OECD. *Journal of Economic Dynamics & Control 47*, 300–316.

Attanasio, O., H. Low, and V. Sanchez-Marcos (2008). Explaining Changes in Female Labor Supply in a Life-Cycle Model. *American Economic Review 98*(4), 1517–1542.

Bar, M. and O. Leukhina (2009). To Work or Not to Work: Did Tax Reforms Affect Labor Force Participation of Married Couples? *B.E. Journal of Macroeconomics 9*(1).

Benabou, R. (2002). Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? *Econometrica 70*(2), 481–517.

Bick, A. (2016). The Quantitative Role of Child Care for Female Labor Force Participation and Fertility. *Journal of the European Association 14*(3), 639–668.

Bick, A., B. Brüggemann, and N. Fuchs-Schündeln (2016). Hours Worked in Europe and the US: New Data, New Answers. Working Paper.

Bick, A. and N. Fuchs-Schündeln (2017). Quantifying the Disincentive Effects of Joint Taxation on Married Women's Labor Supply. Working Paper.

Blundell, R. and T. MaCurdy (1999). Labor Supply: A Review of Alternative Approaches. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Chapter 27, pp. 1559–1695. Elsevier Science B.V.

Bridgman, B., G. Duernecker, and B. Herrendorf (2016). Structural Transformation, Marketization, and Household Production Around the World. Working Paper.

Buera, F. J., J. P. Kaboski, and M. Q. Zhao (2014). The Rise of Services: the Role of Skills, Scale, and Female Labor Supply. Working Paper.

Chade, H. and G. Ventura (2002). Taxes and Marriage: A Two-Sided Search Analysis. *International Economic Review 43*, 955–986.

Chade, H. and G. Ventura (2005). Income Taxation and Marital Decisions. *Review of Economic Dynamics 8*, 565–599.

Chakraborty, I., H. A. Holter, and S. Stepanchuk (2015). Marriage Stability, Taxation and Aggregate Labor Supply in the US vs. Europe. *Journal of Monetary Economics 72*, 1–20.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007). Business Cycle Accounting. *Econometrica 75*(3), 781–836.

Crossley, T. and S.-H. Jeon (2007). Joint Taxation and the Labour Supply of Married Women: Evidence from the Canadian Tax Reform of 1988. *Fiscal Studies 28*, 343–365.

Díaz-Giménez, J., A. Glover, and J.-V. Ríos-Rull (2011). Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update. *Federal Reserve Bank of Minneapolis Quarterly Review 34*(1).

Domeij, D. and M. Flodén (2006). The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased. *Review of Economic Dynamics 9*, 242–262.

Domeij, D. and P. Klein (2013). Should Day Care Be Subsidized? *Review of Economic Studies 80*(2), 568–595.

Duernecker, G. and B. Herrendorf (2016). On the Allocation of Time - A Quantitative Analysis of the US and France. Working Paper.

Eissa, N. O. (1995). Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment. NBER Working Paper No. 5023.

Eissa, N. O. (1996). Labor Supply and the Economic Recovery Tax Act of 1981. In M. Feldstein and J. Poterba (Eds.), *Empirical Foundations of Household Taxation*. NBER Conference Volume.

Erosa, A., L. Fuster, and G. Kambourov (2012). Labor Supply and Government Programs: A Cross-Country Analysis. *Journal of Monetary Economics 59*, 84–107.

Greenwood, J. and A. Seshadri (2002). The U.S. Demographic Transition. *American Economic Review, Papers and Proceedings 92*(2), 153–159.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005). Engines of Liberation. *Review of Economic Studies 72*, 109–133.

Guner, N., R. Kaygusuz, and G. Ventura (2012a). Taxation and Household Labor Supply. *Review of Economic Studies 79*, 1113–1149.

Guner, N., R. Kaygusuz, and G. Ventura (2012b). Taxing Women: A Macroeconomic Analysis. *Journal of Monetary Economics 59*, 111–128.

Guner, N., R. Kaygusuz, and G. Ventura (2016). Child-Related Transfers, Household Labor Supply and Welfare. Working Paper.

Guvenen, F., B. Kuruscu, and S. Ozkan (2014). Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis. *Review of Economic Studies 81*, 818–850.

Jones, L. E., R. E. Manuelli, and E. R. McGrattan (2014). Why Are Married Women Working So Much? *Journal of Demographic Economics 81*(1), 75–114.

Kaygusuz, R. (2010). Taxes and Female Labor Supply. *Review of Economic Dynamics 13*(4), 725–741.

Keane, M. P. (2011). Labor Supply and Taxes: A Survey. *Journal of Economic Literature 49*, 961–1075.

Knowles, J. (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.

Krueger, D., F. Perri, L. Pistaferri, and G. L. Violante (2010). Cross-Sectional Facts for Macroeconomists. *Review of Economic Dynamics 13*(1), 1–14.

McDaniel, C. (2011). Forces Shaping Hours Worked in the OECD, 1960-2004. *American Economic Journal: Macroeconomics 3*, 27–52.

McDaniel, C. (2012). Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD 1950-2003. Working Paper.

Mulligan, C. B. and Y. Rubinstein (2008). Selection, Investment, and Women's Relative Wages over Time. *Quarterly Journal of Economics 123*(3), 1061–1110.

Ngai, L. R. and B. Petrongolo (2014). Gender Gaps and the Rise of the Service Economy. Working Paper.

Ngai, L. R. and C. A. Pissarides (2011). Taxes, Social Subsidies, and the Allocation of Work Time. *American Economic Journal: Macroeconomics 3*(4), 1–26.

Ohanian, L., A. Raffo, and R. Rogerson (2008). Long-Term Changes in Labor Supply and Taxes: Evidence from OECD Countries, 1956-2004. *Journal of Monetary Economics 55*, 1353–1362.

Olivetti, C. (2006). Changes in Women's Hours of Market Work: The Role of Returns to Experience. *Review of Economic Dynamics 9*, 557–587.

Olivetti, C. and B. Petrongolo (2008). Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps. *Journal of Labor Economics 26*, 621–654.

Prescott, E. C. (2004). Why Do Americans Work So Much More Than Europeans? *Federal Reserve Bank of Minneapolis Quarterly Review 28*, 2–13.

Ragan, K. (2013). Taxes, Transfers and Time Use: Fiscal Policy in a Household Production Model. *American Economic Journal: Macroeconomics 5*(1), 168–192.

Rendall, M. (2015). Female Market Work, Tax Regimes, and the Rise of the Service Sector. Working Paper.

Restuccia, D. and R. Rogerson (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics 11*(4), 707–720.

Rogerson, R. (2006). Understanding Differences in Hours Worked. *Review of Economic Dynamics 9*, 365–409.

Rogerson, R. (2007). Taxation and Market Work: Is Scandinavia an Outlier? *Economic Theory 32*, 59–85.

Rogerson, R. (2008). Structural Transformation and the Deterioration of European Labor Market Outcomes. *Journal of Political Economy 116*, 235–259.

Rogerson, R. (2009). Market Work, Home Work, and Taxes: A Cross-Country Analysis. *Review of International Economics 17*, 588–601.

Wallenius, J. (2013). Social Security and Cross-Country Differences in Hours: A General Equilibrium Analysis. *Journal of Economic Dynamics and Control 37*, 2466–2482.

A Appendix: Applying Tax Codes from Different Countries

The third row (τ_l) of Table 4 shows the model predicted hours if the labor income tax system is set country-specific, while gross household income y_{hh} in Equation (4) remains at the US level, i.e.

$$y_{hh} = w_m^{US}(z)h_m + w_f^{US}(x)h_f. \tag{A.1}$$

Progressive tax systems are in some way defined relative to the income level in a country. For example, the US mean wage ($\bar{w}^{US}$) is around four times higher than the mean wage in Hungary (which has the lowest wage). Simply applying the Hungarian tax system one to one to the US would imply that the average household would end up in a range of the tax code featuring a much higher tax rate than the average Hungarian household. We account for this in the following way. First, for each combination of husband-wife hours choices, we calculate the tax rate in country i using the US gender-specific education premia and the country-specific mean wage ($\bar{w}^i$). Second, we apply this tax rate to the US gross earnings implied by the same husband-wife hours choices to obtain the household's income tax liability τ_l, and set Equation (5) equal to:

$$\tau_l = y_{hh} \frac{\tau_l^i \left(\frac{w_m^{US}(z)}{\bar{w}^{US}} \bar{w}^i h_m, \frac{w_f^{US}(x)}{\bar{w}^{US}} \bar{w}^i h_f \right)}{\frac{w_m^{US}(z)}{\bar{w}^{US}} \bar{w}^i h_m + \frac{w_f^{US}(x)}{\bar{w}^{US}} \bar{w}^i h_f} \tag{A.2}$$

in the household optimization problem.

We proceed in a similar fashion when we analyze the effects of country-specific gender-education premia in row 6 of Table 4. Household income in Equation (4) in this case is replaced by

$$y_{hh} = \frac{w_m^i(z)}{\bar{w}^i} \bar{w}^{US} h_m + \frac{w_f^i(x)}{\bar{w}^i} \bar{w}^{US} h_f, \tag{A.3}$$

and the household's income tax liability in Equation (5) by

$$\tau_l = \tau_l^{US} \left(\frac{w_m^i(z)}{\bar{w}^i} \bar{w}^{US} h_m, \frac{w_f^i(x)}{\bar{w}^i} \bar{w}^{US} h_f \right). \tag{A.4}$$

Thus, the mean wage remains unchanged, but only the gender-education premia are set country-specific.