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

Alexander Bick
Arizona State University

Nicola Fuchs-Schündeln
Goethe University Frankfurt, CFS and CEPR *

February 9, 2016

Abstract

We document contemporaneous differences in the aggregate labor supply of married couples across 17 European countries and the US along the extensive and the intensive margin. 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, as well as the educational composition and gender wage gaps and educational premia, to the international differences in the data. Our model replicates the comparatively small cross-country differences of married men's hours worked well. Moreover, taxes, wages, and the educational composition account for a large part of the observed large differences in married women's labor supply between the US and Europe. The non-linearity of labor income taxes leads to substantially different effects of taxation on married men and women, and explains a significant part of the variation in married women's labor supply within Europe. Consumption taxes on the other hand are the main factor causing the US-Europe difference in married women's hours worked.

Keywords: Taxation, Two-Earner Households, Hours Worked

JEL classification: E60, H20, H31, J12, J22

*Email: alexander.bick@asu.edu and fuchs@wiwi.uni-frankfurt.de.

We thank Domenico Ferraro, Dirk Krueger, José-Víctor Ríos-Rull, Richard Rogerson, Kjetil Storesletten, seminar participants at Arizona State University, Carlos III, CEMFI, Chicago Fed, DIW Berlin, IZA Institute for the Study of Labor, LMU Munich, Minneapolis Fed, Northwestern University, Pompeu Fabra, Princeton, Stanford University, Tilburg University, Tinbergen Institute, Universitat Autònoma de Barcelona, University of Michigan, and University of Nürnberg-Erlangen, and conference participants at the CASEE Quantitative Macroeconomis Conference at Arizona State University for helpful comments and suggestions. Bettina Brüggemann 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, where part of this paper was written, for the hospitality, and the Fritz Thyssen Foundation for the financial support of that stay. All errors are ours.

1 Introduction

The international labor supply of married men and women differs substantially. While married men in the core age group 25 to 54 work on average between 9 and 17 percent fewer hours in Europe than in the US, the picture is substantially more heterogeneous for married women. Married women in Eastern Europe and Scandinavia work only 3 and 9 percent fewer hours, respectively, than US married women, whereas married women in Western and Southern European work 26 and 31 percent fewer hours, respectively.[1] In other words, for Scandinavia and Eastern Europe, hours worked differences relative to the US are half the size or less for married women than for married men, while for Southern and Western Europe they are twice the size or more for women than for men. 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 respective cross-country correlations are 0.75 and 0.70). 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 differences to the US should be similar for both genders, possibly with larger differences for women, which is not consistent with the data.

In this paper, we build a simple model of joint household decision making that incorporates international differences in wages, educational composition, and taxation, and show that it is in fact largely able to replicate the above mentioned international differences in the hours worked of married men and married 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 within-Europe variation in married women's hours worked, and breaking the correlation of married men's and women's labor supply, is the explicit modeling of non-linearities in the labor income tax code together with the tax treatment of married couples, which ranges from mostly joint to mostly separate taxation, with many countries falling in between the two extremes. This tax treatment interacts with the progressivity of the tax system in affecting labor supply decisions of both spouses.

To explain 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 household tax burden is then determined as the sum of the tax burdens on these two hypothetical equal incomes. Due to the progressivity of the German tax code, this lowers the overall tax burden of the household, but increases the marginal tax rate of the secondary income earner, whose contribution to the household income is less than half and who would therefore face a lower marginal tax rate under individual taxation. By contrast, it decreases the marginal tax rate of the primary

[1] The country groups are made up as follows: Scandinavia – Sweden and Norway; Western Europe – Austria, Belgium, France, Germany, Ireland, Netherlands, United Kingdom; Eastern Europe – Czech Republic, Hungary, Poland; Southern Europe – Greece, Italy, Spain. Portugal and Denmark are also included in our sample but we discuss them separately as they differ along two important dimensions from the other countries in their respective region.

income earner. Thus, the treatment of couples in labor income taxes leads to additional non-linearities in the tax code compared to singles, which differ substantially for husband and wife. While the US system works differently, it also features joint taxation, i.e. the tax rate of one spouse is increasing in the income of the other spouse.[2]

The key importance of the tax treatment of married couples can be best illustrated with a concrete example. Take the case 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 *individual* 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 on average slightly larger in Sweden than in Germany, and substantially lower in the US. If the husband works the mean hours of US married men, indicated by a vertical line at 1970 hours, the average tax rate is with 20.5% more than 10 percentage points lower in the US than in Germany and Sweden with 31.0% and 32.8%, respectively. 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. Except for very low levels of hours worked, at which special taxation rules apply, 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 30.1% in the US and Sweden, respectively, but is drastically higher in Germany with 50.3%. Comparing the left and the right tax schedule, one can see that tax rates are very similar for both spouses in Sweden, a country with separate taxation, but substantially higher for wives than for husbands 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, and around 15% fewer hours than in the US, while married women in Sweden work only slightly fewer hours than US married women (4%), but German married women work 34% fewer hours. The model is able to replicate these observations because it features both differences in the average tax rate and in the tax structure, combining the progressivity and the tax treatment of couples.

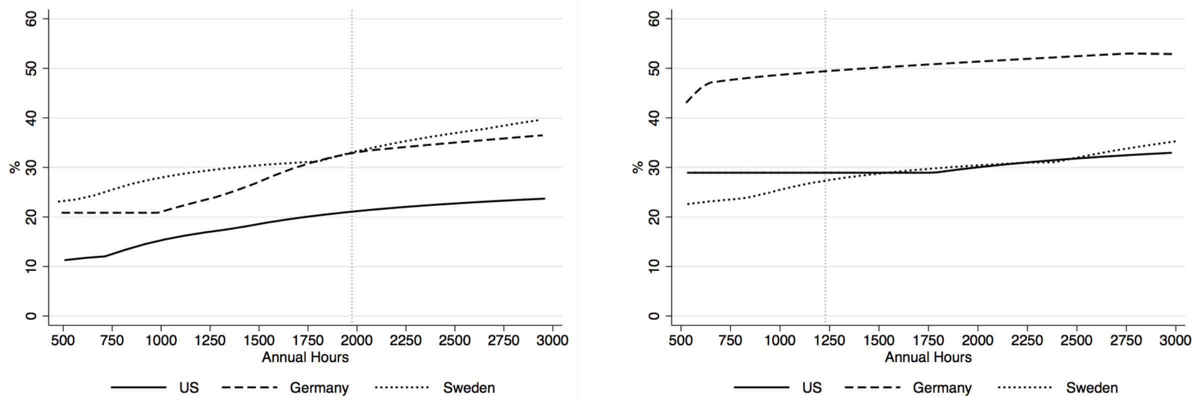
Our quantitative framework is based on the model of joint labor supply of married couples developed in Kaygusuz (2010), Guner et al. (2012a), and Guner et al. (2012b), which features an extensive and an intensive margin of female labor supply. As typical for cross-country studies in macroeconomics, we calibrate the model to match the labor supply behavior in a benchmark country, namely the US. We then predict labor supply behavior in the 17 European countries of the sample, holding preferences fixed but using the country-specific economic environment. The latter comprises non-linear labor income taxes, consumption taxes, wages, specifically the gender wage gap and educational premia, and the educational distribution plus the degree of assortative matching into couples. For the non-linear labor income taxes, we use OECD tax modules which capture country specific features of average and marginal income tax rates of married cou-

[2] In the US system, the incomes of both spouses are summed up, and the income thresholds between which different marginal tax rates occur are multiples of the thresholds for singles. In 2007, e.g.,these thresholds are multiplied by two only for the two lower tax brackets, and by 1.67, 1.21, and 1 for the three subsequent thresholds. Therefore, households with a household income above roughly the 90th percentile face a marriage penalty, which cannot occur in the German system.

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

(a) Average Income Tax Rate for Single Earner Married Households

(b) Average Marginal Income Tax Rate on Wives' Earnings



Vertical dashed line represents mean hours worked per married man (left panel) and married woman (right panel) in the US.

ples in detail, along with standard exemptions, tax credits and benefits, etc., as well as the tax treatment of married couples.

The model correctly predicts lower, but relatively homogeneous hours worked of married men in Europe compared to the US. For married women, the model is able to replicate the small hours worked differences between the US and Eastern Europe and Scandinavia, explains three fourth of the large hours worked difference between the US and Western Europe, and 46 percent of the difference between the US and Southern Europe. Crucial for the results is the fact that the US features a system of joint taxation of married couples, while most Scandinavian, Eastern, and Southern European countries rely on separate taxation, with a mixture of both systems in Western Europe. Higher consumption and average labor income tax rates in Europe mostly lead to *lower* optimal hours relative to the US, an effect which works qualitatively similarly for both married men and married women, but is larger for married women. However, the different tax treatment of married couples and the implied differences in the marginal tax rates of the primary (male) and secondary (female) income earner imply *higher* hours worked of married women in the majority of European countries than in the US, counteracting the effect of the average tax rate to a large degree. This effect tilts married women's hours in Europe in the direction consistent with the data, and is thereby crucial for explaining the within-Europe differences in female labor supply. Differences in the gender wage gap and educational premia across countries play a relatively minor role in explaining cross-country differences in hours worked of married men and women, while the educational composition additionally helps in explaining the within-Europe variation of female hours.

While the model is thus quite successful in explaining hours worked per married woman in our sample countries, it cannot replicate the decomposition into the extensive and the intensive margin well. In contrast to the US and Southern and Eastern Europe, Scandinavia and Western Europe feature relatively high employment rates and low hours worked per employed, which the model does not generate. Through the lens of our model, this is not surprising, since the effects of taxes and wages work qualitatively the same way on both margins. To understand the driving forces of the different margin decompositions better, we introduce wedges into the model, and show that they highly correlate with the part-time generosity ranking of countries provided by the OECD. We show that the incentive effects of our model factors are very similar whether the wedges are introduced or not, i.e. they do not depend on the margin decompositions.

A series of papers (Prescott (2004), Rogerson (2006), Rogerson (2008), Rogerson (2009), Ohanian et al. (2008)) have shown that differences in average tax rates can largely explain differences in the developments of *aggregate* hours worked across European countries and the US, with the exception of Scandinavia. We differentiate 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. Two features that we abstract from in our model 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. **?**, Erosa et al. (2012), and Wallenius (2014) analyze international differences in social security programs, and specifically in retirement systems, and conclude that they can explain large international differences in the timing of retirement, while having almost no effect on labor supply behavior before retirement, i.e. in the age group we focus on.[3]

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 complements to home production (e.g. child care) as the main explanation for this apparent puzzle, a point that was already raised theoretically by Rogerson (2007). They therefore explicitly model home production in addition to market work and distinguish between the taxation of sectors that are substitutes or complements to home production.[4] By contrast, we can replicate Scandinavian hours well, and even better than Ragan (2013), by taking the non-linearity of the labor income tax code into account.[5] Moreover, in the data Scandinavian labor supply is especially high exactly for married women: married women work 9 percent less than US married women in Scandinavia, while single women work 22 percent less, married men 17 percent less, and single men 8 percent less.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. (2012). They build a comprehensive life-

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

[4]Duernecker and Herrendorf (2015) also show the importance of home production in explaining aggregate hours worked, but focus on differential productivity improvements in the home production sector, rather than subsidies to sectors that complement home production.

[5]Our predictions for hours worked in the three Scandinavian countries are off by 7 percent for Sweden, 6 percent for Norway, and 28 percent for Denmark, whereas the corresponding numbers in Ragan (2013) are 39 percent, 41 percent and 50 percent in her benchmark calibration, and 24 percent, 28 percent and 41 percent 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.

cycle model with idiosyncratic income risk to investigate the cross-country variation in hours worked for married and single men and women. Besides taxes, they concentrate on exogenous marriage and divorce probabilities as driving forces of labor supply differences, but do not model international differences in divorce legislation and alimony regulations. Chakraborty et al. (2012) estimate polynomial tax functions for married couples and thus allow for some non-linearity, but base their estimates on the sum of household earnings, and thus cannot exploit the differential effects of tax progressivity vs. tax levels on husbands and wives under systems of joint vs. separate taxation. Their model fit is worse than ours for men and women with the exception of married women in Portugal and Denmark. This again stresses the importance of modeling the tax system in all its details.

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, attributing it e.g. to technological improvement in the household sector (Greenwood and Seshadri (2002), Greenwood et al. (2005)), changes in the gender wage gap (e.g. Jones et al. (2003) and Albanesi and Olivetti (2009)), which also led to changes in the bargaining power within the household (Knowles (2013)), the decline in child care cost (Attanasio et al. (2008)), changes in the return to experience for women (Olivetti (2006)), or improvements in maternal health and the introduction of infant formula (Albanesi and Olivetti (2009)). Within this literature, Guner et al. (2012a) and Guner et al. (2012b) focus explicitly on the taxation of married women. 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. 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) and Eissa (1996) analyze 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 well-earning husbands significantly. Kaygusuz (2010) evaluates the effects of the same US tax reforms on the labor supply of married women with a quantitative model. Rendall (2015) similarly 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.

The paper is organized as follows. The next section presents the micro data sources, explains the construction of the relevant data series, and presents our sample selection criteria. Section 3 shows some facts on the labor supply of married couples. The following section introduces the model, as well as its parametrization and calibration. Section 5 presents the results, discusses the wedges needed to match extensive and intensive margin decompositions, and investigates the relative role of the various model inputs, specifically of the non-linear labor income tax schedule. The last section concludes.

2 Micro Data

2.1 Data Sets on Hours Worked

We work with three different micro data sets to construct hours worked, namely the European Labor Force Survey, the Current Population Survey, and the German Microcensus. A detailed description of the data work, as well as a comparison of the resulting aggregate data to similar series from the OECD and the Conference Board, can be found in Bick et al. (2015). There, we argue that our data are likely more suitable for international comparisons than OECD and Conference Board data.

2.1.1 European Labor Force Survey

The European Labor Force Survey (ELFS) is a collection of annual labor force surveys from different European countries, with the explicit goal of making them comparable across countries. Our ELFS sample comprises the Scandinavian countries Denmark, Norway, and Sweden, the Eastern European countries Czech Republic, Hungary, and Poland, the Southern European countries Greece, Italy, Portugal, and Spain, and the Western European countries Austria, Belgium, France, Ireland, Netherlands, and the UK. The sample size of the ELFS varies across countries and within a country over time, but is always of considerable magnitude, with the minimum annual sample size being more than 15,000 for Denmark, a country with roughly 5.5 million inhabitants. The weeks used as reference week in the survey vary from country to country and year to year, mostly covering a period of between 1 and 12 weeks in the first half of the year up to the year 2004, and the entire year from 2005 on. Appendix A.1.1 describes some data modifications that we have to apply to specific years and countries of the ELFS.

2.1.2 Current Population Survey

For the US, we use the Current Population Survey (CPS), which is a monthly survey of around 60,000 households. Specifically, we work with the CPS Merged Outgoing Rotation Groups data provided by the National Bureau of Economic Research. This data set includes only those interviews in which the households are asked about actual and usual hours worked, namely the fourth and eighth interview of each household. The data cover the entire year, with the reference week always including the 12th of a month, and comprise individual data for about 300,000 individuals per year.

2.1.3 German Microcensus

The German Microcensus covers a one percent random sample of the population of Germany and is an administrative survey. Participation is mandatory. We use the scientific use files, which are a 70 percent random subsample of the original sample. This leaves us with a sample size of between 400,000 and 500,000 individuals per year. Until 2004, the Microcensus was carried out in the last week without a public holiday in April or the first week without a public holiday in May, and from 2005 on continuously over the year.

2.2 Calculation of Average Hours Worked per Person

For each individual, we have information on four key variables: actual hours worked in the main job during a specific reference week, actual hours worked in additional jobs during the reference week, usual hours worked in the main job during a working week, and reasons why the individual worked more or less hours than usual in the reference week.

The main challenge in generating average annual hours worked per person lies in the fact that the reference weeks are not spread representatively across the entire year. In Bick et al. (2015), we show that reference weeks mostly exclude typical vacation periods, and report evidence that vacation days and public holidays are underreported even during the years in which reference weeks cover the entire year. Therefore, we collect information on the number of vacation days and public holidays by country and year from external data sources. The main disadvantage from using external data sources is that we cannot account for heterogeneity in the population when it comes to vacation days.

To generate annual hours worked per person, we first construct individual weekly hours worked by adding up actual hours worked in the reference week in all jobs. For individuals who report having worked less hours than usual in the reference week due to vacation or public holidays, we use usual hours worked instead of actual hours worked. We then multiply these weekly hours worked by 52 minus the weeks lost due to vacation days and public holidays, i.e. the number of these days divided by five, in the respective country, and then average over all individuals.

2.3 Sample Selection

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. This is for example the case in the Netherlands, where individuals living in a civil union are recorded as married in the ELFS. Next, we include only couples for which both partners are observed and fit our sample restrictions. Since clear identifiers for husbands and wives are missing for many years and countries, we define couples consistently as two people of opposite sex who are both married and live in the same household, and drop households in which more than two married adults live. 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 education systems, degrees of youth unemployment, and early retirement programs. Wallenius (2014) concludes that international differences in social security programs are an important driving force in the timing of retirement, but have almost no effect on labor supply behavior in the core age group.

We concentrate on the sample period 2001 to 2008. We use a sample period of more than one year 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 of the sample period is determined by the availability of the OECD tax modules. Last, since we model heterogeneity through differences in education levels, we exclude individuals with missing information on own education or partner's education.

There are three reasons why a married individual aged 25 to 54 might be dropped from our sample, namely first because we cannot identify the partner due to more than two married adults or no other married adult living in the household, second because the partner might be younger than 25 or older than 54, and third because education information might be missing for the respondent or the partner. On average, around 6% of male and 10% of female observations are dropped because of these restrictions (see Table A.6 in the online appendix). The percentage is always larger for women than for men, because it is more likely for women that the partner is older than 54. Variation across countries arises because of variation in the number of missing education observations, variation in the age structure in marriage and age at marriage, and variation in the number of married adults living in one household.[6]

3 Hours Worked of Married Couples

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. On average, married men aged 25 to 54 work around 730 hours more than married women in the same age group. Single women work 190 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: in fact, the standard deviation of hours worked of married women is more than 70 percent higher than 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 an order of magnitude bigger. Married women contribute on average 23 percent of total hours worked, but account for 29 percent of the variance of total hours worked. Moreover, the cross-country correlation of hours worked of married men with the one of single men or single women is 0.75 and 0.70, respectively, while the correlation with hours worked of married women amounts only to 0.05. 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.

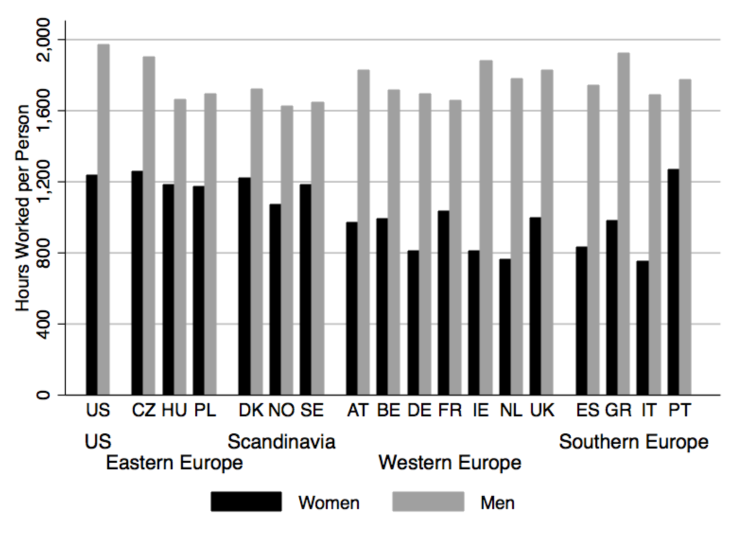
Since from now on we focus on married couples, the issue of selection into marriage arises. While we do not model this selection, we report in Figure A.4 in the Online Appendix the fraction of women in the core age group who are married. It amounts on average to 64 percent, with a standard deviation of 7.5 percent. The extremes are Sweden with 48 percent of women being married, and Poland with 77 percent. For the majority of countries, the fraction of married women lies between 60 and 70 percent. 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,

[6] E.g. Poland, which has the largest number of observations dropped, exhibits an unusually large percentage of individuals married to someone younger than 25, as well as an unusually large number of households consisting of three or more married adults. In the data for the Scandinavian countries, it is in most years impossible to identify households. Therefore, the only reason why married individuals might be dropped is missing information on their own education, leading to a small fraction of observations dropped. Appendix A.1.2 explains how we deal with missing household identifiers in Scandinavian data.

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

	Men		Women	
Country	Married	Single	Married	Single
Mean	1761.6	1484.5	1028.2	1217.0
Standard Deviation	104.2	108.8	179.5	96.6
Coefficient of Variation	0.059	0.073	0.175	0.079
Var(log hours)	0.003	0.005	0.033	0.006
Correlation with HWP_m^{Mar}	1.00	0.75	0.05	0.70

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

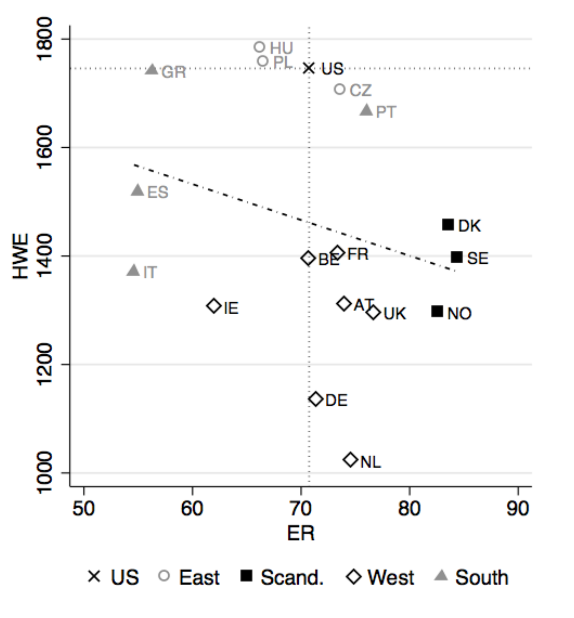


the tax treatment of couples, i.e. whether a country employs a system of joint or separate taxation, is not correlated with the marriage rate.[7] Last, Chade and Ventura (2002) and Chade and Ventura (2005) show in a quantitative equilibrium model of the marriage market for the US that the marriage rate would barely change if the US would replace the current system of joint taxation with one of separate taxation.

Figure 2 shows average hours worked of married women (dark bars) and men (light bars) aged 25 to 54 over the period 2001 to 2008 for all eighteen countries in our sample in a bar chart. The European countries are grouped into four regions, namely Eastern Europe (Czech Republic, Hungary, Poland), Scandinavia (Denmark, Norway, Sweden), Western Europe (Austria, Belgium, Germany, France, Ireland, Netherlands, United Kingdom), and Southern Europe (Spain, Greece, Italy, Portugal). The ordering of the regions is according to the mean hours worked per person of married women, and the countries within each region are

[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 columns 2 and 3 of Table 2 is 0.08, and the correlation between the marriage rate and the size of the tax structure effect shown in Table 7 is 0.13.

Figure 3: Average Employment Rates (ER) and Hours Worked per Employed (HWE) of Married Women (Ages 25-54)



ordered alphabetically.

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, Hungary, and France. Norwegian married men work 350 hours less than, or only 82 percent of, US married men. There is no clear pattern in terms of married men's hours worked among Western, Southern, Eastern, and Northern European countries.

By contrast, for married women there is a clear regional pattern of hours worked per person, which are high in the US, Eastern Europe, and Scandinavia with 1240, 1200 and 1160 hours, respectively, and much lower in Western and Southern Europe with 910 and 960 hours. Portugal is a notable exception, which despite being a Southern European country actually features the highest hours worked of married women.[8] 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, 490 hours less than, or only 61 percent of, US married women. The graph reflects the finding of Table 1 that the differences in hours worked of married women are much larger than for married men.

Figure 3 decomposes hours worked per married woman 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. In this and all following scatter plots, the US is marked with an x, Eastern European countries with a circle, Scandinavian countries with a square, Western European countries with a diamond, and Southern European

[8]Without Portugal, Southern Europe features lower hours of married women than Western Europe, namely 850.

countries with a triangle. While Figure 2 showed that Scandinavia and Eastern Europe both exhibit high hours worked of married women, Figure 3 documents very different behavior in the decomposition into an intensive and an extensive margin: 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 and Western Europe both exhibit low hours worked per married woman, but again are clearly distinguishable in the decomposition into an intensive and an extensive 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 which are lower than in the US, but higher than in Western Europe. The resulting cross-country correlation of hours worked per employed and employment rates of married women is negative with a value of -0.37. Figure A.5 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 around 90 percent. 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.

One might be worried that part of the international differences in married women's labor supply across countries comes from differential effects of children on mothers' hours worked due to different child care availability and cost, maternity leave policies, cultural factors, etc. As we show and describe in more detail in Section A.1 in the online appendix, the effect of having a school aged child on employment rates and hours worked per employed is very similar across countries.[9]The effect of preschool children (aged 0-4) on hours worked per employed is slightly larger for married women in Western Europe than in other countries. Most importantly, having a child of preschool age reduces employment rates for married women in Hungary and the Czech Republic substantially more than in other countries, but has on average a somewhat effect in Southern Europe than in other countries. While the behavior of women with preschool children thus shows country-specific idiosyncracies, likely caused by differences in child care as well as in maternity leave policies, the share of such women is relatively small, never exceeding 30% except for Ireland. Thus, the group is not large enough to change the aggregate picture. In fact, focusing on women without preschool children strengthens the patterns we see in the data: it increases the hours worked of married women in Eastern Europe compared to the US by 4 percentage points, and decreases the ones in Southern Europe by 5 percentage points. The difference in hours worked per married woman between Western Europe and the US is virtually unchanged if women with preschool children are excluded.[10]

Similarly, life-cycle or cohort effects might cause part of the observed differences in the data. Given the short sample period, we cannot distinguish between both types of effects. In Section A.2 in the Online Appendix, we document labor supply of married women by 10 year age groups. After controlling for common age effects, hours worked per employed differences across the age groups within each country

[9]We have to exclude Scandinavian countries from these analyses, due to missing information on households (see Appendix A.1.2).

[10]In Section A.5 in the online appendix, we also show results from our model focusing exclusively on women without children, and the overall model fit is very similar.

are very small. Regarding the employment rate, European married women aged 45-55 show uniformly around 5 percentage point larger differences to US married women of the same age than European women aged 35-45. If one can attribute this to a cohort effect, one might expect somewhat smaller Europe-US differences in married women's hours worked in the future. For women aged 25-34, only exceptionally low employment rates in Hungary and the Czech Republic stand out, likely caused by the effect of preschool children documented above. Thus, we conclude that life-cycle or cohort effects might explain a small part of the Europe-US difference, but are not a major driver of the variation in hours worked within Europe.

Last, we assume in our analysis that differences in hours are driven by the supply side, and not by the demand side. However, international differences in unemployment rates could indicate different probabilities of finding a job. To see how large this effect could potentially be, we take the extreme view that all unemployment differences are driven by the labor demand side. We therefore exclude all unemployed individuals (and partners of unemployed individuals) from the sample, and recompute our labor supply measures. Relative to the US, male and female hours worked per person increase in Eastern Europe and Southern Europe. However, the changes are not dramatic: hours worked per married woman differences to the US increase from -3% to +3% in Eastern Europe, and from -23% to -19% in Southern Europe (see Table A.5 in the online appendix).[11]

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, the educational composition, and education-gender-specific wage premia contribute to the cross-country differences in male and female labor supply presented in Figures 2 and 3. The model framework is based on Kaygusuz (2010) and features a maximization problem of a two person household which jointly determines male and female labor supply.[12]

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 $w^i_f(x)$ and $w^i_m(z)$, where the superscript i represents a given country. We denote the fraction of households of type x, z by $\mu^i(x,z)$ with

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

Households draw a utility cost of joint work q from a distribution $\zeta(q|z)$ which depends on the husband's

[11] These results are in line with unemployment rates obtained from the World Development Indicators. Some countries, especially Poland, Greece, Italy, Spain, France, and Belgium, suffer from very high youth unemployment rates during the sample period, but unemployment rate differences for the core age group are less dramatic.

[12] Guner et al. (2012a) and Guner et al. (2012b) embed the same preference structure in a life-cycle setting to evaluate different tax policies within the US.

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 an explicit extensive margin choice for men since their participation rates in our sample are above 90 percent 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 costs of working are so high that the woman chooses not to work, i.e. $h_f = 0$.

The household faces two types of taxes, namely a consumption tax rate τ_c and a non-linear labor income tax τ_l, which depends on both spouses' gross incomes, as well as the number of children in the household k. The maximization problem of a type $\{x,z\}$ household in country i is given by

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

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

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

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

where $\mathbf{I}_{h_f>0}$ takes the value one if the wife is working and zero otherwise, c represents household consumption, y_{hh} represents gross household income, and τ_l the household's income tax liability, which depends on the gross incomes of husband and wife, as well as the number of children in the household k^i through tax credits and/or child benefits. T represents a lump-sum transfer from the government which redistributes a share $\lambda^i \in [0,1]$ of all government revenues:

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

where * denotes the optimal hours choice given the draw of q.[13] As e.g. in Prescott (2004) and Rogerson (2008), households do not internalize that their choices affect the transfer, but their expectation about the transfer is consistent with the realization.

[13]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. $R = \tau_l + \tau_c^i(c - T)$. T is subtracted in this calculation since the transfer is not subject to the consumption tax. Replacing c from the budget constraint ($c = \frac{1}{1+\tau_c^i}(y_{hh} - \tau_l) + T$), yields $R = \frac{1}{1+\tau_c^i}\left[\tau_l + \tau_c^i y_{hh}\right]$. The transfer T is then the fraction λ^i from government revenues R.

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 differences in the gender-specific education premia matter for labor supply decisions. 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.

4.2 Model Inputs

As inputs into the model, we need country-specific information on consumption tax rates τ_c^i, non-linear labor income taxes τ_l^i, the educational composition and matching into couples $\mu^i(x,z)$, male hourly wages by education $w_m^i(z)$, and female hourly wages by education $w_f^i(x)$. Last, we calibrate the two preference parameters in the utility function, α and ϕ. When used in the model, wages and taxes are converted into 2005 US-Dollars, using PPP-adjusted exchange rates obtained from the Penn World Tables.

4.2.1 Consumption Taxes

Consumption tax rates for our sample countries are provided by McDaniel (2012), who calculates consumption tax rates from NIPA data. 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 column 1 of Table 2. Differences in consumption tax rates are large between the US and Europe, amounting to more than 10 percentage points, and consumption tax rates are highest in Scandinavia, where they are below 30 percent only in Norway.

4.2.2 Non-Linear Labor Income Taxes

The non-linear labor income tax systems are taken from the "OECD Taxing Wages" modules. These are very similar to the NBER TaxSim module for the US, but cover all OECD countries. The OECD provides annual household net income based on the respective country's and year's tax laws, taking income taxes plus employees' social security contributions, as well as cash benefits and standard deductions into account.[14] Tax modules are available online from the year 2001 onwards and capture the full non-linearities of the tax codes. Using these modules, we can assign an annual net household income to each combination of male and female annual gross earnings. We calculate the exact values for an earnings grid with 101 steps for men, ranging from 0 earnings to four times the average annual earnings in the country, and for an earnings grid with 201 steps for women, ranging from 0 earnings to three times the average annual earnings in the country.[15] We then linearly interpolate in two dimensions to assign a net annual household income to each

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

[15] 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

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 their tax burdens for any pair of hours choices. Figure 1 gives an exemplary impression of these tax codes for the US, Germany, and Sweden, each time holding the earnings of the spouse fixed at one specific level.

Columns 2 and 3 of Table 2 give an overview of the model inputs relating to labor income taxation. Clearly, it is impossible to summarize the complex non-linear labor income tax systems in a few numbers. We want to stress that we are exploiting the full non-linearity of the tax code in our exercise, and here just present some suggestive numbers of the tax code. Columns 2 and 3 show two possible measures that reflect three aspects of the labor income tax schedule: column 2 ($\tau_l(0)$) shows 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 male wage and that the wife does not work, and thus gives one of many possible measures of an average tax rate. Column 3 ($\tau_l'(h_f^{US})$) shows 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 female wage, thereby capturing one possible measure of progressivity, as well as jointness of taxation.[16] 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. Since female US hours are smaller than male US hours, and mean wages are always smaller for women than for men, one would expect the tax rates in the third column to be smaller than the ones in the second column in countries with a progressive tax code and separate taxation of married couples. A substantially higher value in the third column than in the second one by contrast indicates that a country's tax code features strong elements of joint taxation of married couples. These are of course only indicative values, since we only show two specific tax rates here, rather than exploiting the full non-linearity, as we do in our quantitative analysis.

The US average tax rate as calculated in column 2 amounts to 20.5%, whereas the corresponding Danish married couple would have to pay an average tax rate of 38.9%, and the Irish couple a tax rate of only 14.9%. The average tax rates are lowest in the US and Southern Europe, followed by Eastern and Western Europe, and highest in Scandinavia. The measure of the average (marginal) tax rate of the secondary income earner shown in column 3 amounts to 29.1% in the US, peaking at 50.3% in Germany, a country with high progressivity and joint taxation of married couples, and 50.0% in Belgium. This measure is again on average lowest in Southern Europe, followed by Eastern Europe and the US, and shows significantly higher levels in Western Europe and Scandinavia. In Scandinavia, Denmark stands out with a high average (marginal) tax rate of the secondary income earner of 48.7%, while Norway and Sweden have levels similar to the

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.

[16] We define this average (marginal) tax rate as $\tau_l'(h_f^{US}) = [\tau_l(w_m^i h_m^{US}, w_f^i h_f^{US}) - \tau_l(w_m^i h_m^{US}, 0)]/[w_f^i h_f^{US}]$. All tax rates in this table are calculated for couples without children. Children decrease $\tau_l(0)$ via tax credits etc., but hardly affect $\tau_l'(h_f^{US})$.

Table 2: Model Inputs

Country	τ_c	$\tau_l(0)$	$\tau_l'(h_f^{US})$	μ_f^{low}	μ_f^{high}	$\frac{w_f}{w_m}$	$\frac{w_f^{high}}{w_f^{low}}$
Czech Republic	14.8	21.8	23.1	9.2	12.1	0.76	1.8
Hungary	23.6	29.5	18.2	20.8	18.9	0.82	2.0
Poland	18.6	29.3	32.2	10.8	18.9	0.80	3.8
Mean	*19.0*	*26.9*	*24.5*	*13.6*	*16.7*	*0.79*	*2.5*
Denmark	32.0	39.9	49.2	19.3	35.6	0.81	1.3
Norway	24.3	29.7	30.2	13.9	38.3	0.78	1.2
Sweden	32.5	33.5	28.0	13.9	37.1	0.76	1.2
Mean	*29.6*	*34.4*	*35.8*	*15.7*	*37.0*	*0.78*	*1.2*
Austria	18.7	31.2	22.8	23.8	14.7	0.76	1.7
Belgium	20.7	34.5	48.2	28.7	34.7	0.84	1.6
France	23.8	23.9	32.8	29.7	28.0	0.79	1.7
Ireland	24.0	16.9	16.9	27.7	29.7	0.70	2.6
Germany	15.4	33.4	49.6	16.9	20.4	0.73	1.5
Netherlands	21.3	31.0	34.7	28.9	24.5	0.75	1.5
United Kingdom	17.1	26.5	19.2	29.4	33.0	0.74	2.1
Mean	*20.1*	*28.2*	*32.0*	*26.4*	*26.4*	*0.76*	*1.8*
Spain	15.9	17.3	18.5	50.8	27.8	0.68	2.6
Greece	14.9	25.9	16.0	34.4	21.4	0.70	2.5
Italy	22.1	27.8	28.3	45.8	12.1	0.87	1.6
Portugal	19.0	18.4	25.4	72.6	13.8	0.83	2.7
Mean	*18.0*	*22.4*	*22.1*	*50.9*	*18.8*	*0.77*	*2.3*
United States	7.4	21.4	29.1	7.8	45.7	0.79	2.2

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 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 female wage, i.e. $[\tau_l(w_m^i h_m^{US}, w_f^i h_f^{US}) - \tau_l(w_m^i h_m^{US}, 0)]/[w_f^i h_f^{US}]$. μ_f^{low} is the share of low educated women, and μ_f^{high} the share of high educated women. w_f/w_m is the average gender wage gap. w_f^{high}/w_f^{low} is the female education premium (i.e. the wage of high educated women divided by the wage of low educated women).

US. On average, according to this measure, Eastern Europe, Scandinavia with the exception of Denmark, and Southern Europe seem to feature systems of separate taxation of married couples, while the US and Denmark exhibit joint taxation. For Western Europe, the picture is mixed, with separate taxation in Austria and the UK, and elements of joint taxation of various degrees in the rest of the countries, most strongly in Germany and Belgium. For each European country group, the difference in either of the two labor income tax rates to the US is smaller than the difference in consumption tax rates in column 1.

4.2.3 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.[17] The percentages of women with low and high education, omitting the group of medium education, are shown in columns 4 and 5 of Table 2. There are substantial differences in the educational composition: in Southern Europe, on average more than half of the married women exhibit low education, while in Eastern Europe and the US only around 10 percent do. Higher education rates are largest in the US with around 45 percent, followed by Scandinavia and Western Europe, and smallest in Eastern and Southern Europe with between 15 and 20 percent. In Eastern Europe, around two third of married women have medium education. Online appendix Table A.7 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.

4.2.4 Hourly Wages

To calculate hourly wages, we have to divide earnings by hours. Unfortunately, the ELFS does not provide earnings data, and the German Microcensus only net earnings. Therefore, we recur to the EU Statistics of Income and Living Conditions (EU-SILC), which is a European household data set that captures income and usual hours but features a sample size two orders of magnitude smaller than the ELFS. We then calculate country- and year-specific mean wages for married men aged 25 to 54 in the EU-SILC and the CPS for the three different education groups. For comparability reasons, we cap hours and earnings in the EU-SILC as in the CPS, and then construct hourly wages by dividing gross annual individual earnings by annual hours. Online appendix A.3 gives further information on the EU-SILC data and its comparability to ELFS data. We construct annual hours by multiplying usual weekly hours with 52 minus vacation/public holiday weeks from external data sources. Last, we drop observations with wages less than half the minimum wage (as in the Review of Economic Dynamics 2010 special issue on cross-country heterogeneity facts, see Krueger et al. (2010) for details), and the top 1 percent of observations, which are mostly driven by low hours rather than high earnings and seem to be due to measurement error. The EU-SILC starts only in 2004, and for some sample countries even later; we extrapolate wages for the missing years based on Eurostat growth rates of mean wages.

[17]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 6). 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 appendix A.1.2 for details how we proceed for Scandinavia.

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 wages overestimate the distribution 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 influence directly the wage of a woman.

Columns 6 and 7 of Table 2 show 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.[18] The average gender wage gap is quite similar across European regions, though with some differences across countries. Southern Europe exhibits the largest degree of heterogeneity, with the largest gender wage gap in Spain and the smallest one in Italy. The educational premia tend to be higher in Eastern and Southern Europe, as well as the US, than in Western Europe, and are lowest in Scandinavia.

4.3 Redistribution of Government Revenues

The government redistributes a fraction $\lambda^i \in [0,1]$ 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 $\lambda^i = 1$. In Online Appendix Section A.5, we show results from two alternative specifications with either no redistribution of government revenues (i.e setting $\lambda^i = 0$), or from setting λ^i equal to 1 minus twice the share of expenditures on military from all government expenditures, similar to the specification used by Prescott (2004).

4.4 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.

Again following Kaygusuz (2010), Guner et al. (2012a) and Guner et al. (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:

$$q60\zeta(q|z) \equiv q^{k_z-1}\frac{\exp(-q/\theta_z)}{\Gamma(k_z)\theta_z^{k_z}}, \tag{7}$$

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' three own education levels $x \in \{low, medium, high\}$.

[18]Corresponding education premia for men are similar and are shown in Table A.7 in the online appendix.

[19]Robustness checks with respect to this parameter are shown in Online Appendix Section A.5.

Table 3: Data Targets and Calibrated Preference Parameters

	Parameters	Data	Model	$\Delta_{\text{Model-Data}}$
Hours Worked:	$\alpha = 0.463$			
HWP_m		1970	2019	49
HWE_f		1746	1703	-42
Female Employment Rates by Husband's and Own education (in %)				
Low educ. husband:	$k_{low} = 1.051$, $\theta_{low} = 0.267$			
Low educ. woman		42.4	44.9	2.5
Medium educ. woman		63.3	59.9	-3.3
High educ. woman		76.1	77.3	1.2
Medium educ. husband:	$k_{med} = 1.079$, $\theta_{med} = 0.157$			
Low educ. woman		48.4	50.8	2.4
Medium educ. woman		71.2	68.2	-3.0
High educ. woman		83.2	85.7	2.4
High educ. husband:	$k_{high} = 0.521$, $\theta_{high} = 0.335$			
Low educ. woman		49.6	51.8	2.2
Medium educ. woman		66.2	62.7	-3.5
High educ. woman		73.7	75.1	1.3

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 participation rates of the women's labor force participation by their own education conditional on the husband's education.

Table 3 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 2.5% lower in the model than in the data, and hours worked per married man are 2.4% higher than in the data. On average by husband's education, employment rates are matched almost perfectly. Employment rates of low and high educated women are however slightly higher in the model than in the data, and of medium educated women slightly lower.

4.5 Model Validation

Regarding the untargeted gradients by education of hours worked per employed for women, and hours per person for men, the model replicates the gradients qualitatively and also matches the hours of high and medium educated women and men quantitatively well (see Table A.8 in the Online Appendix). However, it underestimates hours of low educated women by 8%. For men, by contrast, the model overestimates the low hours of low educated men by 19%. This is likely due to high non-employment rates among low educated men suppressing their hours worked, which the model does not capture.

We do two further validity checks of the model. First, we analyze 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 parameter fixed. The model correctly predicts hardly any change in hours worked of married men over the period of three decades. This exercise is in the spirit of Kaygusuz (2010), who analyzes only the time period from 1980 to 1990, which we extend to 2008. 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 A.4. Thus, this simple model can capture the time-series development of hours in the US very well.

Secondly, we analyze what the model would predict if the US were to go from the current system of joint taxation of married couples to one of strictly separate taxation. A similar exercise is carried out by Guner et al. (2012a) in a much richer general equilibrium life-cycle model that features marriage and fertility, as well as a social security system. By comparing our result to theirs, we can see whether our simple model is able to capture the main disincentive effects of joint taxation. In this exercise, we compare the labor supply of two singles living together in one household with the one of a married couple living in one household. We add a proportional tax/subsidy when going from the current system to a system of strictly separate taxation, such that tax revenues are kept constant in both cases. Our results are very similar to the results in Guner et al. (2012a). While Guner et al. (2012a) find that the employment rate of married women would increase by 10.4 percent, we find an increase of 9.2 percent, and their predicted increase in hours worked per employed married woman of 0.3 percent compares to ours of 0.14 percent.

5 Results

Keeping the preference parameters fixed across countries, we use country-specific taxes, the educational composition (i.e. the educational distribution by gender and the degree of assortative matching), and wages in order to obtain predicted hours worked of married couples across countries. We first present the cross-sectional predictions of hours worked per person of married men and women, before we decompose hours worked of married women further into an extensive and an intensive margin. Using the US as the benchmark country, we always compare deviations from US hours in model and data. In a decomposition analysis, we evaluate the relative importance of taxes, the educational composition, and wages in explaining the cross-country variations of married individuals' hours worked. We further analyze the effects of labor income taxes by decomposing them into differences in tax levels and differences in marginal tax rate schedules. We discuss Denmark and Portugal separately in the last subsection: as shown before, Denmark is an important outlier within Scandinavia in terms of the taxes, while Portugal is an outlier within Southern Europe in terms of actual hours worked of married women. Thus, all statistics in this Section exclude Denmark and Portugal in both model and data.

5.1 Hours Worked per Person in Model and Data

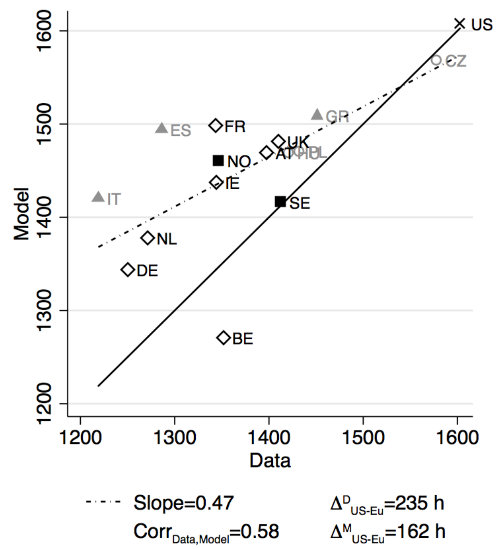
Table 4 shows in the first column the percent difference in married couples' hours worked per person (i.e. adding up hours of husband and wife) between the respective country group and the US in the data, and in the second column the model predicted percent differences. While the table only presents country-group averages, country-specific results presented in level of hours can be seen in Panel (a) of Figure 4. This figure also adds information on the average US-Europe hours difference in the data and the model, the correlation between hours in the data and the model, and the slope of a regression line of model predicted hours on data hours.[20] The model correctly predicts uniformly lower hours in Europe than in the US. Moreover, in both model and data the Europe-US difference is smallest for Eastern Europe, followed by Scandinavia, and then followed by Western Europe. However, while in the data the differences are largest for Southern Europe, that is not the case in the model; we will later see that this is driven by women, and specifically women in Italy. While through the lens of the model, differences in taxes, wages, and the educational composition can account for on average three fourth of the differences between Europe and the US in Eastern Europe, Scandinavia, and Western Europe, they can only account for one half of the difference for Southern Europe. The correlation between model and data is 0.58. To get an impression of the goodness of fit, we replicate the exercise in Prescott (2004) on the aggregate level including all individuals aged 15 to 64 for the years 2001 to 2008, and using linear labor income taxes. In this case, the model explains on average only one half of the aggregate Europe-US hours difference, while we explain slightly more than two thirds. Moreover, the correlation between model and data in the Prescott (2004) exercise is with 0.28 less than half the correlation in our model. This underlines the good fit of our model.

The next columns of Table 4 decompose the previous results by gender. Country-specific results can be seen in panels (b) and (c) of Figure 4. Differences in taxes, wages, and the educational composition explain the cross-country variation in married men's hours worked quite well. Again, on average the model somewhat underpredicts the differences to the US. While married men in Eastern Europe and Scandinavia work 11 and 17 percent less, respectively, than married men in the US in the data, the model predicts on average a difference of minus 6 and minus 11 percent, respectively. For Western Europe, the fit is best with a

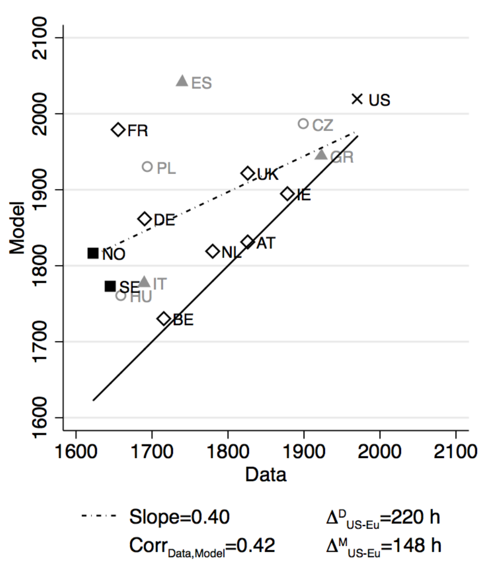
[20] In the regression and the calculation of the correlation, we exclude the US. Thereby, the correlation is the same whether applied to the level of European hours as presented in Figure 4, or the deviation from US hours as presented in Table 4 .

Figure 4: Country-Specific Results

(a) HWP: Men + Women



(b) Men



(c) Women

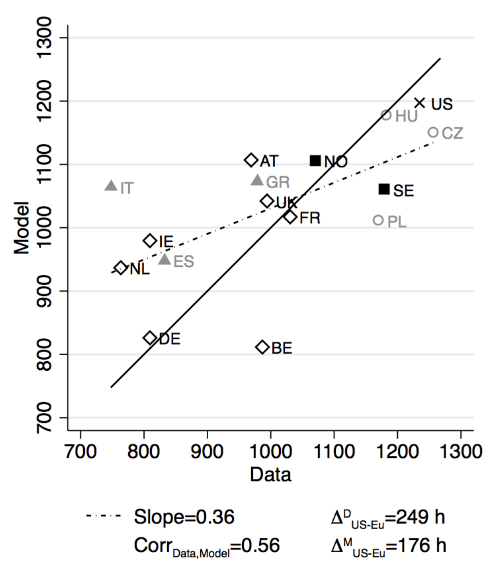


Table 4: Hours Worked per Married Person – Percent Differences Relative to the US

Country	Total		Men		Women	
	Data	Model	Data	Model	Data	Model
Eastern Europe	−7.8	−6.5	−11.1	−6.3	−2.6	−7.0
Scandinavia	−13.9	−10.5	−17.1	−11.1	−8.9	−9.5
Western Europe	−16.5	−12.2	−10.3	−7.8	−26.3	−19.8
Southern Europe	−17.7	−8.4	−9.4	−4.9	−30.9	−14.3

model predicted difference of 8 percent, compared to 10 percent in the data. For Southern Europe, the model explain half of the difference to the US. However, this is almost entirely driven by the bad fit for Spain, which exhibits 12 percent lower male hours worked than the US, while the model predicts 1 percent higher hours than in the US. Focusing on individual countries, for 9 countries the differences between prediction and data amount to 3 percentage points or less. The worst fit is achieved for France, which features 16 percent lower hours of married men than the US, while the model predicts a difference of only 2 percent. Overall, the correlation between male hours worked per person in data and model amounts to 0.42.

We now turn to hours worked of married women in the last two columns of Table 4. Eastern European married women work almost as many hours as US married women with a difference of on average minus 3 percent in the data, while the model predicts a small, but somewhat larger in the absolute sense, difference of minus 7 percent. For Scandinavia, the model generates a similarly good average fit with a difference to the US of minus 9 percent in the data, compared to minus 10 percent in the model. Thus, the model is able to replicate the small differences to the US in married women's hours worked for both Eastern Europe and Scandinavia. For Hungary, Norway, and Sweden, the model correctly replicates larger differences in hours worked relative to the US for men than for women.

For Western Europe the model explains three fourth of the observed large difference in married women's hours worked of minus 26 percent in the data. The fit is very good for France and Germany, with a deviation between model and data of only 1.5 and 3.4 percentage points, respectively. This excellent fit is quite remarkable, given that married women work 17 percent (200 hours) and 34 percent (425 hours) less, respectively, in these two countries than in the US. Only for Austria, where married women work 22 percent less than in the US, can the model explain less than half of the difference. For Belgium, the model overshoots in predicting the difference.

For Southern Europe, the model explains on average 46 percent of the difference in hours worked relative to the US, predicting 14 percent lower hours than in the US, while in the data Southern European married women work 31 percent less than their US counterparts. This is largely driven by Italy, for which only a third of the difference to the US can be explained by the model. For Greece, the model generates half of the observed difference, and for Spain even two third. Austria and Italy are the only two countries in Western

Table 5: Female Employment Rate (ER) and Female Hours Worked per Employed (HWE) Differences Relative to the US

	ER		**HWE**	
Country	**Data**	**Model**	**Data**	**Model**
Eastern Europe	−1.9	−3.0	0.2	−2.8
Scandinavia	12.8	−3.8	−22.9	−4.4
Western Europe	1.1	−8.6	−27.4	−8.8
Southern Europe	−15.4	−9.8	−11.6	−0.4

Note: For the employment rate we show percentage point differences, and for hours worked per employed percent differences.

and Southern Europe for which the model does not replicate the fact that hours worked differences to the US are larger for married women than for married men.

Summarizing, through the lens of the model, differences in taxation, wages, and the educational composition can account for a substantial fraction of the hours worked difference of married women in Eastern Europe, Scandinavia, and Western Europe, compared to the US, and for 46 percent of the large differences in hours worked per married woman between the US and Southern Europe.

The correlation between married women's hours worked per person in model and data amounts to 0.56 and is thus higher than for men. Also, the model is able to replicate the low correlation of male and female hours worked per person in the data: while the correlation is slightly negative with -.12 in the data, it is slightly positive with .06 in the model.

5.2 Extensive and Intensive Margins of Married Women's Hours Worked

5.2.1 Results for Both Margins

For married women, we further compare the performance of the model in explaining extensive and intensive margin differences relative to the US. Table 5 shows the corresponding results. The first two columns show employment rate differences relative to the US (in percentage points) in the data and the model, respectively, the next two columns show hours worked per employed differences (in percent).

The model generates a close fit on both margins for Eastern Europe. Remember that Eastern Europe was the only region that showed a similar decomposition into both margins as the US. For Southern Europe, the model reproduces 60% of the employment rate difference, but does not generate any difference in the intensive margin to the US, compared to 12 percent lower hours per employed women in the data.

The decomposition into the extensive and intensive margin does not work well for Scandinavia and Western Europe. Scandinavia and Western Europe exhibit higher employment rates of married women than the US, namely 13 percentage points and 1 percentage point higher ones, respectively. By contrast, the model predicts lower employment rates by 4 and 9 percentage points, respectively. Regarding hours worked per employed, the opposite picture arises, with large negative differences in the data, namely 23 percent for Scandinavia and 27 percent for Western Europe, but smaller predicted negative differences by the model, namely 4 and 9 percent, respectively. For Scandinavia and Western Europe, the model thus has difficulties explaining the decomposition of hours worked per married woman into an intensive and an extensive margin. It does however an overall decent job in establishing the relative ordering of countries in hours worked per employed and employment rate differences relative to the US: the correlations between model and data amount to 0.58 and 0.36, respectively. Similarly, the model captures the cross-country correlations between male and female labor supply. It not only replicates the almost zero correlation between male and female hours worked per person fairly well (as already reported in the previous subsection), but also the negative correlation between male hours worked per person and female employment rates (-.30 in the data and -.26 in the model), and at least qualitatively generates the positive correlation between male hours worked per person and female hours worked per employed (.13 in the data vs. .55 in the model).

It is not entirely surprising that taxes, wages, and the educational composition have difficulties explaining the vastly different decomposition into intensive and extensive margins in Scandinavia and Western Europe on the one hand, and Eastern and Southern Europe and the US on the other hand shown in Figure 3. Taxes and wages affect both margins in the same direction, and thus the predictions relative to the US are qualitatively similar across both margins. As a consequence, the model features a positive correlation of .35 between married women's employment rates and hours worked per employed, while in the data the correlation is -.37.

5.2.2 Wedges and the Margins of Labor Supply

The results so far indicate that there are driving forces outside of our model factors of taxes, wages, and educational composition that generate the different decompositions of married women's hours worked into the extensive and the intensive margin across countries. To get some insights into what these driving forces are, we introduce two country-specific wedges into the model, $\hat{\tau}_E$ and $\hat{\tau}_{h_f}$, in order to match the country-specific employment rate and hours worked per employed. The introduction of these two wedges is a reduced form exercise in the spirit of Chari et al. (2007) and **?**. The wedges could potentially capture very different factors like restrictions on hours or employment from the labor demand side, regulation, child care, parental leave, etc. with which we correlate them later on.

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

$$\max_{h_m,h_f}\left\{\ln c-\alpha h_m^{1+\frac{1}{\phi}}-\alpha h_f^{1+\frac{1}{\phi}}-\mathbf{1}_{\{h_f>0\}}q+\mathbf{1}_{\{h_f>0\}}\hat{\tau}_E+\hat{\tau}_{h_f}\cdot h_f\right\} \tag{8}$$

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 labor supply. 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 affects primarily the intensive margin choice: a negative $\hat{\tau}_{h_f}$ decreases the utility of work for every additional hour of work. Thereby, it also has an indirect effect on the extensive margin choice, possibly inducing some marginal women to stop working. We calibrate both wedges to match the country-specific employment rates and hours worked per employed of married women.

The resulting wedges are shown in Figure 5 on the vertical axis. The wedge is negative for all countries but the Czech Republic, Greece, Poland, and Hungary. It thus implies for the majority of countries a tax on hours worked per employed, driving down the optimal hours choice. The wedge is most negative for the Netherlands, for which hours worked per employed womanare lowest among our sample countries. The employment wedge, by contrast, is positive for most countries, implying a subsidy on employment. Only Poland, Greece, and Hungary exhibit a negative employment wedge. The wedge is largest again for the Netherlands, closely followed by Norway and Sweden. Both wedges are highly negatively correlated with a value of -0.82. This indicates that there might be one common underlying factor that drives the different decompositions into the extensive and the intensive margin across countries.

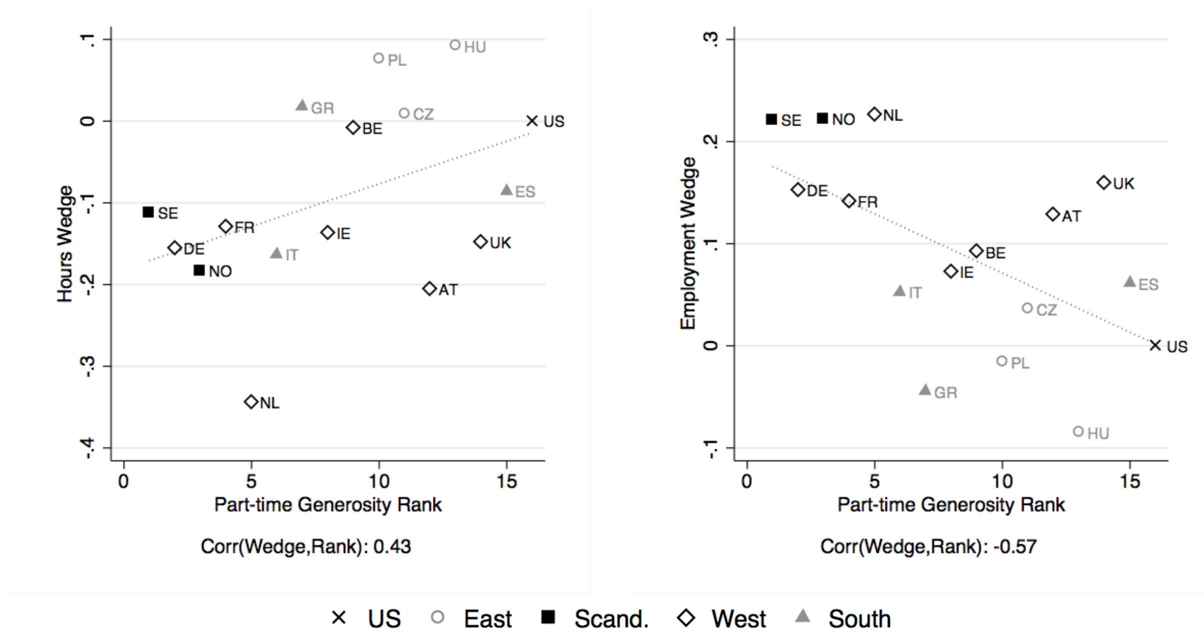
To understand what this common factor could be, we correlate the wedge with country-specific variables relating to child care arrangements, maternal leave policies, weeks of vacation/public holidays, and part-time generosity. The last factor, namely a measure of part-time generosity, correlates strongest with the two wedges, and also intuitively can explain both wedges well. Suppose that there exist a group of women in a country who find it optimal to work but provide relatively few hours, e.g. because of family commitments. If part-time jobs are not available in this country, these women might prefer not to work to working full-time, resulting in a low average employment rate and high average hours worked per employed. If the rights to work part-time in this country are expanded, these women will enter the work force at the desired low hours, thereby increasing average employment rates and decreasing average hours worked per employed. For the empirical exercise, we rely on the part-time generosity ranking of countries provided by the OECD. 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.

Figure 5 shows both wedges on the vertical axis plotted against the OECD part-time generosity rank on the horizontal axis. The hours wedge in the left panel exhibits a positive correlation with the part-time generosity rank of 0.42. The employment wedge in the right panel, on the other hand, features a negative correlation of -0.58 with the part-time generosity rank. 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. Incorporating part-time regulation into the model is however difficult for two reasons: First, the OECD only provides an ordinal country ranking that relies on many different features of part-time regulation. Secondly, modeling part-time regulation without assuming what one wants to find by imposing a restrictive hours grid is not trivial. It requires a better understanding of the sources of differences in part-time regulations, which is beyond the scope of this paper.

Figure 5: Wedges and OECD Part-time Generosity Rank

(a) Hours Wedge

(b) Employment Wedge



Rank 1 corresponds to the highest part-time generosity.

Table A.9 in the online appendix shows correlations of the wedges with other potential driving forces, namely the number of vacation days and public holidays, two measures of parental leave length (one relating to paid maternal leave from the Family Policy Database of the Max-Planck-Institute for Demographic Research, which provides historic data but does not cover Eastern European countries; the other one relating to paid plus unpaid parental leave from the OECD Family Database, which refers however to the years 2012 to 2014), two measures of pay replacement rates during these leaves from the same data sources, public child care expenditure over GDP from the OECD, and net child care cost as percent of average female pay from the OECD. The number of vacation days and public holidays exhibit a positive correlation with 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 correlations with the wedges are less strong. The correlations of both wedges with leave policies are generally low. The correlations with public child care expenditure over GDP are higher and go in the expected direction: higher public expenditure correlates positively with the employment wedge, i.e. acting as a subsidy on employment, and has a weaker but negative correlation with the hours wedge. Thus, child care issues might be an additional potential driving force of the different margin decompositions. However, net child care cost exhibit very weak correlations with both margins, which both go in the unexpected direction (high child care costs are associated with a positive employment wedge and a negative hours wedge). This indicates that other factors than cost, for example quality or availability of slots, might play an even more important role than the cost of child care itself. These factors are hard to model and to measure in the data. Therefore, similarly to part-time regulation, modeling child care across countries is not trivial, and we abstain from doing it. Moreover, informal child care arrangements also seem to matter and are hard to model: according to the OECD, on average across our sample countries 27% of preschool children and 20% of school children receive some form of informal child care, with the average hours conditional on receiving this form of child care by 15 and 10 hours, respectively.

We now proceed by analyzing the effect of taxes, wages, and demographics on hours worked per married woman in more detail without analyzing the two margins any further. One might however be worried that there are important interactions between these factors and the factors driving the different decomposition into both margins, which would make our exercise less insightful. Importantly, at the end of the next section, we come back to the wedges exercise and show that the effects of the driving forces in our model are very similar qualitatively and quantitatively whether country-specific wedges are included into the model or not. We therefore conclude that the effects of taxes, wages, and demographics on married women's hours worked per person can be meaningfully analyzed in the benchmark model despite the model's difficulties in replicating the extensive and intensive margins.

5.3 Decomposition Analysis

To understand the relative importance of wages, taxes, and the demographic composition in explaining the cross-country differences in hours worked per person, we simulate the model changing the input factors to country-specific levels one by one. In the first subsection, we start out doing so by setting only one feature of the economic environment country-specific and leaving all others as in the US. We present the results focusing on country-group averages, and explain the workings of the incentive effects of the different model components on hours of married men and married women. Subsection 5.3.2 focuses explicitly on the labor income tax effects and decomposes them further into an average tax rate effect and a tax structure effect. In the following subsection, we show country-specific results by changing model inputs in a cumulative fashion, and discuss the ability of the model to explain the large variation in married women's hours worked within Europe. Last, in Subsection 5.3.4 we reintroduce wedges into the model to show that the incentive effects of our model factors are independent of the model's difficulties in replicating the extensive and intensive margins.

5.3.1 Effects on Married Couples

We start out confronting households living in the US environment with one specific model input of the European countries at each step. For each exercise, we adjust the transfers such that the government always maintains a balanced budget. Results from this decomposition analysis are shown in Table 6. Columns "Data" and "Model" in Table 6 replicate the results from Table 4, in which we use the full country-specific economic environment described in Equations (1) to (6). The next four columns set only one single element in Equations (1) to (6) specific to country i, namely consumption taxes (τ_c), non-linear labor income taxes (τ_l), the demographic composition and matching into couples ($\mu(x,z)$), or wages (w), while keeping all

Table 6: Decomposing the Effects of Different Model Inputs on Labor Supply (% difference to the US)

(a) Male HWP

Country	**Data**	**Model**	τ_c	τ_l	$\mu(x,z)$	w
Eastern Europe	−11.1	−6.3	−2.7	−4.9	2.1	−1.7
Scandinavia	−17.1	−11.1	−4.7	−9.6	0.0	1.2
Western Europe	−10.3	−7.8	−3.0	−6.4	1.8	−0.1
Southern Europe	−9.4	−4.9	−2.3	−4.6	2.4	−0.2

(b) Female HWP

Country	**Data**	**Model**	τ_c	τ_l	$\mu(x,z)$	w
Eastern Europe	−2.6	−7.0	−7.2	1.6	−2.1	−1.9
Scandinavia	−8.9	−9.5	−12.6	−2.0	0.9	1.7
Western Europe	−26.3	−19.8	−7.9	−4.9	−6.8	0.6
Southern Europe	−30.9	−14.3	−6.3	6.4	−10.9	−5.0

Note: For the decomposition in columns 3 to 7, exactly one model input is set country-specific, and the rest are left unchanged at their US values.

others at the US level. When applying the country-specific labor income tax schedule, we account for the fact that tax systems are defined relative to the income level in a country by calculating tax rates associated with hours choices in each country and applying these to the same hours choices in the US. 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.2. Panel (a) of Table 6 shows differences to the US in married men's hours worked per person, while panel (b) presents the results for married women's hours worked per person.

Column 3 starts the decomposition with applying the country-specific consumption tax rates. Not surprisingly, the disincentive effects of consumption taxes on hours worked of both married men and women are largest for Scandinavia, where consumption taxes are highest. Consumption taxes alone predict that Scandinavian married women should work 12.6 percent less hours than US married women. For all other European regions, the effects of consumption taxes alone on married women's labor supply are also sizeable, predicting between 6 and 8 percent lower hours worked than in the US. For married men, the disincentive effects of European consumption taxes are smaller, with 2 to 5 percent. 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.[21] For each region, the consumption tax effects on male hours worked are only one third as large as on female ones. The correlation between male and female hours if only the

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

consumption tax is set country-specific is extremely high with 0.98. Thus, the consumption tax works in exactly the same way on male and female hours worked.

While consumption taxes are on average the main factor driving European female hours worked down, labor income taxes play the largest role in explaining on average low European hours worked relative to the US for men, as can be seen in column 4. Labor income taxes uniformly predict lower hours worked of married men in Europe than in the US by 4 to 10 percent. Their disincentive effect on male hours worked is thus around twice as large as the disincentive effect coming from consumption taxes. As for consumption taxes, the disincentive effect is largest in Scandinavia. This is in line with the consumption tax rates and average income tax rates for men shown in Table 2, which are both largest for Scandinavia, followed by Western Europe. For married women, however, the picture now looks very different. Labor income taxes in fact predict lower hours worked for married women than in the US only for Western Europe and Scandinavia, but *higher* ones for Eastern Europe and Southern Europe. Even for Western Europe, the disincentive effect of labor income taxes is only around two thirds of the size of the disincentive effect of consumption taxes, while for Scandinavia, it is only one sixth of the size of the consumption tax effect. The low correlation of 0.33 of the labor income tax effect between men and women stands in striking contrast to the high correlation of the consumption tax effect. We analyze the reason for this low correlation between men and women in the next Subsection 5.3.2, which investigates the labor income taxes in more detail.

Turning to the educational composition and matching into couples in column 5, one can see that they induce a small increase in hours worked for married men for all European regions. This is at the first view surprising, given that the share of high-educated individuals is highest in the US, and the share of low-educated lowest. There are two reasons for the result. First, the hours gradient by education is relatively small for men. Secondly, and more importantly, the educational composition has a much larger effect on the labor supply of married women given that they also adjust along the extensive margin. Turning to women in panel (b), one can see that the educational composition alone predicts 11 percent lower hours worked per person in Southern Europe than in the US. Applying the Southern European educational composition substantially increases the share of low-educated women, reducing the employment rate. This effect is so large that it induces men to on average increase their hours, despite the fact that the less favorable educational composition also applies to them. As it turns out, the educational composition alone can explain one third of the lower labor supply of Southern European married women, and is the main model force behind the decrease in hours there. For Western Europe, the educational composition alone also predicts 7 percent lower hours than in the US, which amounts to one third of the model predicted difference. Western Europe is the region in which the share of low-educated individuals is second largest after Southern Europe. The country-specific educational composition leads to a strong negative cross-country correlation of -0.6 between male and female hours, but causes only a small variation in male hours.

The effect of the country-specific gender and educational wage premia on hours worked in Europe is on average relatively small, as can be seen in the last column of Table 6. It is largest for Southern European married women: wages alone predict 5 percent lower hours worked than in the US. For Scandinavia and

Western European women, the wage effect goes in the opposite direction as the data, namely predicting higher hours worked in Europe than in the US, but the effects are small.

Summarizing, the decomposition analysis shows that labor income taxes are of biggest importance in explaining the lower hours worked of married men in Europe than in the US. By contrast, consumption taxes, which work qualitatively very similar on male and female labor supply, are the main driver of low hours worked of married women in Europe. The educational composition substantially drives down hours of Southern European married women, and to a lesser degree of Western European ones. The effect of wage differences on hours worked differences is relatively small for both men and women.

5.3.2 Investigating Labor Income Taxes in More Detail

In the previous subsection, we saw that labor income taxes uniformly predict lower hours worked in Europe than in the US for married men, but for married women only in Western Europe and Scandinavia. In this subsection, we derive more insights into why this is the case by further decomposing the actual non-linear labor income tax schedule two components: first, the average tax rate, i.e. the level of the tax schedule, and second, the actual marginal tax schedule along with the tax treatment of married couples, which together define how the marginal tax rate for each spouse changes with the own and the spousal income. We call this second component tax structure.

To distinguish between the average tax rate and the tax structure, we conduct 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 such that government revenues are left unchanged at the US level.[22] 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.

Table 7 shows the resulting decomposition of the labor income tax effect into the average tax rate and the tax structure. As in the previous table, the upper panel shows results for hours worked per person of married men. For married men, the tax structure uniformly predicts lower hours worked in Europe than in the US, while the average tax rate has a smaller effect, and even a slightly positive one in Eastern Europe. This indicates the higher progressivity of taxes in Europe. The mechanism is reinforced since the degree of joint taxation is more pronounced in the US than in most other countries, reducing the marginal tax rate of American married men. On top of that comes an additional disincentive effect through the higher average tax rates in Scandinavia and Western Europe.

For married women, the results are very different, with the tax structure on average predicting *higher* hours worked in Europe than in the US, rather than lower hours worked. Focusing first on the average tax rate, we see that, as for consumption taxes, the relative ordering of the regional effects is the same for men

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

Table 7: Effects of Average Tax Rate vs. Tax Structure on Labor Supply (% difference to the US)

(a) Male HWP

Country	**Data**	τ_l	**Average** τ_l	**Structure** τ_l
Eastern Europe	−11.1	−4.9	0.5	−5.4
Scandinavia	−17.1	−9.6	−3.9	−5.7
Western Europe	−10.3	−6.4	−2.2	−4.2
Southern Europe	−9.4	−4.6	−0.8	−3.8

(b) Female HWP

Country	**Data**	τ_l	**Average** τ_l	**Structure** τ_l
Eastern Europe	−2.6	1.6	−0.2	1.9
Scandinavia	−8.9	−2.0	−9.4	7.4
Western Europe	−26.3	−4.9	−4.5	−0.4
Southern Europe	−30.9	6.4	−2.0	8.5

Note: Columns 4 and 5 add up to Column 3. Column 3 corresponds to column 4 in Table 6.

and women, with again larger effects for women due to the higher implied female elasticity. Moreover, the correlation between male and female hours if only the average tax rate is set country-specific is with 0.98 almost perfect and resembles the one of consumption taxes. The tax level predicts 9 and 5 percent lower hours worked by married women in Scandinavia and Western Europe, respectively, with a small effect of -2 percent in Southern Europe and almost no effect in Eastern Europe. This ordering is in line with the measure of the average tax rate shown in Table 2, which was highest in Scandinavia, followed by Western Europe, then Eastern and Southern Europe, and is also in line with the average tax rate effect for married men. Thus, the disincentive effects of average labor income tax rates and of consumption taxes are quite easy to deduce from the data, are qualitatively exactly the same for men and women, but always larger for women than for men. The disincentive effects of consumption taxes are larger than the disincentive effects of the average labor income tax rates for both men and women, which is again in line with the data shown in Table 2, in which the Europe-US differences in consumption tax rate is larger than the difference in the exemplary average tax rate from column 2.

By contrast, the tax structure itself predicts uniformly *higher* hours worked of married women in Europe than in the US, with the exception of Western Europe. For Western Europe, the tax structure alone would predict 0.4 percent lower hours worked per married woman than in the US, for Eastern Europe 2 percent higher hours, and for Southern Europe and Scandinavia 7 and 9 percent higher hours. In fact, there are only four countries for which the tax structure alone would predict lower hours worked of married women in Europe than in the US, namely Belgium, Germany, and Ireland, all located in Western Europe, plus the

Czech Republic. Western Europe is the region that exhibits the largest heterogeneity in the tax structure effect: it ranges from predicting 21.3 percent lower hours than in the US for Germany, and 5.9 percent higher hours than in the US for Austria.[23]

The reason behind this positive effect of the tax structure on married women's hours worked in the majority of European countries compared to the US lies in the joint taxation of married couples in the US. Joint taxation makes the marginal tax rate of the each spouse depend on the other spouse's income. As a result, if a woman starts working, she faces a relatively high marginal tax rate, which is increasing in her husband's income. Table 2 presented as an exemplary measure the average marginal tax rate if a woman starts working the same hours as the average US woman. This rate exceeds 50% in Germany, but is close to only 30% in the US, reflecting the higher progressivity of the general labor income tax schedule in Germany even though both countries feature joint taxation. This explains the larger disincentive effect of the tax schedule on married women's hours worked in Germany than in the US predicted by the model. In contrast, most of the other European countries feature systems closer to separate taxation than the US system. Therefore, the tax structure predicts higher hours worked of married women for the majority of European countries. In contrast to the almost perfect correlation of male and female hours if only the consumption tax or only the average labor income tax rate are set country-specific, this correlation is only 0.16 if only the tax structure is set country-specific. This again underlines the different workings of the tax structure for both genders.

Thus, the decomposition of the labor income tax into the tax schedule and the average tax rate makes it clear that the average tax rate alone is not a good approximation of the incentive effects of income taxes on married women's labor supply. [24] Especially to generate the comparatively high hours worked of married women in Eastern Europe and Scandinavia, incorporating the tax structure is crucial. For Southern Europe, by contrast, the tax structure would predict higher hours worked than in the US, and thus makes it more difficult to predict hours worked differences in the data.

5.3.3 Married Women: Variation Within Europe

While so far we have concentrated on country group averages and differences to the US, but discussed results for both married men and women, we now shift our attention more in detail to country-specific results and analyze within European variation in greater detail focusing on women only. Figure 6 presents scatterplots of hours worked per married woman in the data on the x-axis and in the model on the y-axis. 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 model inputs sequentially into the model. Thus, instead of changing only one input at a time, the inputs are adjusted in a cumulative fashion. We start with the two factors that are less of a focus of the paper, educational composition and wages, and then add the taxes. There are some non-linearities

[23]The average tax rate effect is negative in all Western European countries but Ireland.

[24]Online Appendix Section A.6 compares the model performance relying on the non-linear labor income tax schedule or relying on linear tax rates as calculated by McDaniel (2011) and Ragan (2013), and confirms that a model with linear taxes cannot match hours worked of married women in Europe well.

between wages and the labor income tax code, but results remain qualitatively the same if wages are added at the end.

Panel (b) introduces the country-specific educational composition into the model. The educational composition tilts the model predicted hours in the right direction: given the higher share of low- and medium-educated workers, it reduces predicted hours in Southern and Western Europe, and to a much smaller degree in Eastern Europe, leaving hours in Scandinavian countries on average unchanged. The correlation of hours in data and model is very high with 0.79, with a slope of 0.23 if the model output is regressed on the data.[25] Applying only the country-specific educational composition plus matching into couples predicts 67 lower hours on average in Europe than in the US, compared to a difference of 249 hours in the data, i.e. it explains around a quarter of the average hours difference.

Panel (c) adds country-specific wages to the educational composition. Adding the country-specific education-gender-wage premia leaves the average hours difference to the US almost unchanged, as well as the regression slope, but worsens the correlation between data and model, which drops from 0.79 to 0.37.[26] Thus, differences in the wage premia do not help in explaining hours worked by married females any further.

Panel (d) then adds on top of this the country-specific consumption tax rates. Consumption taxes shift predicted hours in Europe substantially downward compared to US hours, by on average 100 hours, compared to a difference of 249 hours in the data. They thus explain 40% of the average US-Europe hours difference, and are the most important driver of low European hours of married women. However, they do nothing further to explain the variation *within* Europe: the slope of the regression line of model hours on data hours falls slightly, as the correlation between hours in model and data remains virtually unchanged. The only noticeable effect on the variation within Europe is that the predicted hours in Norway and especially Sweden fall more than in the other countries.

Panel (e) now adds the average labor income tax rate from the decomposition exercise as a country-specific model input. Average European hours fall by 30 hours compared to panel (c), 12 percent of the total hours difference. While thus average tax rates on average predict lower hours in Europe than the US, predicted hours in Ireland increase substantially, as to a smaller degree do hours in the Czech Republic, Hungary, and Spain. The slope of the regression line increases somewhat from 0.21 to 0.27, and also the correlation increases mildly from 0.36 of 0.41. Thus, adding the average tax rates improves the model fit a bit in all dimensions.

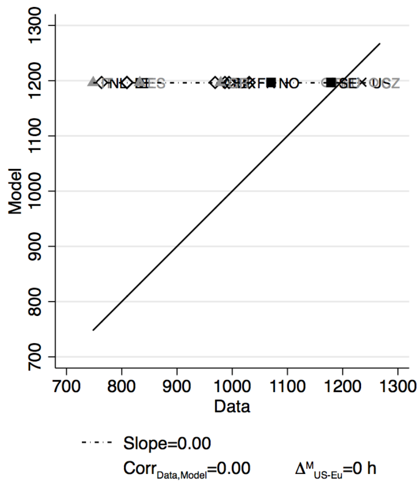
Panel (f) finally adds the tax structure of the labor income tax code and presents the full country-specific results, already shown in Panel (c) of Figure 4. Predicted hours in Europe slightly fall compared to panel (e), i.e. on average the effect of the tax structure goes into the wrong direction, as seen before. However, the tax structure substantially helps to improve the fit within Europe. The slope increases from 0.27 to 0.36, and the correlation from 0.41 to 0.56. A few countries move however further away from the hours in the data at

[25]A perfect fit would imply a slope of 1 along the 45 degree line. Since US female hours are not matched perfectly in the calibration, the US is slightly below the 45 degree line.

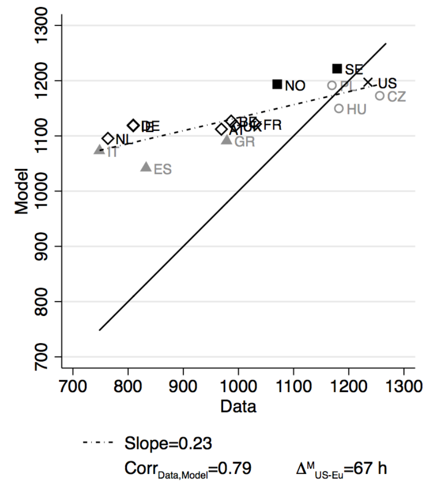
[26]Given log utility of consumption, differences in mean wages do not affect choices at all, and only the wage premia matter.

Figure 6: Decomposition: Cumulative Country-Specific Results

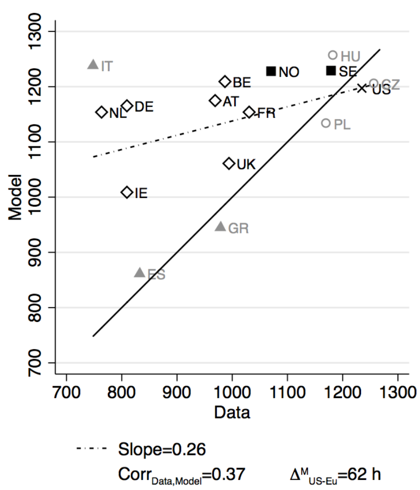
(a) US Environment



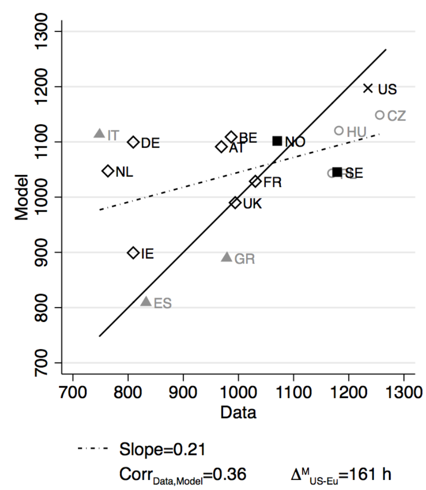
(b) + Educational Composition



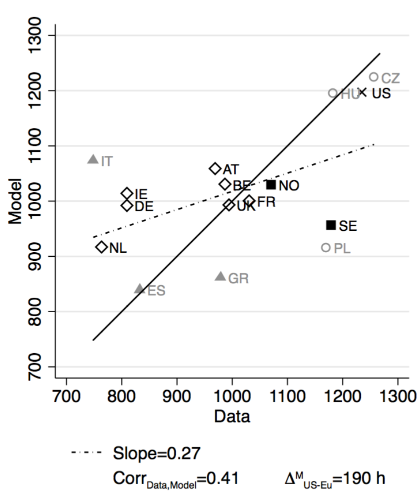
(c) + Wages



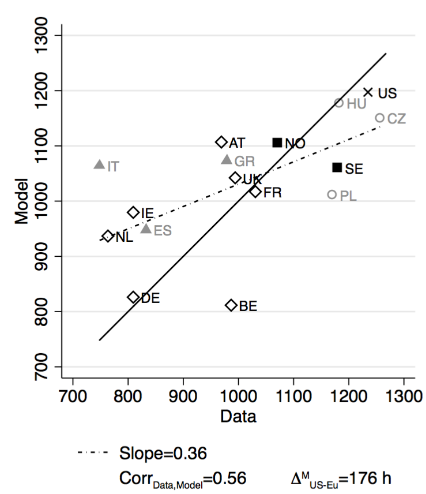
(d) + Consumption Tax



(e) + Average Tax Rate



(f) + Tax Structure (=full country-specific)



this step. Specifically, one can see the positive effect of the tax structure on Spain , which makes it harder to predict their low hours. The jointness of taxation in Belgium and the Czech Republic by contrast decreases their hours too much.

Summarizing, Figure 6 confirms that consumption taxes are the major driving force of *on average* low hours worked of married women in Europe. Analyzing average labor income tax rates alone would bring no success in explaining female hours in Europe. Taking the non-linearity of the tax codes seriously and dealing with the jointness of tax systems is crucial to explain the *variation* within Europe. The educational composition of the workforce adds further to explaining this variation.

5.3.4 Decomposition With and Without Wedges

One might be worried that the incentive effects of taxes, wages, and the educational composition would work differently if both margins of labor supply would be matched perfectly across countries, i.e. in the model with wedges. To dispel these concerns, we repeat the decomposition exercise from Table 6 in the model with country-specific wedges, and show in Figure 7 the effect on hours of setting either only the consumption tax rate (Panel (a)) or the labor income tax rate (Panel (b)) country-specific in the model without wedges and in the model with wedges.[27] In the model with wedges, 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. As one can see, the incentive effects of the two taxes are very similar in the model with or without wedges, both qualitatively and quantitatively. The largest differences arise for Sweden in the case of the consumption tax, and the Netherlands in case of the labor income tax. Overall, these figures indicate that our decomposition analysis delivers meaningful insights even if we do not explicitly model the factors that lead to the different margin decomposition of married women's hours worked across countries. Figure A.6 in the Online Appendix confirms that this is also the case for wages and the educational composition, as well as for the decomposition of labor income taxes into the average tax rate and the tax structure.

5.4 Two Outliers: Denmark and Portugal

As explained in the introduction, there are two outliers that we shortly discuss separately in this section, namely Denmark within Scandinavia and Portugal within Southern Europe. Danish married women work essentially the same number of hours as US married women, but the model predicts a difference of minus 30 percent (see Table 8). The failure of the model to replicate the labor supply behavior of Danish married women comes from the joint effect of consumption and labor income taxes. In contrast to the other Scandinavian countries, the labor income tax also predicts substantially lower hours in Denmark than in the US: Denmark features the highest average tax rate in Scandinavia and combines this with a tax system that features strong elements of joint taxation, which is not the case in the other Scandinavian countries. While

[27] The country-specific values in the model without wedges are exactly the ones that generate the country-group average effects in Table 6.

Figure 7: Decomposition With and Without Wedges

(a) Consumption Tax — US Env. & $\hat{\tau}$s (b) Labor Income Tax — US Env. & $\hat{\tau}$s

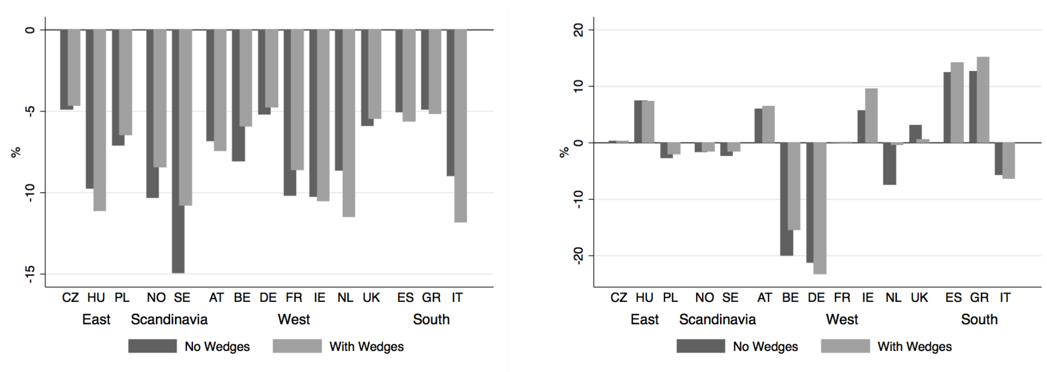


Table 8: Decomposing Hours Worked Differences Relative to the US for Denmark and Portugal

	Data	Model	μ	τ_c	τ_l	w	Tax Avg. τ_l	Tax Str. τ_l
Denmark								
HWP$_m$	−12.8	−16.0	0.4	−5.4	−13.9	0.0	−9.4	−4.5
HWP$_f$	−1.4	−29.7	−2.4	−14.2	−23.4	4.5	−23.7	0.3
Portugal								
HWP$_m$	−10.2	−3.0	4.0	−2.8	−1.0	−1.5	1.2	−2.2
HWP$_f$	2.6	−18.9	−16.6	−7.2	3.3	−10.8	2.9	0.4

the tax structure predicts essentially the same hours in Denmark as in the US, the effect is relatively weak compared to the other Scandinavian countries. Coupled with the much stronger negative effect of the high average tax rate, this leads to a predicted difference of minus 23 percent based on the income taxes alone.

Portugal, on the other hand, is a clear outlier when it comes to Southern Europe from the data side. While Greek, Spanish, and Italian married women work between 20 and 40 percent fewer hours than US married women, Portuguese married women work even slightly more hours than US ones. While the Portuguese labor income tax system alone correctly predicts higher hours worked in Portugal than in the US, consumption taxes, wages, and the demographic composition all predict lower hours worked, such that in the end the model predicts 19 percent lower hours worked of married women in Portugal than in the US. Note, however, that the tax structure effect goes into the correct direction for both countries, closing the gap between model and data.

6 Conclusion

Relying on three micro data sets, we document average hours worked of married couples for a sample of eighteen European countries and the US over the time period 2001 to 2008. We find that hours worked vary significantly across countries, and the largest variations can be found for married women. Whereas European married men work relatively homogeneously between 9 and 17 percent fewer hours than US married men, the picture for married women is much more heterogeneous, with Eastern European and Scandinavian married women working only 3 and 9 percent fewer hours than US married women, but Western and Southern European women working 26 and 31 percent fewer hours.

We investigate in how far international differences in consumption taxes, labor income tax systems, gender wage gaps and educational premia, and the educational composition and matching into couples 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, holding preferences constant across countries. The model only slightly underpredicts the observed international hours differences for married men. Moreover, it does a good job in explaining international differences in hours worked by married women. Specifically, the model is able to replicate the low hours worked per married woman differences between the US and Eastern Europe and Scandinavia, as well as the large differences between the US and Western Europe. Only for Southern Europe can the model only explain 46 percent of the observed difference.

A decomposition analysis shows that consumption taxes offer significant disincentives to work in all European countries and for both sexes. The effect of non-linear labor income taxes is however much more complicated: for married men, labor income taxes always predict lower hours worked in Europe than in the US, which is however not true for married women. While the on average higher tax rates in Europe also provide a disincentive effect for married women, the tax structure is with a few exceptions more favorable for European women than for US ones, since the US features a system of joint taxation of married couples. The labor income tax rates explain a large part of the variation of married women's hours worked within Europe, aided by the educational composition.

The model that we use in this paper is a very simple one. It does e.g. not incorporate a home production sector in competition with the service sector of the economy, a subsistence level of consumption, a life cycle component, or income risk. The success of the model is thus quite remarkable. While the model results leave some scope for other factors explaining hours worked of married women, they come quite close to replicating the data. Taxes, wages, and the educational composition thus have large explanatory power for international differences in hours worked of married women. 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 into the extensive and the intensive margin for married women in Scandinavia and Western Europe remain as an open question for future research.

References

Albanesi, S. and C. Olivetti (2009). Gender Roles and Medical Progress. Working Paper, Columbia University and Boston University.

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.

Bick, A., B. Brüggemann, and N. Fuchs-Schündeln (2015). Hours Worked in Europe and the US: New Data, New Answers. Working Paper, Goethe University Frankfurt.

Blundell, R. and T. MaCurdy (1999). *Labor Supply: A Review of Alternative Approaches*, Chapter 27, pp. 1559–1695. Elsevier Science B.V.

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 (2012). Marriage Stability, Taxation and Aggregate Labor Supply in the US vs. Europe. Working Paper, Southern Methodist University, Uppsala University and Magyar Nemzeti Bank.

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.

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

Duernecker, G. and B. Herrendorf (2015). On the Allocation of Time A Quantitative Analysis of the US and France. Working paper, University of Mannheim and Arizona State University.

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.

Jones, L. E., R. E. Manuelli, and E. R. McGrattan (2003). Why Are Married Women Working So Much? Working paper, Federal Reserve Bank of Minneapolis.

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. Special issue: Cross-Sectional Facts for Macroeconomists.

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, Arizona State University.

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 Womens 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, University of Zurich.

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. (2014). Social Security and Cross-Country Differences in Hours: A General Equilibrium Analysis. *Journal of Economic Dynamics and Control 37*, 2466–2482.

A Appendix

A.1 Data Appendix

A.1.1 Data Issues in the ELFS

For data reasons detailed in Bick et al. (2015), we exclude the years 2001 for the UK and 2005 for Spain from the analysis. Furthermore, we exclude the year 2001 for Italy and the year 2008 for Ireland because the OECD TaxBen Module does not produce the corresponding tax rates. We exclude households in the ELFS of which at least one member lives in an institution, since the CPS does not cover individuals living in institutions. This leads to the deletion of a negligent number of observations. A detailed description of all issues involving setting up the data sets can be found in Bick et al. (2015).

A.1.2 Dealing with Missing Household Identifiers in Scandinavia

- Omit Scandinavia from analyses with children
- We do not have any information on children in Scandinavian countries for our sample period. For these countries, we calculate the distribution of children from the EU-SILC data. We use the 2004 value for the years 2001 to 2003.
- For the Scandinavian countries we do not have any information on spousal education. We therefore base the matching into couples on data from EU-SILC applying the same sample selection criteria.
- How to impute matched data in Scandinavia from EU-SILC into ELFS

A.2 Model Decomposition Analysis

The fifth column (τ_l) of Table 6 shows the results if the 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, we calculate for any combination of male and female hours choices the country-specific tax rate using the US gender-specific education premia and the country-specific mean wage ($\bar{w}^i$) . Second, we apply the resulting average tax rate to the corresponding US household income 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 column 6 of Table 6. 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.