USC FBE MACROECONOMICS &
INTERNATIONAL FINANCE WORKSHOP
presented by: Nicola Fuchs-Schundeln
FRIDAY, April 15, 2016
3:30 - 5:00 pm, Room: HOH-506

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 *

March 31, 2016

Abstract

We document contemporaneous differences in the aggregate labor supply of married couples across 17 European countries and the US. Based on a model of joint household decision making, we quantify the contribution of international differences in non-linear labor income taxes and consumption taxes, as well as the educational composition, and gender wage gaps and educational premia, to the international differences in hours worked in the data. Through the lens of our model, taxes, wages, and the educational composition account for a large part of the small differences in married men's and the large differences in married women's hours worked between the US and Europe in the data. 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, Paul Klein, Dirk Krueger, José-Víctor Ríos-Rull, Richard Rogerson, Kjetil Storesletten, the editor and four anonymous referees, as well as seminar participants at Arizona State University, Carlos III, CEMFI, Chicago Fed, DIW Berlin, Free University Berlin, ITAM, IZA, LMU Munich, LSE, Minneapolis Fed, Northwestern, Pompeu Fabra, Princeton, Stanford, St. Louis Fed, Tilburg, Tinbergen Institute, Universitat Autònoma de Barcelona, UC Berkeley, Universities of Alberta, Edinburgh, Hamburg, Houston, Konstanz, Mainz, Michigan, Montreal, and Nürnberg-Erlangen, UT Austin, Uppsala University, and Yale, and conference participants at the Annual Meetings of the Society for Economic Dynamics (2014) and the German Economic Association (2011), Barcelona GSE Summer Forum Workshop on Micro and Macro Perspectives of Taxation (2014), Canadian Macro Study Group (2015), CASEE conferences at Arizona State University on "Quantitative Macroeconomis Conference" (2012) and "Household and Female Labor Supply" (2015), Christmas Meeting of German Economists Abroad (2012), NBER Summer Institute Macro Public Finance Workshop (2012), North American Winter Meetings of the Econometric Society (2013), Midwest Macro Meetings (Fall 2013), and the Philadelphia Fed/NBER Conference on Marcoeconomics across Time and Space (2014) for helpful comments and suggestions. Bettina Brüggemann and Hannah Paule-Paludkiewicz provided truly outstanding research assistance with the data, and Enida Bajgoric, Pavlin Tomov, and Desiree Winges with the tax codes. The authors gratefully acknowledge financial support from the Cluster of Excellence "Formation of Normative Orders" at Goethe University and the European Research Council under Starting Grant No. 262116. Alexander Bick further thanks the Research Department of the Federal Reserve Bank of Minneapolis for the hospitality, and the Fritz Thyssen Foundation for the financial support of that stay. All errors are ours.

1 Introduction

The 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 0.79 and 0.64). Explaining these large differences in the labor supply behavior of married men and married women with linear (average marginal) taxes – a prominent and successful predictor in explaining aggregate hours worked differences in the literature pioneered by Prescott (2004) – seems challenging. Even if women have a higher labor supply elasticity than men, the relative country ordering of 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, the educational composition, and taxation. These features can account for 55 percent of the US-Europe gap in male hours worked, and 74 percent of the US-Europe gap in female hours worked, and also replicate the cross-country variation well: the cross-country correlation between hours worked in the model and the data is 0.44 for men and 0.63 for 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 in 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, which comprise both the tax treatment of married couples and the progressivity of the tax code. The tax treatment of married couples across countries ranges from joint to separate taxation, with most 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. Differences in the educational composition across countries further help explain low female hours worked in Southern and Western Europe.

To understand the workings of joint taxation, consider the case of Germany. The incomes of husband and wife are summed up and divided by two, and the household tax burden is then determined as the sum of

[1] The country groups are made up as follows: Eastern Europe – Czech Republic, Hungary, Poland; Scandinavia – Sweden and Norway; Western Europe – Austria, Belgium, France, Germany, Ireland, Netherlands, United Kingdom; 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.

the tax burdens on these two hypothetical equal incomes. Due to the progressivity of the German tax code, this implies that the overall tax burden of the household is therefore lower than under separate taxation. Moreover, the secondary income earner (whose contribution to the household income is less than half) faces a higher marginal tax rate than under separate taxation, while the opposite is true for the primary earner. Thus, the treatment of couples in the labor income tax code leads to additional non-linearities in their tax codes compared to singles, which differ substantially for husband and wife. While the US system works differently, it also features joint taxation, i.e. the marginal tax rate of one spouse is increasing in the income of the other spouse.[2]

The key importance of the tax treatment of married couples can best be 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 *separate* taxation. Figure 1 shows in the left panel the average labor income tax rate of a household in which the wife does not work, plotted against the hours worked by the husband. This average labor income tax rate is slightly larger in Sweden than in Germany, and substantially lower in the US. 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 21.4 percent more than 10 percentage points lower in the US than in Germany and Sweden with 33.4 percent and 33.5 percent, 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, keeping the husband's hours fixed at 1970 hours. This average marginal tax rate is substantially higher in Germany, but similar in Sweden and the US. At the average hours of US married women, indicated with a vertical line at 1235 hours, it amounts to 29.1 percent and 28.0 percent in the US and Sweden, respectively, but is drastically higher in Germany with 49.6 percent. Comparing the left and the right tax schedule, one can see that tax rates are slightly lower for secondary income earners in Sweden, a country with separate taxation, but substantially higher for secondary income earners than for primary income earners in the US and Germany, countries with joint taxation of married couples. In the data, married men in Germany and Sweden work roughly the same hours, around 15 percent fewer hours than in the US. Married women in Sweden, in turn, work only slightly fewer hours than US married women (4 percent), while German married women work 34 percent fewer hours. The model generates this variation because it features both differences in the average tax rate and in the tax structure, combining the progressivity and the tax treatment of couples.

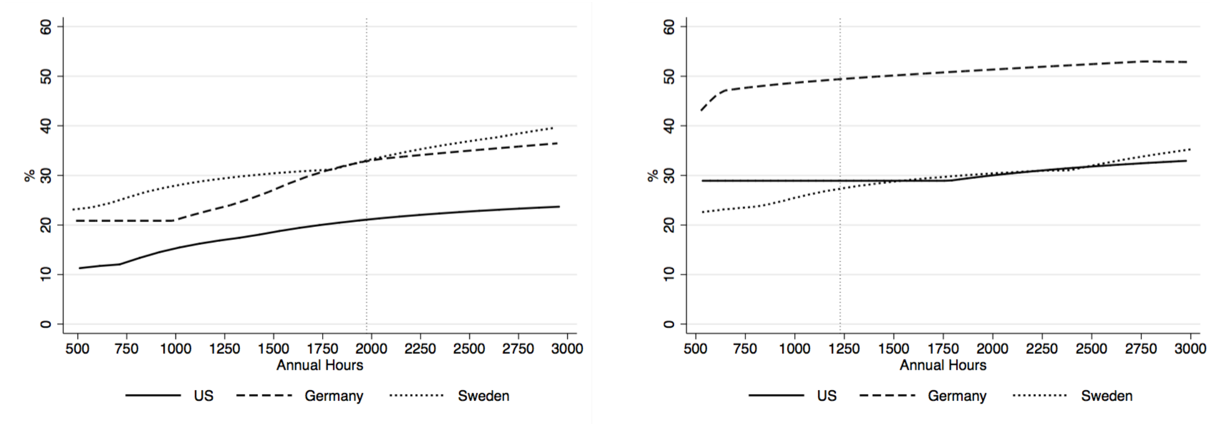
Our quantitative framework is based on the model of joint labor supply of married couples developed in Kaygusuz (2010), Guner et al. (2012a), and Guner et al. (2012b), which features an extensive and an

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

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

(a) Average Income Tax Rate if Only the Husband Works

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



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

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 sample countries, 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 the OECD Taxing Wages modules, which implement the country-specific statutory labor income tax codes in detail, specifically also 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, and explains two thirds of the large hours worked differences between the US and Western Europe or Southern Europe. For married men, both European labor income taxes and to a lesser extent the higher European consumption taxes contribute to explaining the average US-Europe hours difference. For women, in turn, consumption taxes are a crucial factor for explaining the *average* US-Europe hours difference, while non-linear labor income taxes are crucial for explaining the variation in married women's hours worked *within* Europe, and for generating the low observed correlation of hours worked between married men and married women. Consumption taxes and *average* labor income taxes have qualitatively the same effects on men and women, with quantitatively larger effects on married women due to their higher implied labor supply elasticity. However, the non-linearities in the labor income tax codes affect married men's and women's labor supply very differently. 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. The different tax treatments of married couples and the implied differences in the marginal tax rates of the primary (male) and secondary (female) income earner alone imply *higher* hours worked of married women in most European countries than in the US, but significantly lower hours in some Western European countries. As a consequence, the fit of the within-Europe variation in married women's hours worked improves once the non-linearities of the labor income tax codes are taken into account. For married men, this effect, together with the higher progressivity of European tax systems, further decreases European hours relative to US hours. The educational composition and the gender wage gap and educational wage premia play an important role in explaining the low hours worked of Southern European married women. The educational composition also depresses married women's hours worked in Western Europe somewhat, while the effects of wages and the educational composition on female labor supply in Eastern Europe and Scandinavia relative to the US are rather small.

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. Compared 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 our model factors (taxes, wages, and the educational composition) affect both margins qualitatively in the same way. 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, as well as with the availability and quality of child care. We show that the effects on hours worked per person of the various model ingredients considered, most prominently taxes, are quantitatively very similar in our benchmark model, which does not replicate the negative cross-country correlation between the extensive and intensive margin, and in the model with wedges, which by construction exactly replicates this negative correlation.

A series of papers (Prescott (2004), Rogerson (2006), Rogerson (2008), Ohanian et al. (2008), Rogerson (2009)) have shown that differences in consumption and average income tax rates, combined into one linear tax rate on income, can largely explain differences in the time series of *aggregate* hours worked across European countries and the US after World War II, with the exception of Scandinavia. We distinguish explicitly between consumption and labor income taxes, and show that it is crucial to take the full non-linearity of labor income tax schedules into account. 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), Alonso-Ortiz (2014), 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. However, Wallenius (2014) finds that differences in retirement systems have almost no effect on labor supply behavior before retirement, i.e. in the age group we focus on.[3] Guvenen et al. (2014) investigate the role of cross-country differences of non-linear labor income tax codes for singles in driving post- and pre-tax inequality. They show that these tax differences can explain a sizable part of the differences in the US-Europe hours gap for men.

High hours worked in Scandinavia despite high consumption and labor income taxes there have been raised as a puzzle in the literature. Ragan (2013), and in a similar fashion Ngai and Pissarides (2011), suggest government subsidies in sectors that serve as substitutes to home production, especially child care, as the main explanation for this apparent puzzle, a point that was already raised theoretically by Rogerson (2007). They therefore explicitly model home production in addition to market work, and analyze the effects of international differences in government subsidies.[4] By contrast, we can replicate Scandinavian hours well, and even better than Ragan (2013), by taking the non-linearity of the labor income tax code into account.[5] Our approach thus offers a complementary explanation to Ragan (2013) and Ngai and Pissarides (2011) for high hours worked by Scandinavian married women.

The paper most closely related to ours is Chakraborty et al. (2015). They build a comprehensive life-cycle model with idiosyncratic income risk to investigate the cross-country variation in hours worked for married and single men and women. Besides taxes, they concentrate on exogenous marriage and divorce probabilities as driving forces of labor supply differences. Chakraborty et al. (2015) estimate polynomial tax functions for married couples and thus allow for some non-linearity, but base their estimates on the sum of household earnings, and thus cannot fully exploit the differential effects of tax progressivity vs. tax levels on husbands and wives under systems of joint vs. separate taxation. They find that marriage and divorce rates generate variation in female hours that are in line with the data, but that taxes cannot explain any of the variation in female hours worked across countries. The latter result shows how important it is to take the full non-linearities of the tax codes into account. Their model fit is worse than ours for the majority of countries for both men and women.[6]

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

[3] Erosa et al. (2012) and Alonso-Ortiz (2014) only present results for agents from age 50 onwards. By starting at the age 25, we also abstract from international differences in educational systems and youth unemployment rates.

[4] Duernecker and Herrendorf (2015) also show the importance of home production in explaining international aggregate hours worked differences, but focus on differential productivity improvements in the home production sector, rather than government subsidies. Bridgman et al. (2016) report hours of home production for a large set of countries.

[5] Our predictions for hours worked in the three Scandinavian countries are off by less than 1 percent for Sweden, 9 percent for Norway, and 12 percent for Denmark, whereas the corresponding numbers in Ragan (2013) are 39 percent, 41 percent, and 51 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.

[6] Exceptions are Denmark and Portugal for women, and Norway, Sweden, and Spain for men. Moreover, for men the model fit is equally good in their model as in ours for Belgium, the UK, Greece, and Portugal. The comparison metric is the absolute deviation between hours differences to the US in model and data. 14 countries are covered in both Chakraborty et al. (2015) and this paper.

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. Albanesi and Olivetti (2009), and Jones et al. (2014)), which also lead 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)), improvements in maternal health and the introduction of infant formula (Albanesi and Olivetti (2015)), and structural change (Buera et al. (2014), Ngai and Petrongolo (2014), and Rendall (2015)).

Within this literature Rendall (2015) analyzes the effect of taxation on the increase in female labor supply in the US over time, and links both phenomena to structural change by focusing on the size of the service sector. Crossley and Jeon (2007) study in a difference-in-differences approach a Canadian tax reform in 1988 which reduced the "jointness" of the labor income tax system, while Eissa (1995) 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 high-earning husbands significantly. Kaygusuz (2010) analyzes the effects of the same US tax reforms on the labor supply of married women with a quantitative model. Guner et al. (2012a) and Guner et al. (2012b) evaluate counterfactual policy reforms in an elaborate quantitative life-cycle model. They find that going from joint to individual taxation would increase the labor supply of married women in the US substantially.

This 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 across countries. The following section introduces the model, as well as its parametrization and calibration. Section 5 presents the results, discusses the wedges needed to replicate the extensive and intensive margin decomposition, 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

International hours worked are only available on the aggregate level, and provided by the OECD and the Conference Board. We work with three different micro data sets, namely the European Labor Force Survey, the Current Population Survey, and the German Microcensus, to construct internationally comparable hours worked for demographic subgroups. A detailed description of the data work can be found in Bick et al. (2015). There, we also contrast our resulting aggregate hours to those from the OECD and the Conference Board, and argue that also on the aggregate level our data are likely more suitable for international comparisons.

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 Eastern European countries Czech Republic, Hungary, and Poland, the Scandinavian countries Denmark, Norway, and Sweden, the Western European countries Austria, Belgium, France, Ireland, Netherlands, and the UK, and the Southern European countries Greece, Italy, Portugal, and Spain. 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 10,000 individuals aged 15 to 64 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 conducted every month. 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 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 260,000 individuals aged 15 to 64 per year.

2.1.3 German Microcensus

The German Microcensus covers a 1 percent random sample of the German population 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 around 320,000 individuals aged 15 to 64 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.

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. In Bick et al. (2015), we document that this adjustment solves the problem that the reference weeks are not spread representatively across the entire year.

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 educational 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 Taxing Wages modules, which are a key input into our quantitative analysis. 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: 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; third, because education information might be missing for the respondent or the partner. On average, around 11 percent of male and 19 percent of female observations are dropped because of these restrictions (see Table W.11 in the Online Appendix). The percentage is always larger for women than for men, because it is more likely

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

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

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 of couples and age at marriage, and variation in the number of married adults living in one household.[7] In the ELFS data for the Scandinavian countries we miss household identifiers (see Appendix A.1.2 for further information on how we deal with missing household identifiers in Scandinavia). 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 for the Scandinavian countries.

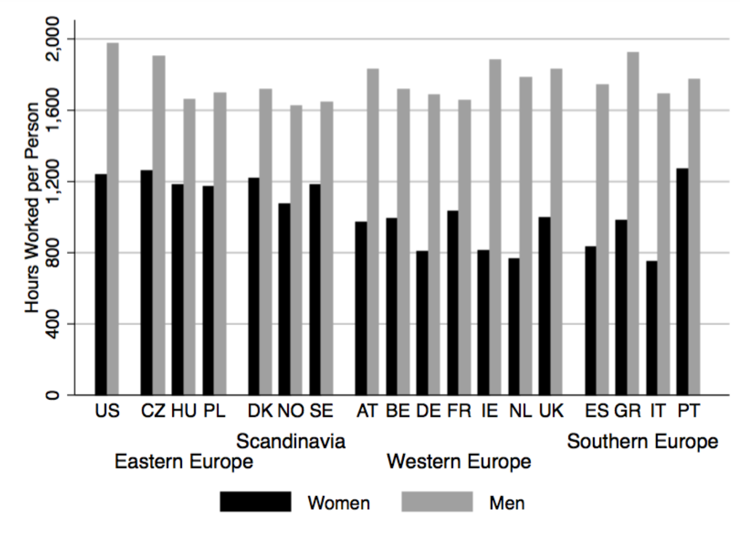
3 Hours Worked of Married Couples

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 (including divorcées) work 160 hours more than married women, and single men 280 hours less than married men. While married women are thus clearly the group with the lowest hours worked, they exhibit the largest cross-country standard deviation in mean hours worked per person: the standard deviation of hours worked of married women is more than 60 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 is up to an order of magnitude bigger. Married women contribute on average 20 percent to total hours worked, but account for 41 percent of the variance of total hours worked. Moreover, the cross-country correlation of hours worked of married men with hours worked of single men or single women is 0.79 and 0.64, respectively, while the correlation with hours worked of married women amounts only to 0.06. Thus, there is clearly something special about married women, and investigating the sources of the different behavior of married men and married women is of great interest if one wants to understand international differences in hours worked among core aged individuals.

Since from now on we focus on married couples, the issue of selection into marriage arises. While we do

[7] E.g., Poland, which has the largest share of observations dropped, exhibits a higher than average percentage of individuals married to someone younger than 25, as well as an above average number of households consisting of three or more married adults.

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



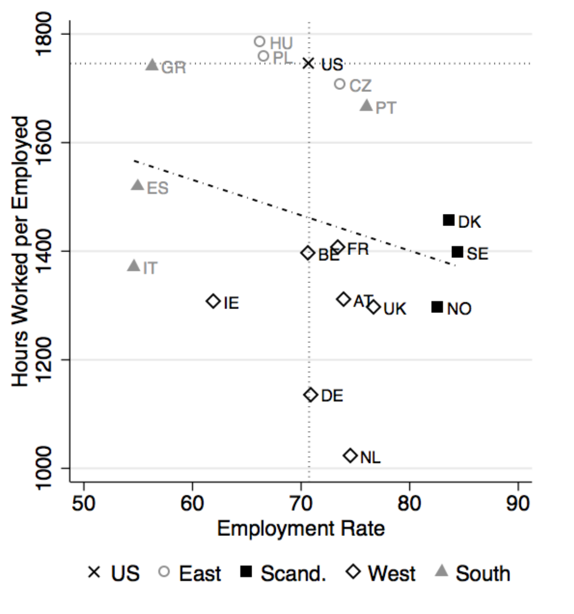
not model this selection, we report in Figure W.7 in the Online Appendix the fraction of women in the core age group who are married. It amounts on average to 64 percent, with a standard deviation of 7.5 percent. The extremes are Sweden with 48 percent of women being married, and Poland with 77 percent. 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, 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.[8] Last, Chade and Ventura (2002) and Chade and Ventura (2005) show in a quantitative equilibrium model of the marriage market for the US that the marriage rate would barely change if the US replaced the current system of joint taxation with one of separate taxation.

Figure 2 shows average hours worked of married women (dark bars) and men (light bars) aged 25 to 54 over the period 2001 to 2008 for all eighteen countries in our sample. The European countries are grouped into four regions, namely Eastern Europe (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).

Hours worked of married men are highest in the US, followed closely by Greece, the Czech Republic, and Ireland. At the lower end of the sample are Norway, Sweden, France, and Hungary. Norwegian married men work 350 hours less than, or only 82 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.

[8] 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 3 and 2 of Table 2 is 0.07, and the correlation between the marriage rate and the size of the tax structure effect shown in Table 8 is 0.09.

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



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 1235, 1200 and 1160 hours, respectively, and much lower in Western and Southern Europe with 910 and 960 hours. Portugal is a notable exception among the Southern European countries and actually features the highest hours worked of married women among all countries in our sample. Western Europe is somewhat divided with Germany, Ireland, and the Netherlands having relatively low hours worked, comparable to Italy and Spain, whereas France, Belgium, the UK, and Austria have higher hours worked, but still below the level of Scandinavian and Eastern European countries. The lowest hours worked arise in Italy with 750 hours, 485 hours less than, or only 61 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 scatterplots, 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 countries with a triangle. While Figure 2 showed that Scandinavia and Eastern Europe both exhibit high hours worked of married women, Figure 3 documents stark differences for the extensive and intensive 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 the extensive and intensive margin. Relative to the US, Western Europe features high employment rates and low hours worked per employed, while Southern Europe has low employment rates, with hours worked per employed 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 -0.26. Low hours worked per employed married woman in the data are largely driven by a high share of part-time working women, and vice versa. The cross-country correlation between hours worked per employed woman and the share of women working less than 30 hours is -0.9. We get back to the issue of part-time work in Section 5.2.2. Figure W.8 in the Online Appendix shows the decomposition into both margins for married men. There is little variation in employment rates of married men, which lie always above 90 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 hours worked per person 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 W.1 in the Online Appendix, the effect of having a school aged child relative to an older child or no children at all on employment rates and hours worked per employed is very similar across countries. The labor supply behavior of women with preschool children, on the other hand, exhibits country-specific idiosyncracies. However, since the share of women with preschool children is with on average 25.5 percent quite small, these idiosyncracies do not drive the aggregate hours worked differences. Moreover, in Section W.5 in the Online Appendix, we also show results from our model focusing exclusively on women without children, and the overall model fit is qualitatively and quantitatively very similar.

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 W.2 in the Online Appendix, we document labor supply of married women by 10 year age groups. For each country, the hours worked per employed difference to the US is similar for the three age groups. Regarding the employment rate, European married women aged 45-55 show uniformly around 5 percentage points 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, exceptionally low employment rates in Hungary and the Czech Republic stand out, caused by the effect of preschool children that we document in Online Appendix Section W.1. 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 these differences could potentially be, we take the extreme view that all unemployment differences are driven by the labor demand side. We therefore exclude all households in which the wife and/or the husband is unemployed 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 change from -3 percent to 3 percent in Eastern Europe, and decrease in absolute terms by 4 percentage points in Southern Europe, see Table W.12 in the Online Appendix.[9]

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 optimally determines male and female labor supply.[10]

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

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

Households draw a utility cost of joint work q from a distribution $\zeta(q|z)$ which depends on the husband's education level. This cost is only incurred if the wife participates in the labor market, and thus introduces an explicit extensive margin choice for women. We abstract from modeling fixed costs of work for men. As a consequence, they always optimally choose to provide positive hours. We follow this approach because male participation rates in the data in our sample countries are above 90 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. $h_f = 0$.

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

[9]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.

[10]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.

is given by

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

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

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

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

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

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

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

where * denotes the optimal hours choice given the draw of q.[11]

As usual in the literature explaining aggregate hours worked differences between Europe and the US, consumption and labor supply are assumed to be separable, and utility from consumption is logarithmic. Therefore, cross-country differences in mean wages are irrelevant, and only cross-country differences in the gender and education premia matter for labor supply decisions. 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, non-linear labor income taxes τ_l, the educational composition and matching into couples $\mu(x,z)$, male hourly wages by education $w_m(z)$, and female hourly wages by education $w_f(x)$. Last, we calibrate the two preference parameters in the utility function, α and ϕ, and the parameters governing the fixed cost distributions. Table W.3 in Online Appendix Section W.3 lists all data sources for all countries.

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

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.[12] 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 Europe and the US, amounting to 14 percentage points on average. The Czech Republic has the lowest consumption tax rate in Europe with only 14.8 percent, while they are highest in Scandinavia, where only the Norwegian one is below 30 percent.

4.2.2 Non-Linear Labor Income Taxes

The labor income tax codes come from the OECD Taxing Wages modules. These are very similar to the NBER TaxSim module for the US, but cover all OECD countries from 2001 onwards. The OECD Taxing Wages module implements the statutory labor income tax code, including employees' social security contributions and cash benefits, by marital status and number of children in the household. It calculates a household's net income for any combination of husband's and wife's earnings. Standard deductions (i.e. basic allowances, allowances for children, deduction of social security contributions) are included in these calculations, whereas individual non-standard deductions (e.g. mortgage payment deductions, deductions of child care expenses, deductions for expenses on household helpers) are not.[13] Cash benefits include those obtained for children. We compute net household earnings for a grid of wives' annual earnings with 201 grid points, ranging from 0 earnings to three times the average annual earnings in the country, and for an earnings grid with 101 grid points for men, ranging from 0 earnings to four times the average annual earnings in the country.[14] We then linearly interpolate in two dimensions to assign a net annual household income to each possible annual hours choice of husband and wife. One additional input into the tax codes are the number of children. From the micro data, we calculate the percentage of married couples with 0, 1, 2, 3, or 4+ children conditional on the educational match, and then take the weighted average over these tax burdens for any pair of hours choices. Figure 1 gives an exemplary impression of the resulting tax schedules for the US, Germany, and Sweden, holding the earnings of the other spouse fixed at one specific level. When we use the tax schedules as model inputs, we vary of course the incomes of both spouses at the same time.

While it is impossible to summarize the complex non-linear labor income tax systems in a few numbers, Table 2 presents two exemplary measures that reflect key 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,

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

[13] 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.

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

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	0.74	1.84
Hungary	23.6	29.5	18.2	20.8	18.9	0.82	1.97
Poland	18.6	29.3	32.2	10.8	18.9	0.82	3.56
Mean	*19.0*	*26.9*	*24.5*	*13.6*	*16.7*	*0.79*	*2.45*
Denmark	32.0	40.1	49.2	19.3	35.6	0.85	1.20
Norway	24.3	29.7	30.2	13.9	38.3	0.79	1.14
Sweden	32.5	33.5	28.0	13.9	37.1	0.78	1.15
Mean	*29.6*	*34.4*	*35.8*	*15.7*	*37.0*	*0.81*	*1.16*
Austria	18.6	31.4	22.6	23.8	14.7	0.76	1.83
Belgium	20.7	34.5	48.2	28.7	34.7	0.85	1.53
France	23.8	23.9	32.8	29.7	28.0	0.78	1.69
Germany	15.4	33.4	49.6	16.9	20.4	0.74	1.44
Ireland	24.2	17.2	16.7	27.7	29.7	0.77	2.46
Netherlands	21.3	31.0	34.7	28.9	24.5	0.76	1.54
United Kingdom	17.1	26.5	19.2	29.4	33.0	0.77	1.81
Mean	*20.2*	*28.3*	*32.0*	*26.4*	*26.4*	*0.78*	*1.76*
Greece	14.9	25.9	16.0	34.4	21.4	0.76	2.05
Italy	22.1	27.8	28.3	45.8	12.1	0.68	2.47
Portugal	19.0	18.4	25.4	72.6	13.8	0.81	2.84
Spain	15.9	17.3	18.5	50.8	27.8	0.66	2.54
Mean	*18.0*	*22.4*	*22.1*	*50.9*	*18.8*	*0.73*	*2.48*
United States	7.4	21.4	29.1	7.8	45.7	0.79	2.20

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. $[\tau_l(w_m h_m^{US}, w_f h_f^{US}) - \tau_l(w_m h_m^{US}, 0)]/[w_f h_f^{US}]$. 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).

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 the interaction between progressivity and jointness of taxation.[15] 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 lower than male US hours, and mean wages are always lower 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 strictly 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. Let us stress again that these two tax rates shown here are only indicative values, and that we exploit the full non-linearity of the labor income tax code in our quantitative analysis.

The US average tax rate as calculated in column 2 amounts to 21.4 percent, whereas the corresponding Danish married couple would have to pay an average tax rate of 40.1 percent, and the Irish couple a tax rate of only 17.2 percent. 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 percent in the US, peaking at 49.6 percent in Germany, a country with high progressivity and joint taxation of married couples, and 48.2 percent 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 49.2 percent, while Norway and Sweden have levels similar to the US. Taking the difference between columns 3 and 2 as a suggestive measure of the jointness of taxation, Eastern Europe, Scandinavia, and Southern Europe feature largely systems of separate taxation of married couples, with Denmark and Portugal featuring elements of joint taxation. For Western Europe, the picture is mixed, ranging from clear separate taxation in Austria and the UK to the strongest impact of joint taxation in Belgium and Germany. 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.[16] 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

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

[16] See appendix A.1.2 for details how we proceed for Scandinavia. Low education is defined as primary and lower secondary education (ISCED categories 0 to 2), medium education as upper secondary and non-tertiary post-secondary education (ISCED categories 3 and 4), and high education as any tertiary education (ISCED categories 5 and above). In the US, low education is defined by having completed at most 11th grade of high school; medium education by having completed the 12th grade of high school, having a high school diploma, or attended some college; and high education by having at least a college degree.

average half of the married women exhibit low education, while in Eastern Europe and the US only around 10 percent do. The fraction of highly educated married women is the 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 thirds of married women have medium education. Online Appendix Table W.13 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 one order of magnitude smaller than the ELFS. Online Appendix Section W.3 provides more information on the EU-SILC, shows that data in the ELFS and the EU-SILC are similar along several relevant dimensions, and provides further information on how we construct wages. We 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 monthly individual earnings by monthly hours. We construct monthly hours by multiplying usual weekly hours with 52 minus vacation/public holiday weeks from external data sources, divided by 12. 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 in 2004 for 11 out of our 17 European countries, and in 2005 for the remaining 6 countries; we extrapolate wages for the missing years based on OECD growth rates of mean wages.

For married women, the issue of self-selection into employment arises. If high ability women of each education group are more likely to join the labor force, then observed mean wages overestimate the mean of offered wages, see e.g. Olivetti and Petrongolo (2008). We therefore apply a simple two-stage Heckman procedure to impute wages of non-working women. The exclusion restrictions are that the income of the husband as well as the presence of children do not directly influence the wage of a woman, see e.g. Mulligan and Rubinstein (2008).

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.[17] The average gender wage gap by region is quite similar in Europe and the US, with the exception of Southern Europe, which features on average a larger gender wage gap. Southern Europe also exhibits the largest

[17] Corresponding education premia for men are shown in Table W.13 in the Online Appendix and are largely similar.

degree of heterogeneity, with a ratio of female to male wages of 0.66 in Spain, the lowest among the sample countries, and 0.81 in Portugal. The overall smallest gender wage gaps arise in Belgium and Denmark, with a ratio of 0.85. 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.[18]

4.3 Redistribution of Government Revenues

The government redistributes a fraction $\lambda \in [0,1]$ of all government revenues back to the households in a lump-sum fashion. In the benchmark calibration, we follow Rogerson (2008), Ohanian et al. (2008), and Ragan (2013) and assume full redistribution of government revenues and thus set $\lambda = 1$. In Online Appendix Section W.5, we show results from two alternative specifications with either no redistribution of government revenues (i.e setting $\lambda = 0$), or from setting λ equal to 1 minus twice the share of military expenditures from all government expenditures, similar to the specification used by Prescott (2004).

4.4 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:

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

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' own education levels $x \in \{low, medium, high\}$. For given preference parameters α and ϕ, and conditional on being married to a type z husband, the three different education levels x and implied wages generate three different threshold levels $\bar{q}(x,z)$ at which a woman of type x is indifferent between working and not working. Assume for simplicity that all type z husbands work the same amount of hours. Women with more education, i.e. a higher wage, will have a higher threshold q, and therefore a higher labor force participation rate, for any

[18] We conduct two robustness checks with respect to wages, which are shown in Online Appendix Section W.5. Table W.9d in the Online Appendix shows results if no Heckman-correction is applied on female wages and only observed wages are used. On average, results are very similar, except for Southern Europe. Table W.9e in the Online Appendix shows results if we allow for wage heterogeneity within each education group. Our country-specific estimate of heterogeneity is the standard deviation of the residuals from a regression of log male wages on year and education dummies. Results are qualitatively identical, but quantitatively the explained difference between Europe and the US is smaller.

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

Table 3: Data Targets and Calibrated Preference Parameters

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

given distribution of q. This pattern is also prevalent in the data, i.e. conditional on the husband's education, the female labor force participation rate is increasing in the woman's own education. The parameters k_z and θ_z are then selected to ensure that the mass of women below these thresholds corresponds to the empirically observed female participation rates by female education conditional on the husband's education.

Table 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 2.6 percent lower in the model than in the data, and hours worked per married man are 1.7 percent higher than in the data. The average female employment rates by the husband's education are matched almost perfectly. Employment rates of low and high educated women are however slightly higher in the model than in the data, and those of medium educated women are slightly lower.

4.5 Model Fit

Even though we only target average hours by gender, the model also replicates the gradients by education fairly well, see Table W.14 in the Online Appendix. The model is only somewhat off for the low educated:

hours per employed of low educated women are underpredicted by 9 percent, and hours per person of low educated men are overpredicted by 18 percent. This is due to high non-employment rates among low educated men suppressing their hours worked per person, 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 parameters fixed. This exercise is in the spirit of Kaygusuz (2010), who analyzes only the time period from 1980 to 1990, which we extend through 2008. The model correctly predicts hardly any change in hours worked of married men over the period of three decades. For married women, the model captures both the increase in the employment rate and in hours worked per employed from 1979 to 2008 almost perfectly, with some deviations of the employment rates in model and data in the 1990s. Details are shown in Online Appendix Section W.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. 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 a married couple with the one of two singles living together in one household. We lay a proportional tax/subsidy on the households of two singles, such that tax revenues are identical to the ones collected from married households. 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 when going from joint to separate taxation, 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 (i.e. consumption tax rates and labor income tax systems), 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-country predictions of hours worked per married person, then break them down by gender, and afterwards show extensive and intensive margin predictions for married women. 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 couples' hours worked. We further analyze the effects of labor income taxes by decomposing them into differences in average tax rates 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 labor income taxes, while Portugal is an outlier within Southern Europe in terms of actual hours worked of married

Table 4: Hours Worked per Married Person (% difference to the US)

Region	Total Data	Total Model	Men Data	Men Model	Women Data	Women Model
Eastern Europe	−7.8	−6.2	−11.1	−5.1	−2.6	−8.0
Scandinavia	−13.9	−9.7	−17.1	−10.3	−8.9	−8.6
Western Europe	−16.5	−11.3	−10.3	−7.4	−26.4	−17.9
Southern Europe	−17.7	−8.5	−9.4	−2.0	−30.9	−19.5
Europe	−14.7	−9.5	−11.2	−6.2	−20.2	−15.0

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

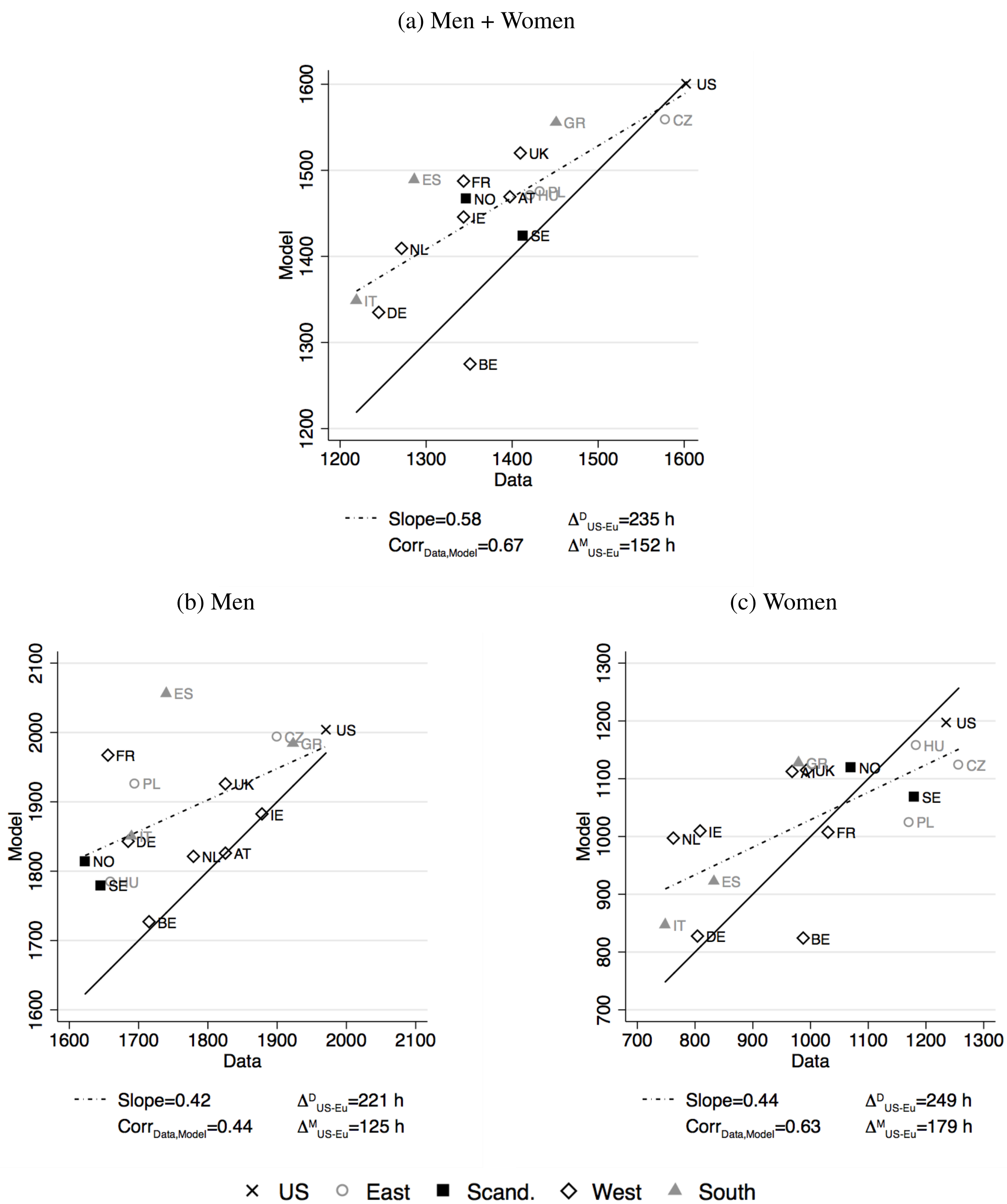
Table 4 shows in the first column the percent difference in married couples' hours worked per person (i.e. adding up hours of husband and wife) between each country group and the US in the data, and in the second column the model predicted percent differences. While the table only presents country-group averages, country-specific results presented in levels 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 US-Europe difference is smallest for Eastern Europe, followed by Scandinavia, and then followed by Western Europe. Southern Europe is the exception, because the difference to the US is largest in the data but not in the model. While through the lens of the model differences in taxes, wages, and the educational composition can account on average for more than three quarters of the hours worked difference between the US and Eastern Europe, and for slightly more than two thirds of the hours worked differences between the US and Scandinavia and Western Europe, respectively, they can only account for one half of the difference for Southern Europe. The correlation between model and data is 0.67.

To get an impression of the goodness of fit, we replicate the exercise in Prescott (2004), who analyzes the role of *linear* taxes for *aggregate* hours worked (i.e. hours of all individuals aged 15 to 64). Specifically, we take the same neo-classical growth model, calibrate it to match US aggregate hours, and then predict hours for the European countries and compare them to hours of all individuals aged 15 to 64 in the data. The Prescott exercise with *linear* taxes can only explain 53 percent of the US-Europe difference in aggregate hours, while our exercise with *non-linear* taxes explains two thirds of the US-Europe hours difference for for married couples aged 25 to 54. Moreover, the correlation between model and data in the Prescott exercise is with 0.28 less than half the correlation in our exercise. This underlines the good fit of our model.

The next columns of Table 4 break the previous results down by gender. The corresponding country-

[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: Hours Worked per Married Person



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 –5 and –10 percent, respectively. For Western Europe, the fit is best with a model predicted difference of 7 percent, compared to 10 percent in the data. For Southern Europe, the model explains only one fifth of the difference to the US. This is mostly driven by Spain, where male hours worked are 12 percent lower than in the US, while the model predicts 3 percent higher hours than in the US. Focusing on individual countries, for 7 countries the differences between prediction and data amount to 3 percentage points or less. The fit is worst for France: hours of married men are 16 percent lower than in the US in the data, 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 0.44.

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 –3 percent in the data, while the model predicts a small, but somewhat larger difference (in absolute terms) of –8 percent. For Scandinavia, the model generates an almost perfect average fit with a difference to the US of –9 percent in the data and 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 even correctly replicates larger differences in hours worked relative to the US for married men than for married women.

For Western Europe the model explains two thirds of the observed large difference to the US in married women's hours worked of –26 percent in the data. The fit is particular good for France and Germany, with a deviation between model and data of only 0.6 and 3.9 percentage points, respectively. This excellent fit is quite remarkable, given that married women work 17 percent (200 hours) and 35 percent (430 hours) less, respectively, in these two countries than in the US. For Ireland and the UK, the model explains only around two fifths of the observed differences to the US. The worst relative fit is for Austria, the only country in Western 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. For Belgium, the model overpredicts the difference.

For Southern Europe, the model explains slightly less than two thirds of the difference in hours worked relative to the US, predicting 20 percent lower hours than in the US, while in the data Southern European married women work 31 percent less hours than their US counterparts. This is largely driven by Greece, for which only one quarter of the difference to the US can be explained by the model.

Summarizing, through the lens of the model, differences in taxes, wages, and the educational composition can account for a substantial fraction of the hours worked differences of married women in Eastern Europe, Scandinavia, Western Europe, and Southern Europe to the US. The correlation of married women's hours worked in model and data is with 0.63 higher than for married men. Also, the model is able to repli-

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

	ER		HWE	
Region	**Data**	**Model**	**Data**	**Model**
Eastern Europe	−1.9	−2.8	0.2	−4.0
Scandinavia	12.8	−3.3	−22.9	−4.2
Western Europe	1.1	−7.7	−27.4	−8.0
Southern Europe	−15.4	−11.4	−11.6	−4.2
Europe	−1.3	−6.9	−18.1	−5.9

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

cate the low correlation of male and female hours worked per person in the data, which amounts to 0.15 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 analyze 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, while 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 is the only region that shows a similar decomposition into both margins as the US. For Southern Europe, the model reproduces three quarters of the employment rate difference, but only one third of the intensive margin difference. The decomposition into the extensive and intensive margin does not work well for Scandinavia and Western Europe. The respective employment rates in the data exceed those in the US by 13 percentage points and 1 percentage point. By contrast, the model predicts lower employment rates by 3 and 8 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 8 percent, respectively. Consistent with the data, Western Europe has the largest model predicted differences to the US in female hours worked per employed among the European regions, but the model accounts only for one third of the actual differences. Hence, for Scandinavia and Western Europe the model has difficulties explaining the decomposition of hours worked per married woman into an intensive and an extensive margin.

It is not entirely surprising that taxes, wages, and the educational composition have difficulties explaining the vastly different decompositions into the intensive and extensive margin in Scandinavia and Western Europe on the one hand, and Eastern and Southern Europe and the US on the other hand (compare Figure

3). The model factors 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 0.44 between married women's employment rates and hours worked per employed, while in the data the correlation is -0.37.[21]

5.2.2 Wedges and the Margins of Labor Supply

Many institutional, policy, and regulatory factors other than those featured in our model (i.e. taxes, wages, and the educational composition) affect hours worked, and especially the decisions whether to work or not, and how many hours to work conditional on working. In order to get some insights into these factors, we follow the approach pioneered by Chari et al. (2007) and Restuccia and Rogerson (2008), and introduce two country-specific wedges into the model, $\hat{\tau}_E$ and $\hat{\tau}_{h_f}$, in order to match the country-specific female employment rates and hours worked per employed. The wedges pick up any differences between the model's predictions and data that are not driven by taxes, wages, and the educational composition. They could potentially capture very different factors, such as restrictions on hours or employment from the labor demand side, child care, or maternal leave. Below we correlate these non-modeled factors with our wedges and explore which of them may be particularly important for understanding the discrepancy between the model and the data in the decomposition into both margins. Moreover, and especially important, we investigate in how far the quantitative effects on hours worked of our model factors differ between our benchmark model, which does not replicate the negative cross-country correlation between the extensive and intensive margin, and the model with wedges, which by construction exactly replicates this negative correlation in the data. We delegate this second exercise to Section 5.3.4, where we investigate in more detail the role of each model factor in explaining hours worked. For now, we focus on the wedges themselves and how they correlate with country-specific factors outside the model.

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

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

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 hours worked. It only affects the extensive margin choice of married women. $\hat{\tau}_{h_f}$, on the other hand, represents an implicit subsidy or tax on female hours worked. It directly affects the intensive margin choice by entering the first order condition for female hours conditional on working: a negative $\hat{\tau}_{h_f}$ increases the disutility of work for every hour of work. Thereby, it also has an indirect effect on the extensive margin choice, by changing the utility of working a certain amount of hours vis-à-vis the utility of not working. We calibrate both wedges to match the country-specific employment rates and hours worked

[21] The correlation in the data including Denmark and Portugal is -0.26, which is the number we report in Section 3.

per employed married woman.

By construction, for countries for which the model predicts too high (low) hours or employment rates relative to the data, the respective wedge is negative (positive). The larger the difference between the predictions of the model without wedges and the data, the larger is the absolute value of the calibrated wedge, see Figure W.9 in the Online Appendix. Moreover, since the model predicts a positive correlation between employment rates and hours worked per employed rather than the observed negative correlation, the wedges are highly negatively correlated with a value of -0.85. This indicates that there might be common underlying factors that drive the two margins in opposite directions.

To understand what such factors could be, we follow the idea in Ohanian et al. (2008) and correlate the wedges with country-specific variables that are potentially important for women's labor supply choices and the decomposition into both margins. Table W.15 in the Online Appendix shows the country-specific wedges, as well as the country-specific values of all variables that we correlate with the wedges. The variable that features one of the strongest correlations with both wedges is an OECD country ranking of part-time generosity. A higher part-time generosity rank implies more extensive rights of the worker to demand part-time work, fewer grounds for refusal from part of the employer, the right to go back to full-time work, as well as part-time legislation being in place for a longer time period. As Table 6 shows, the OECD part-time generosity rank exhibits a positive correlation of 0.46 with the hours wedge, and a negative correlation of -0.60 with the employment wedge. Part-time generosity also intuitively explains why both margins could be negatively correlated in the data. Suppose that there exist a group of women who find it optimal to work but supply relatively few hours. If part-time jobs are not available in a country, these women might prefer not working over working full-time, resulting in a low average employment rate and high average hours worked per employed. In a country with more generous rights to work part-time, these women will enter the work force at the desired low hours, thereby increasing average employment rates and decreasing average hours worked per employed. To give a concrete example, the US, our benchmark country, has the lowest OECD part-time generosity rank, while Sweden has the highest rank and has one of the highest employment and hours wedges (in absolute values). Thus, the differential regulation of part-time work across countries is a promising feature that could potentially explain the negative correlation of hours worked per employed and employment rates that we see in the data. Incorporating part-time regulation explicitly 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; second, modeling part-time regulation without assuming what one wants to find by imposing a different fixed utility cost distribution of working part-time 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.

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

Table 6: Correlation of Wedges with Different Factors

Policies	$\hat{\tau}_h$	$\hat{\tau}_e$
Part-time generosity rank	0.46	−0.60
Annual leave + public holidays	−0.33	0.37
Maternity leave: paid weeks	0.29	−0.19
Avg. pay during maternity leave	0.10	−0.05
Net child care costs (% of avg. earnings)	−0.31	0.13
Public child care expenditure (% of GDP)	−0.34	0.64
% of preschool child. in informal care	0.04	−0.33
Avg. hours of informal care for preschool child.	0.39	−0.47
Divorce rate (per 1000 persons per year)	0.14	0.13

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

care), and divorce rates. 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. Somewhat surprisingly, the same holds true for a measure of child care cost.[22] However, the correlation of the employment wedge with public child care expenditure over GDP is much higher. Public child care expenditure intuitively has the potential to drive both margins in the opposite direction, since it might draw women into the labor market who prefer to work low hours. Thus, child care policies might be an additional potential driving force of the different margin decompositions, but other factors than net costs, for example quality or availability of slots, seem to play a crucial role. Moreover, informal child care arrangements also play an important role in practice. According to the OECD, on average across our sample countries 28 percent of preschool children receive some form of informal child care, with the average hours conditional on receiving this form of child care being 15. Informal child care is strongly negatively correlated with public expenditures on child care.[23] This suggests that private arrangements serve at least to some degree as substitutes for government child care policies, see e.g. Bick (2016) for the case of Germany. As this discussion shows, similarly to part-time regulation, modeling and measuring child care policies across countries is far from trivial, and we abstain from doing it in this paper. Last, we correlate the wedges with divorce rates, as suggested by Chakraborty et al. (2015), but correlations are very low.

[22] This is especially surprising in light of the findings of Domeij and Klein (2013) and Guner et al. (2014) that subsidizing child care is an effective policy if the government wants to increase female labor supply.

[23] This correlation with public expenditures on child care is -0.55 for the fraction of children receiving informal care, and -0.40 for average hours in informal care.

We now proceed by analyzing the effects of taxes, wages, and the educational composition 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 on married women's hours worked per person are qualitatively and quantitatively very similar whether country-specific wedges are included into the model or not.

5.3 Decomposition Analysis

To understand the relative importance of wages, taxes, and the educational composition in explaining cross-country differences in hours worked of married men and women, 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 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. Afterwards, 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 of married women's hours worked within Europe. Last, in Subsection 5.3.4 we reintroduce wedges into the model and show that the effects on hours worked of our model factors are quantitatively very similar between our benchmark model, which does not replicate the negative cross-country correlation between the extensive and intensive margin, and the model with wedges, which by construction exactly replicates this negative correlation.

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 7. Columns "Data" and "Model" in Table 7 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 the respective country, 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 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: for each combination of husband-wife hours choices, we calculate the country-specific tax rate using the US gender-specific education premia and the country-specific mean wage, and then apply this tax rate to the US gross earnings implied by the same husband-wife hours choices. Similarly, when applying the gender-education-specific wages from a European country, we leave the mean wage at the US level such that the US labor income tax

Table 7: Decomposing the Effects of Different Model Inputs on Labor Supply (% difference to the US)

(a) Male Hours Worked per Person

Region	**Data**	**Model**	τ_c	τ_l	$\mu(x,z)$	w
Eastern Europe	−11.1	−5.1	−2.5	−3.8	2.4	−2.0
Scandinavia	−17.1	−10.3	−4.5	−6.6	0.3	1.2
Western Europe	−10.3	−7.4	−2.8	−5.5	2.2	−0.3
Southern Europe	−9.4	−2.0	−2.1	−2.7	2.8	0.2
Europe	−11.2	−6.2	−2.8	−4.8	2.1	−0.3

(b) Female Hours Worked per Person

Region	**Data**	**Model**	τ_c	τ_l	$\mu(x,z)$	w
Eastern Europe	−2.6	−8.0	−7.2	1.8	−2.8	−2.6
Scandinavia	−8.9	−8.6	−12.9	0.4	0.5	2.5
Western Europe	−26.4	−17.9	−8.0	−7.4	−7.3	2.0
Southern Europe	−30.9	−19.5	−6.5	7.4	−11.2	−7.4
Europe	−20.2	−15.0	−8.2	−1.6	−6.1	−0.7

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.

schedule can be meaningfully applied. Details of these procedures are in Appendix A.2. Panel (a) of Table 7 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 work 13 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.[24] 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 consumption tax is set country-specific is with 0.99 extremely high. Thus, the consumption tax affects male and female hours in exactly the same way qualitatively.

While for women consumption taxes are on average the main factor driving the lower European hours, labor income taxes are the main factor for explaining low European male hours: labor income taxes alone

[24] 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.

predict 2 to 7 percent lower hours of married men in Europe than in the US. Their disincentive effect on male hours worked is thus around 70 percent larger than 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 shown in columns 1 and 2 of 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, but *higher* ones for Eastern Europe, Scandinavia, and Southern Europe. Even for Western Europe, the disincentive effect of labor income taxes is smaller than the disincentive effect of consumption taxes. The low correlation of 0.46 of the labor income tax effect between men and women stands in striking contrast to the essentially perfect correlation of the consumption tax effect. We analyze the reason for this low correlation 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. Second, 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 average female employment rate. Since husbands of non-working women work higher hours than those of working women, this compositional change increases average male hours, despite the fact that the less favorable educational composition also applies to men. 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 female hours than in the US, which amounts to one quarter of the difference in the data. 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 -0.64 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 7. It is largest for Southern European married women: wages alone predict 7 percent lower hours worked there than in the US. For Scandinavian 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 contribute the most to explaining the lower hours worked of married men in Europe than in the US. By contrast, consumption taxes, which work qualitatively very similarly 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, except for Southern European 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 into two components: first, the average tax rate, i.e. the level of the tax schedule, and second, the progressivity of the tax schedule along with the tax treatment of married couples, which together define how the marginal tax rate of each spouse changes with the own and the spousal income. We 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 on the household such that government revenues are left unchanged at the US level.[25] 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 8 shows the resulting decomposition of the labor income tax effect into the tax structure and the average tax rate. 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 and Southern Europe. There are two reasons for the strong negative tax structure effect: first, taxes are generally more progressive in Europe than in the US. Secondly, the degree of joint taxation is more pronounced in the US than in most European countries, reducing the marginal tax rate of American married men compared to most European married men. On top of the tax structure effect comes an additional disincentive effect through 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, except for Western Europe. 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 and women, with again larger effects for women due to the higher implied female

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

Table 8: Effects of Average Tax Rate vs. Tax Structure on Labor Supply (% difference to the US)

(a) Male Hours Worked per Person

Region	**Data**	τ_l	**Structure** τ_l	**Average** τ_l
Eastern Europe	−11.1	−3.8	−5.2	1.3
Scandinavia	−17.1	−6.6	−3.8	−2.8
Western Europe	−10.3	−5.5	−4.1	−1.4
Southern Europe	−9.4	−2.7	−4.1	1.3
Europe	−11.2	−4.8	−4.3	−0.5

(b) Female Hours Worked per Person

Region	**Data**	τ_l	**Structure** τ_l	**Average** τ_l
Eastern Europe	−2.6	1.8	0.2	1.6
Scandinavia	−8.9	0.4	7.6	−7.3
Western Europe	−26.4	−7.4	−3.3	−4.2
Southern Europe	−30.9	7.4	4.5	2.9
Europe	−20.2	−1.6	0.4	−2.0

Note: Columns 4 and 5 add up to Column 3. Column 3 corresponds to column 4 in Table 7.

elasticity. The correlation between male and female hours if only the average tax rate is set country-specific is exactly 1, and resembles the one if only consumption taxes are set country-specific. The higher average tax rates predict 7 and 4 percent lower hours worked by married women in Scandinavia and Western Europe, respectively, with small positive effects for Eastern and Southern Europe. Thus, the disincentive effects of average labor income tax rates and of consumption taxes are qualitatively exactly the same for men and women, but always larger for women than for men. Quantitatively, the disincentive effects of consumption taxes are larger than those of the average labor income tax rates for both men and women, which is in line with the large consumption tax rate differences between Europe and the US shown in Table 2.

In contrast to the average tax rate effect, 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 3.3 percent lower hours worked per married woman than in the US, for Eastern Europe 0.2 percent higher hours, and for Southern Europe and Scandinavia 4.5 and 7.6 percent higher hours. In fact, there are only five 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 Italy and the Czech Republic. Western Europe is the region that exhibits the largest heterogeneity in the tax structure effect: it ranges from predicting 13.3 and 11.0 percent lower hours than in the US for Germany and Belgium to 5.1 percent higher hours than in the US for the UK. The average tax rate effect by contrast is negative in all Western European countries but Ireland and Austria.

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, the wife faces a relatively high marginal tax rate, which is increasing in her husband's income. Table 2 presents as an exemplary measure the average marginal income tax rate if a woman starts working the same hours as the average US woman. This rate is close to 50 percent in Germany, but only 29 percent 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. But for the majority of European countries the tax structure predicts higher hours worked of married women because the tax systems there are closer to separate taxation than the US system. In contrast to the almost perfect positive correlation of male and female hours if only the consumption tax or only the average labor income tax are set country-specific, this correlation is only 0.34 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 structure 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.[26] Incorporating the tax structure is particularly important to generate the comparatively high hours worked of married women in Scandinavia. 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 the large hours worked differences in the data.

5.3.3 Married Women: Variation Within Europe

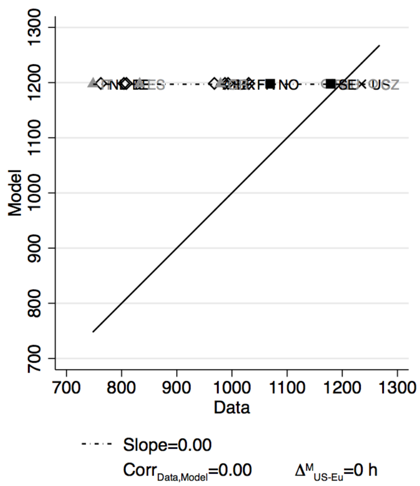
So far we present the results only for the country group averages and report them as differences to the US for men and women. We now shift our attention to country-specific results and analyze the within Europe variation in greater detail, focusing on women only. Figure 5 presents scatterplots of hours 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 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

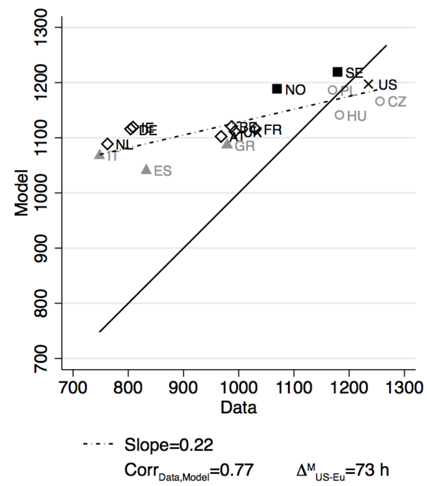
[26] Online Appendix Section W.6 compares the model performance to a model relying on linear tax rates as calculated by Prescott (2004), McDaniel (2011) and Ragan (2013) as inputs, and confirms that a model with linear taxes cannot explain hours worked of married women in Europe well.

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

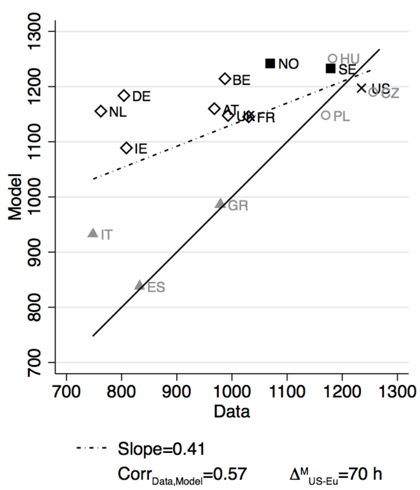
(a) US Environment



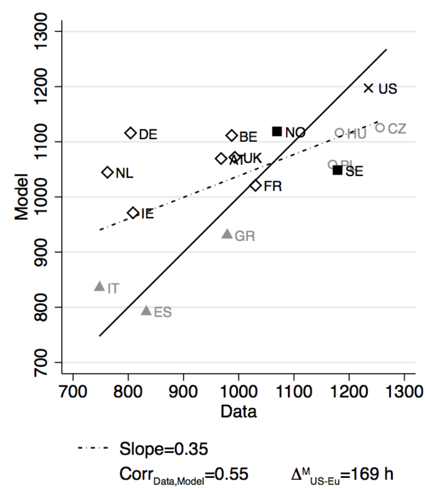
(b) + Educational Composition



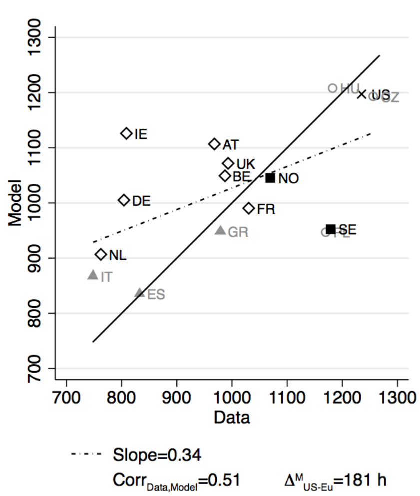
(c) + Wages



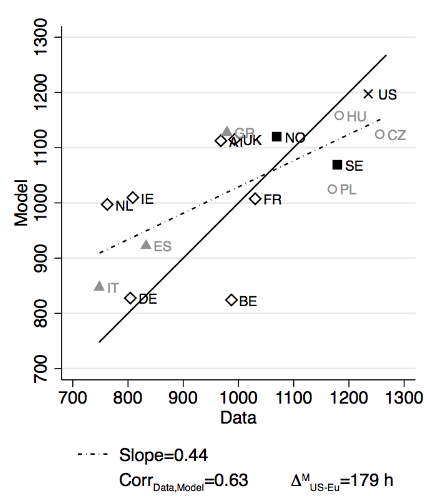
(d) + Consumption Tax



(e) + Average Tax Rate



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



in data and model is very high with 0.77, with a slope of 0.22 if the model output is regressed on the data.[27] Applying only the country-specific educational composition plus matching into couples predicts 73 lower annual 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.[28] Adding the country-specific education-gender-wage premia leaves the average hours difference to the US almost unchanged, but improves the regression slope, which rises to 0.41. As one can see in the figure, this is driven by the negative effect of wages on predicted hours in the Southern European countries, while for most Western European countries the model fit worsens. The correlation between data and model drops from 0.77 to 0.57.

Panel (d) then adds on top of this the country-specific consumption tax rates. Consumption taxes shift predicted hours in Europe substantially further downward, by on average 100 hours. They thus explain 40 percent of the average US-Europe hours difference, and are the most important driver of low hours of European married women. However, they do not explain any additional variation *within* Europe: the slope of the regression line of model hours on data hours falls slightly, as does the correlation between hours in model and data.

Panel (e) now adds the average labor income tax rate as a country-specific model input. Average European hours fall by 12 hours compared to panel (d), 5 percent of the total hours difference. While average tax rates overall predict slightly lower hours in Europe than in the US, predicted hours in Ireland increase substantially as well as to a smaller degree in the Czech Republic, Hungary, and Spain. The slope of the regression line and the correlation both again decrease very slightly. Thus, adding the average tax rates only modestly helps in explaining the average US-European hours difference.

Panel (f) finally adds the tax structure of the labor income tax code and presents the full country-specific results, already shown in Panel (c) of Figure 4. Predicted hours in Europe slightly increase by 2 hours 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 improve the fit within Europe. The slope increases from 0.34 to 0.44, and the correlation from 0.51 to 0.63. The fit improves substantially for Germany, Sweden, Poland, and Ireland. A few countries move however further away from the hours in the data at this step. Specifically, one can see the positive effect of the tax structure on Greece and Spain, which makes it harder to predict low hours in these two countries. The strong impact of joint taxation in Belgium and the Czech Republic, by contrast, decreases hours there too much.

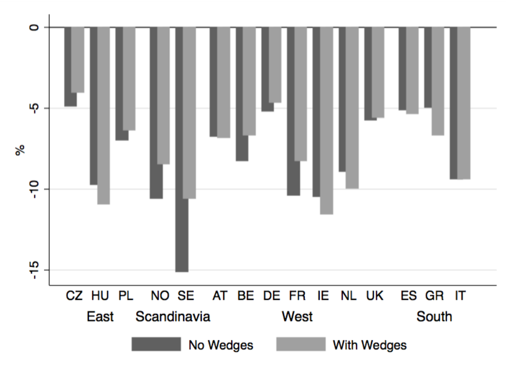
Summarizing, Figure 5 confirms that consumption taxes are the major driving force of *on average* low hours worked of married women in Europe. Analyzing average labor income tax rates alone would bring no further success in explaining female hours in Europe. Taking the non-linearity of the tax codes seriously and dealing with the jointness of tax systems is crucial to explain the *variation* within Europe. The educational

[27]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 (because our calibration features more moments than parameters), the US is slightly below the 45 degree line. As before, in the regression and the calculation of the correlation, we exclude the US.

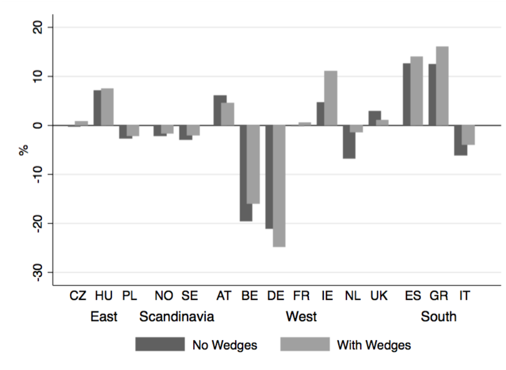
[28]Given log utility of consumption, differences in mean wages do not affect choices at all, and only the education-gender-wage premia matter.

Figure 6: Decomposition With and Without Wedges

(a) Consumption Tax



(b) Labor Income Tax



Note: This figure shows predicted hours worked per person differences of married women to the US if only one model component is set country-specific in model with and without wedges.

composition of the workforce adds further to explaining this variation, and wages are a substantial factor contributing to lower hours in Southern Europe.

5.3.4 Decomposition With and Without Wedges

One might be worried that the effects of taxes, wages, and the educational composition on hours worked would be different if both margins of labor supply were matched perfectly across countries, i.e. in the model with wedges. To address these concerns, we repeat the decomposition exercise from Table 7 in the model with country-specific wedges, and show in Figure 6 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. The country-specific values in the model without wedges are exactly the ones that generate the country-group average effects shown in Table 7. 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 models with or without wedges, both qualitatively and quantitatively. The largest differences arise for Sweden in the case of the consumption tax, and for Ireland and the Netherlands in case of the labor income tax. Overall, these figures indicate that our decomposition analysis delivers quantitatively meaningful insights even if we do not explicitly model the factors that lead to the different margin decompositions of married women's hours worked across countries. Figure W.10 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.

Table 9: Decomposing Hours Worked Differences for Denmark and Portugal (% difference to the US)

	Data	Model	τ_c	τ_l	$\mu(x,z)$	w	Tax Str. τ_l	Tax Avg. τ_l
Denmark								
HWP_m	−12.8	−15.7	−5.3	−12.1	1.0	−0.4	−3.7	−8.4
HWP_f	−1.3	−26.3	−14.3	−18.5	−2.9	7.3	0.7	−19.2
Portugal								
HWP_m	−10.2	−1.0	−2.5	0.4	4.4	−2.7	−2.2	2.6
HWP_f	2.6	−16.6	−7.1	7.0	−18.0	−8.6	0.9	6.1

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 –26 percent (see Table 9). The failure of the model to replicate the labor supply behavior of Danish married women comes from the 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 has 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 the tax structure predicts essentially the same hours in Denmark as in the US, it predicts significantly higher hours than in the US for the other Scandinavian countries. Coupled with the much stronger negative effect of the high Danish average tax rate, this leads to a predicted hours difference of –19 percent based on labor 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 17 percent lower hours worked of married women in Portugal than in the US.

6 Conclusion

Relying on three micro data sets, we document average hours worked of married couples for a sample of 17 European countries and the US over the time period 2001 to 2008. We find that hours worked vary significantly across countries, and 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. Eastern European and Scandinavian married women work only 3 and 9 percent fewer hours than US married women, but Western and Southern European women work 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, calibrated to the US and holding preferences constant across countries. The model features can account for 55 percent of the US-Europe gap in male hours worked, and 74 percent of the US-Europe gap in female hours worked, and also replicate the cross-country variation within Europe well: the cross-country correlation between hours worked in the model and the data is 0.44 for men and 0.63 for women. Specifically, the model is able to replicate the small hours worked per married women differences between the US and Eastern Europe and Scandinavia, and the much larger differences between the US and Western and Southern Europe.

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 in 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, which comprise both the tax treatment of married couples and the progressivity of the tax code. The tax treatment of married couples across countries ranges from joint to separate taxation, with most 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. Differences in the educational composition across countries further help in explaining low female hours worked in Southern and Western Europe. For married men, labor income taxes are the main driver of low European hours worked, aided by consumption taxes.

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

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Duernecker, G. and B. Herrendorf (2015). On the Allocation of Time - A Quantitative Analysis of the US and France. Working Paper.

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

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

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

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

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

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

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

Guner, N., R. Kaygusuz, and G. Ventura (2014). Childcare Subsidies and Household Labor Supply. Working Paper.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Wallenius, J. (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 Taxing Wages module does not produce the corresponding tax rates. We exclude households in the ELFS in 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 negligible number of observations. A detailed description of all issues related to the construction of the data sets can be found in Bick et al. (2015).

A.1.2 Dealing with Missing Household Identifiers in Scandinavia

For the three Scandinavian countries Denmark, Norway, and Sweden, the ELFS does not provide household identifiers in the anonymized data set available to researchers. Therefore, we neither know the age of the husband, nor the education of the husband, nor how many married adults and how many children live in the household. This has three consequences for our analysis based on the ELFS. First, in terms of our sample of married women, all Scandinavian statistics reported in the paper are based on all married women aged 25 to 54 (except if we do not observe their education), rather than on the subset of married women aged 25 to 54 whose husband is aged 25 to 54 as well, for whom we know the education of the husband, and with whom no more than one other married adult lives in the same household. Second, when we analyze hours worked by presence of children in the data, we exclude Scandinavian countries. Third, the simulation of the model uses as inputs the matching into couples based on both spouses' education, as well as the fraction of couples with zero to four children by educational match. For the Scandinavian countries, we calculate the corresponding statistics in the EU-SILC and use them as an input for solving the model. Online Appendix Section W.3 shows that data in the ELFS and the EU-SILC are similar along several relevant dimensions, and that within the EU-SILC relevant data for married women in Scandinavia are similar whether focusing on all married women aged 25 to 54 or on the subset of married women aged 25 to 54 who match our more stringent sample selection criteria.

A.2 Model Decomposition Analysis

The fourth column (τ_l) of Table 7 shows the model predicted hours if the labor income tax system is set country-specific, while gross household income y_{hh} in Equation (4) remains at the US level, i.e.

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

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

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

in the household optimization problem.

We proceed in a similar fashion when we analyze the effects of country-specific gender-education premia in column 6 of Table 7. 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.