The Marriage Earnings Gap *

Elena Herold† Luisa Wallossek‡

March 2025

Latest version: click here.

Abstract

What happens to earnings upon marriage? Linking administrative and survey data from Germany, we show that there is a *marriage earnings gap*. Even after accounting for the child penalty, women's earnings drop by 20% after marriage. We show that the marriage earnings gap results from both the extensive margin (women stop working) and the intensive margin (women work fewer hours), but not from a decrease in hourly wages. Labor supply disincentives from joint taxation can explain about one quarter of the marriage earnings gap, while we find no effect for labor supply incentives from changes in divorce law. Leveraging variation in norms created by the German separation, we find that gender norms are another important driver behind the marriage earnings gap.

Keywords: Marriage, Earnings, Gender Gap, Policy Design, Joint Taxation
JEL Codes: D10, H24, J12

*We are grateful for comments from Youssef Benzarti, Clemens Fuest, Giulia Giupponi, Leonie Koch, Jonas Löbbing, Shelly Lundberg, Dominik Sachs, Jakob Søgaard, Alisa Tazhitdinova, Mirko Wiederholt, Joachim Winter, as well as participants and discussants at the BSE Summer Forum, CESifo Public Economics Week, EALE, EAYE, EDGE Jamboree, IFS NFS Workshop, IIPF, Nordic Public Policy Symposium, and seminar participants at the EU Tax Observatory, GLAB at University of Toronto, NHH, University of Konstanz, University of Mannheim, and University of St. Gallen. Urs Maier provided outstanding research assistance. We gratefully acknowledge the support of the Research Data Centres of the Federal Statistical Office, particularly Hannes Fauser, the Statistical Offices of the Federal States, and the German Pension Insurance.

†Ifo Institute, Munich Graduate School of Economics (herold@ifo.de)
‡University of Oslo (luisa.wallossek@econ.uio.no)

Women have lower earnings than men. When combining labor force participation, hours worked, and hourly wages, the average gender earnings gap in Europe is 36% (Eurostat 2023). It is well established that the earnings gap widens when women become mothers (e.g. Bertrand, Goldin, and Katz 2010; Angelov, Johansson, and Lindahl 2016; Kleven, Landais, and Søgaard 2019), often referred to as "child penalty." In this paper, we show that women's and men's earnings also diverge after another event: marriage.

Although most people marry eventually (United Nations 2019), the role of marriage for earnings, distinct from having children, is not fully understood. Using survey data for a large set of European countries, we document that women's labor force participation drops when they get married and that this marriage gap persists when accounting for the child penalty. For the main part of the paper, we focus on Germany, where we can precisely estimate the marriage earnings gap with administrative data and decompose it into participation, hours worked, and wages. In addition, the German setting allows us to analyze the extent to which financial incentives from public policy design explain the marriage earnings gap.

To estimate the marriage earnings gap, we employ an event-study approach that accounts for both the time relative to marriage and the time relative to childbirth. Although the timing of marriage and childbirth is often correlated, variation in the exact timing between the two events allows us to isolate the marriage earnings gap. We use monthly panel data from the German public pension insurance and link this administrative data with annual survey data from the largest German household survey (SOEPRV). We observe monthly earnings, as well as the month of marriage and childbirth for a subsample of the data. Monthly data is preferred over annual data here because marriage and childbirth rarely occur in the same month but more often in the same year: 17% of married women with children marry and give birth to their first child in the same year, but only 1% do so in the same month. However, we show that in the medium run, yearly and monthly data yield almost identical results for the marriage earnings gap. The linked survey data adds information that is key for understanding gender gaps but is typically not covered by administrative data, including hours worked and attitudes towards gender norms.

Using this linked data set, we estimate a significant and persistent marriage earnings gap. Female and male earnings are on similar trajectories until marriage when female earnings drop while male earnings stabilize. Female earnings decrease continuously in the first years after marriage before the gap stabilizes at a constant level: five to ten years after marriage women experience a marriage earnings gap of 20%.

The marriage earnings gap persists for spouses who are already cohabiting before marriage, ruling out that the gap is explained by couples moving in together in the year of marriage.

We then decompose the marriage earnings gap into the extensive margin (participation), intensive margin (hours worked), and hourly wages. We first first document the extensive margin marriage gap: After marriage, the share of working women decreases by 10%. Employing time-use data from the linked survey, we then estimate the intensive margin marriage gap: Conditional on working, married women reduce their working hours by 18%. For women, spending fewer hours on paid labor market work correlates with spending more hours on unpaid household work, which increases by 16%. By contrast, men do not change their time-use behavior after getting married. In terms of wages, we do not find a marriage wage penalty for women, nor a marriage wage premium for men.

We then analyze the extent to which the marriage earnings gap is explained by incentives from public policy. Individuals get married for various reasons, including love, norms, family planning, and financial reasons – but all enter the same default marriage contract, which is designed by the government. By defining the rules for the contract, policy makers can use marriage as a public policy tool to incentivize or disincentivize spousal labor supply. We study the impact of labor supply incentives from two main areas of marriage-related public policy: joint income taxation and divorce law. We exploit heterogeneity in tax rate changes and a divorce law reform that cuts ex-spousal alimony payments for secondary earners.

We first analyze the role of labor supply incentives from joint taxation. Married couples in Germany are typically taxed jointly. Under joint taxation, spouses face the same marginal tax rate (MTR) because income taxes are defined as a function of household income rather than individual income. With a progressive income tax schedule, this commonly implies a higher MTR for the secondary earner and a lower MTR for the primary earner (similar to joint taxation in other countries such as the US). Joint taxation thus disincentivizes the labor supply of secondary earner – in different sex couples, those are predominantly women. To analyze the effect of tax incentives on the marriage earnings gap, we use administrative income tax data for the full population of German taxpayers from the Taxpayer Panel (TPP). Together with Koch (2024), we developed a novel approach for linking spouses based on their tax identifiers in periods before they are married and file jointly.[1] To the best of

[1]Because individual tax identifiers were only introduced in 2010, the data only allows to link spouses who marry in 2011 or later.

our knowledge, we are the first to provide this linkage in the TPP data. The linkage is necessary to analyze changes in MTR under joint taxation, which is typically not possible because only one spouse can be traced back pre marriage.

We exploit cross-sectional heterogeneity in MTR changes under joint taxation created by the non-linear German income tax schedule to estimate the income elasticity. Building on the standard approach for estimating elasticities (e.g., Gruber and Saez 2002), we instrument for the change in MTR by using a hypothetical mechanical change in MTR based on pre-marriage income as an instrument. The obvious threat to identification here is that the MTR change is endogenous to the income level and within household income distribution, both of which are likely correlated with unobservables that impact labor supply, such as career preferences. To account for this, we include fixed effects for bins of individual income and income share when estimating the elasticity. We show that our results are robust to using different bin sizes. Intuitively, the identifying assumption is that, while having a high or low income and marrying someone with a high or low income is endogenous, the exact amount of own and spousal income can be considered as good as random.

We find that tax incentives from joint taxation increase the marriage earnings gap for women. We estimate an elasticity of gross labor income with respect to the net-of-marginal tax rate that ranges from 0.3 to 1.1 in the first years after marriage. A back-of-the envelope calculation indicates that tax incentives can explain 26% of the marriage earnings gap for women. For men, we do not find an effect of tax incentives on their earnings. Qualitatively, this result is in line with the finding that married women have a high elasticity of taxable income, while married men have a low(er) elasticity (Blau and Kahn 2007; Hermle and Peichl 2018). We also show that women experience a marriage earnings gap even if they are primary earners and thus face an MTR decrease from joint taxation, while men do not reduce their earnings if they are secondary earners. These results suggest that spousal earnings around marriage do not only respond to financial incentives but are impacted by something else, i.e., gender norms, which we exploit in more detail later.

Second, we analyze the impact of labor supply incentives from post-divorce alimony payments by exploiting a reform of the German divorce law in 2008. If eligible for post-divorce alimony, secondary earners receive monthly income transfers from their ex-spouse post divorce. With alimony, marriage acts as insurance for secondary earners who work less during marriage – in most cases women. They face a risk of losing lifetime income, both from working less during marriage and from accumulating less human capital which may translate into lower earnings post divorce. Alimony

insures against this risk by increasing the expected lifetime income via transfers from the ex-spouse in case of a divorce. The 2008 reform drastically cut the eligibility and duration of ex-spousal alimony. We exploit this exogenous variation to analyze how the indirect labor supply incentives from alimony affect the marriage earnings gap. Analyzing the earnings of the full population of German taxpayers who got married in the years around the reform (about two million individuals), we find no change in the marriage earnings gap post reform, neither for new marriages nor for those who were married already when the reform was enacted.

Taken together, these two results draw a nuanced picture of the role of labor supply incentives provided by marriage related public policy. Direct tax incentives from joint taxation impact the marriage earnings gap, while indirect incentives from alimony do not have an effect. One potential explanation for this is the difference in immediacy: income is taxed immediately, but alimony payments accrue only in the future. However, Artmann, Fuchs-Schündeln, and Giupponi (2023) find that female labor supply responds to changes in pension entitlements, suggesting that our sample is in general forward looking and responds to incentives that materialize only in the future. Another potential explanation is that alimony is relevant only if spouses divorce. Limited responsiveness to alimony may be explained by spouses underestimating their divorce probability (Berresheim and Koll 2023).

To test the role of gender norms, we leverage quasi-exogenous variation in gender norms from the German separation and reunification. Having been divided for more than 40 years, East and West Germany had very different institutions, cultures, and gender norms, particularly with respect to women's participation in the labor force. While the West German Federal Republic of Germany (FRG) was characterized by male breadwinner marriages with many stay-at-home wives, the East German Democratic Republic (GDR) had a high female labor force participation and many dual-earner couples. Differences in preferences and norms from East and West are persistent even long after reunification (Alesina and Fuchs-Schündeln 2007; Boelmann, Raute, and Schönberg 2021), allowing us to exploit the quasi exogenous variation in gender norms: We compare the marriage earnings gap for individuals born before reunification and who married post reunification. Being born in the East vs. the West, they grew up with different gender norms and different legal institutions. When they married, they faced the exact same legal institutions in reunified Germany.

We show that, while the east-west income difference for women exhibits no time trend before marriage, an event-time related east-west gap opens at marriage. After marriage, women born in the east have relatively higher income, i.e., a smaller

marriage earnings gap, than women born in the west. This finding highlights the relevance of gender norms for the marriage earnings gap. Using stated views on gender norms from the survey data, we also document a larger marriage earnings gap for women with more traditional views, providing further suggestive evidence for the role of gender norms. These results complement our finding of gender differences in responsiveness to tax incentives under joint taxation, which suggest that the observed spousal labor supply around marriage is affected by an interplay of financial incentives from public policy and gender norms.

The first main contribution of this paper is to estimate the marriage earnings gap net of the child penalty. We analyze how earnings change when individuals get married and show that female earnings drop after marriage. We add to a broad literature on marriage and income that mostly studies income and wage differences between married and unmarried individuals (e.g., Hill 1979; Korenman and Neumark 1991; Juhn and McCue 2016, 2017). An earlier literature attributed wage differences to changes in productivity after marriage, based on the hypothesis of household specialization because of comparative advantages (e.g. Becker 1985, 1991). More recent research has questioned the causal impact of marriage on productivity (e.g., Killewald and Gough 2013; Killewald and Lundberg 2017; Pilossoph and Wee 2021), and the role of comparative advantages (Siminski and Yetsenga 2022). Our findings align with these results: While we do not find evidence for a marriage wage penalty for women, we show that women face a marriage earnings gap because they work less after getting married.

We also contribute to a growing literature that documents a persistent gender earnings gap emerging after the birth of the first child, commonly referred to as the *child penalty* (e.g., Angelov, Johansson, and Lindahl 2016; Kleven, Landais, and Søgaard 2019; Kleven, Landais, Posch, Steinhauer, and Zweimüller 2019; Andresen and Nix 2021).[2] We contribute to the child penalty literature by documenting that in addition to childbirth, women's earnings also drop after marriage. Although childbirth and marriage often coincide closely, the role of marriage has been largely overlooked in the child penalty literature. A very recent exception is concurrent work by Kleven, Landais, and Leite-Mariante (2024), who estimate extensive margin "marriage penalties" together with child penalties and find small extensive margin marriage gaps for high-income countries. This result can be reconciled with ours: using linked adminis-

[2]Our paper also relates to research on expectations about the labor market outcomes of mothers (Kuziemko, Pan, Shen, and Washington 2018; Boneva, Golin, Kaufmann, and Rauh 2022).

trative and survey data from Germany, a high-income country, we show that the total marriage earnings gap is much higher than the extensive margin gap since married women also respond at the intensive margin by reducing their hours worked. Relatedly, Berniell et al. (2022) account for marriage when estimating extensive margin child penalties.[3]

The second main contribution of our paper is to analyze the mechanisms behind the marriage earnings gap. We empirically analyze how spousal earnings respond to labor supply incentives from marriage-related public policy. Here, we contribute to a literature that studies how taxes and transfers related to marital status impact the labor supply of married women. The role of joint income taxation for spousal labor supply has been emphasized by structural estimations (Bick and Fuchs-Schündeln 2018; Borella, De Nardi, and Yang 2023), cross-country comparisons (Kalíšková 2020; Bartels and Shupe 2023), as well as case studies of the introduction (LaLumia 2008; Kalíšková 2014) and abolition (Selin 2014) of joint taxation for spouses. We also contribute to a literature that studies how divorce law impacts individual and spousal decision making, specialization, and labor supply (e.g., Rangel 2006; Stevenson 2007, 2008; González and Özcan 2013; Voena 2015; Chiappori, Iyigun, Lafortune, and Weiss 2017; Foerster 2022). Our paper contributes to these two strands of literature by empirically analyzing how spousal earnings respond to two different types of labor incentives in the same setting. While we find that spouses do not respond to indirect labor supply incentives from post-divorce alimony, the earnings gap is responsive to direct labor supply incentives from income tax incentives.

We show that in addition to labor supply incentives from taxation, gender norms and culture are important mechanisms behind the marriage earnings gap. Here, we contribute to a large literature on the relevance of gender norms for gender gaps in the labor market (for an overview, see Bertrand 2011; Blau and Kahn 2017). More specifically, we also contribute to a literature that exploits differences in maternal labor supply based on gender norms and culture between East and West Germany (Boelmann, Raute, and Schönberg 2021; Jessen 2022). Related to our findings, Giommoni and Rubolino (2022) show that gender norms constitute a friction for spouses' responses to tax incentives.

The remainder of the paper is set up as follows. Section 1 provides information on the institutional background for marriage and divorce law. In section 2, we describe

[3]Intuitively, not accounting for marriage when estimating child penalties confounds both the event time coefficients and the counterfactual that the coefficients are divided by to obtain the child penalty estimate. As a result, the child penalty estimates do not necessarily change (see Berniell et al. 2022) when marriage biases the event time coefficients. We discuss the details in section 3.

the different data sets we use for our analyses and provide some basic descriptive statistics for the different samples. We explain our empirical approach and estimate the marriage earnings gap in section 3, also decomposing the marriage earnings gap and providing results for other countries. We then analyze mechanisms behind this gap in section 4. Section 5 concludes.

1 Institutional Background

Marriage as a Public Policy Marriage is a legal contract between two partners. The government defines the rules of this contract, making marriage a public policy. Typically, marriage is incentivized by granting spouses access to several advantages. These advantages can be monetary (e.g., joint taxation benefits, lower inheritance taxes, survivors' pensions) or non-monetary (e.g., simplified access to citizenship, alleviated parenthood regulations). At the same time, marriage typically also with legal and financial obligations for spouses, both during marriage and after divorce (i.e., property division, ex-spousal alimony). For post-divorce obligations, divorce law acts as a default marriage contract.[4] We study two institutional features of marriage as mechanisms behind the marriage earnings gap: joint taxation and ex-spousal alimony.

Joint Taxation Under joint taxation, income taxes are a function of household income rather than individual income. Joint taxation for married spouses exists in many countries, including the United States and several European countries. In Germany, spouses can choose between joint and individual taxation and the majority choose the former. Under joint taxation, spouses are taxed as if both earn exactly the mean household taxable income. Filing jointly is almost always beneficial for spouses with unequal income. This is different to other countries, such as the US or Switzerland, where joint taxation may lead to a tax penalty.

When taxed jointly, both spouses face the same marginal tax rate (MTR), which, in most cases, implies a decrease in MTR (labor supply incentive) for the primary earner and an increase in MTR (labor supply disincentive) for the secondary earner, compared to individual taxation. The size of the incentives depends on the level of household income and the within household shares of income. A key advantage of the German setting is that the MTR increases linearly for large parts of the income

[4]Spouses can deviate from the default contract in certain domains by signing a prenuptial agreement, but rarely do so: in the sample we use for our main analysis, only 8% have a prenuptial agreement.

distribution (see Figure A.7), which leads to continuous variation in the incentives from joint taxation. We exploit this heterogeneity in subsection 4.1, where we discuss the MTR incentives in more detail.

Ex-Spousal Alimony Upon divorce, there are two types of financial obligations for ex-spouses. First, wealth that accrued during marriage is typically divided according to a pre-defined rule. The German divorce law is a community property law, where any wealth acquired during marriage is split equally between divorced spouses. Second, divorcees may face alimony obligations for their ex-spouse during post-divorce periods. Under the German divorce law, this applies if the economically weaker ex-spouse cannot financially support themselves adequately (e.g., because of childcare, illness, age, or unemployment).[5] Alimony payments then amount to 3/7 of the difference in net income between ex-spouses, with the alimony-paying spouse being entitled to a fixed minimum of their own net income. In 2008, a reform drastically cut alimony eligibility, as a result of which secondary earners can expect to receive alimony for a shorter period – and often none at all – from their ex-spouse after divorce. We exploit this exogenous variation in subsection 4.2, where we also discuss the reform in more detail.

2 Data

For our main analysis, we utilize two data sets: German administrative pension data linked to survey data from the SOEP (SOEPRV) and German administrative income tax data (TPP). We explain the different data sets below. In addition, we use international SHARE survey data to estimate the marriage gap for a large set of European countries. We describe the SHARE data in subsection A.3.

2.1 SOEPRV

The SOEPRV data is jointly provided by the Research Data Centre of the German Pension Insurance (FDZ-RV) and the German Institute for Economic Research (DIW) (Goebel et al. 2022). It combines survey data from the Socioeconomic Panel (SOEP) and administrative data from pension accounts (VSKT). The SOEP is the

[5] Alimony payments are meant to support the ex-spouse. In addition, divorced parents typically face child support obligations if they are the non-custodial parent, i.e. if their child lives with their ex-spouse. Children are eligible for child support irrespective of parental income.

largest German household panel, and has been conducted annually since 1984. It provides household and individual-level data for a representative sample of the German population. The data set includes a unique identifier to link individuals to administrative VSKT data. The record linkage is available for respondents who participated in the 2018 and 2020 SOEP waves and consented to the linkage. Lüthen et al. (2021) provide a detailed description of the record linkage process.

The linked administrative data is a version of the VSKT pension data, provided specifically for the SOEP linkage. The VSKT is a unique administrative data set that combines monthly earnings from the entire biography of individuals with precise marriage and divorce information for a subset of individuals. Because public pension entitlements are split upon divorce, the German Pension Insurance has information on marriage and divorce as soon as individuals get divorced and split their pension entitlements. For this subsample, we use the marriage info from the administrative data. For everyone else, we use the marriage info from the survey data. We benefit from the combination of precise administrative data and rich survey data and use the SOEPRV data to exploit earnings around marriage.

Variables. The SOEPRV contains the full set of variables available from the SOEP as well as a set of variables from the VSKT. Most relevant for our analysis, the VSKT data covers monthly gross wage earnings, as well as marriage and divorce dates for divorced individuals. For women, the data also includes the year of childbirth. All variables are on a monthly level. The Research Data Centre of the German Pension Insurance (FDZ-RV) (2022) provides a detailed description of available variables.

Our main outcome is monthly gross earnings. Earnings are measured in so-called *earnings points* (EP), with 1 EP (1/12 EP) corresponding to the average annual (monthly) gross wage income in a given calendar year. Because EPs are normalized by average annual income, they account for overall income growth and are comparable over time.[6] While our main outcome is earnings from the administrative data, we use survey data to analyze other outcomes, i.e., time use and hourly wages. We also use SOEP variables that reveal individual characteristics, such as marriage information, or information on children, if these are not covered in the administrative data.

Sample. The SOEPRV data covers a total of about 12,000 individuals for whom we can link administrative and survey data. The linked data includes mainly employees, since self-employed individuals, civil servants, and farmers are typically not included in the pension records. Our main sample includes 6291 individuals for whom we

[6]EPs are top-coded at about 2, with no information on the exact income of individuals with income exceeding this cutoff.

Table 1: Descriptive Statistics SOEPRV Sample

	All	Women	Men
Age at marriage	27	26	28
Share married East Germany	0.22	0.22	0.23
Year of marriage	1988	1989	1988
Share divorced	0.18	0.20	0.17
Years married	13	13	13
Share children (ever)	0.92	0.93	0.92
Gross monthly income t-12	1898	1667	2150
Source VSKT-VA	0.18	0.20	0.17
N	6,283	3,276	3,007

Data: SOEPRV, raw data (unweighted), mean values. The sample contains individuals that are married at some point, aged 18 to 45 at marriage, and observable a minimum of four years before and eight years after marriage. If available, we take the marriage date from the administrative VSKT data. If not, we take the marriage date from the SOEP survey data. Income is measured in the year before marriage and reported in 2015 Euros.

observe their first marriage, either in the VSKT (if divorced) and/or in the SOEP (if reported in the survey). We include individuals aged 18 to 45 at the time of marriage, for whom we can observe at least four years prior to and eight years after marriage. Table 1 provides descriptive statistics. We use this sample for our main analysis in subsection 3.2.

2.2 Taxpayer Panel

The *Taxpayer Panel* (TPP) is an annual panel of income tax payers for the years 2001 to 2018 provided by the Research Data Center of the German Federal Statistical Office and the statistical offices of the Länder (RDC, 2022). Our sample covers the full population of German income taxpayers born between 1951 and 1998. The TPP contains income tax return data for tax filers as well as employer provided information for non-filers (similar to the W-2 form in the US, available only since 2012). Taxpayers are either individuals or joint filing married spouses. Upon joint filing, one spouse (typically the woman) is added to the other spouse's (typically the man's) spell. Germany introduced individual tax identifiers only in 2010, allowing us to link spouses also before marriage if they married in 2011 or later. To the best of our knowledge, together with Koch (2024), ours is the first study to link German income tax data for spouses before marriage.

We use the TPP data set to analyze the impact of financial incentives on the marriage earnings gap. The advantage of the TPP data set here is twofold. First, it provides the exact taxable income, which is crucial for analyzing tax incentives in subsection 4.1. Second, the large sample size from the full population data allows us to test for changes in the marriage earnings gap between year-of-marriage cohorts around the 2008 alimony reform in subsection 4.2.

Variables. Data on the full taxpayer population is only provided upon individual request for research projects and made available only with a limited number of variables. Our customized data set contains detailed information on labor income and total income, as well as demographics such as marital status, children, age, state, or religion. Joint filing spouses are treated as one taxpayer unit, but the data reports income, etc. on the individual level. All variables are at the annual level.

Sample. In subsection 4.1 and 4.2, we draw two subsamples of newlyweds for our analyses. To assess the impact of joint taxation on the marriage earnings gap (subsection 4.1), we draw a subsample of newlyweds who married between 2011 and 2015, whom we observe three years before and after getting married. For this sample, we link the spouses before marriage, which is crucial for analyzing changes in MTR from joint taxation.

For our analysis of the reduction of ex-spousal alimony (subsection 4.2), we draw a subsample of newlyweds who married between 2004 and 2012, i.e., +/- 4 years around the 2008 reform date. For this sample, the TPP data does not allow for linking spouses before marriage, since individual taxpayer IDs were only introduced in 2010. Since the event study estimation does not require to link individuals over periods, we can estimate the marriage earnings gap for this sample as well (we provide more details on this in subsection 4.2). Table 3 shows the descriptive statistics.

3 Estimating the Marriage Earnings Gap

3.1 Framework for Estimating the Marriage Earnings Gap

Naive Approach: Not Accounting for Children To estimate the marriage earnings gap, we first employ an event study approach that closely follows the standard event study approach for child penalties (see e.g., Kleven, Landais, and Søgaard 2019): We regress the outcome of interest Y_{it}^g on a full set of dummies for event time t defined relative to the event of marriage. We refer to this as a *naive approach* because it does not account for child-related earnings changes. Since the timing of marriage is

Table 2: Descriptive Statistics TPP – MTR Sample

Men age at marriage	36
Women age at marriage	34
Year of marriage	2013
Share East Germany (women)	0.13
Share East Germany (men)	0.13
Share children pre marriage (women)	0.22
Share children pre marriage (men)	0.22
Share children (ever)	0.82
Share working t-1 (women)	1.00
Share working t-1 (men)	1.00
Mean gross labor income t-1 (women)	35,881
Mean gross labor income t-1 (men)	49,941
Median gross labor income t-1 (women)	33,531
Median gross labor income t-1 (men)	43,777
Female income share t-1	0.43
MTR t-1 (women)	0.2985
MTR t-1 (men)	0.3376
Hypothetical joint MTR t-1	0.3241
N	192,596

Data: TPP, This table shows the descriptive statistics for the TPP subsample of spouses that can be linked pre-marriage and that we use for our analysis in subsection 4.1. For details on the sample restrictions, see subsection 4.1. This is the sample we use for estimating Equation 4 and 5. Numbers are mean values unless specified differently, income in 2015 Euros. N refers to the number of couples.

often correlated with the timing of having a first child (see Figure A.1), children are a confounder this naive estimation does not account for. We estimate the following event study regression for the naive approach:

$$Y_{it}^{g} = \sum_{j \neq r} \beta_j^g \cdot \mathbb{1}[j = t] + \zeta_{cal}^{g} + \eta_{age}^{g} + \epsilon_{it}^{g}, \tag{1}$$

where Y_{it}^{g} denotes the labor market outcome of individual i of gender g at event time t. We estimate Equation 1 separately for women and men. The event dummies capture the dynamics of Y_{it}^{g} at event time $t = j$, yielding the coefficients of interest β_j^g. Event time t is defined relative to the event of marriage at $t = 0$, omitting the reference period r. For estimations with monthly data, we measure t in months

Table 3: Descriptive Statistics TPP – Alimony Sample

Men age at marriage	33
Women age at marriage	30
Year of marriage	2009
Share married East Germany	0.16
Share divorced	0.16
Years married	5
Share children (ever)	0.82
Men gross labor income t-1	30,848
Women gross labor income t-1	23,844
N	2,165,423

Data: Taxpayer Panel (TPP), mean values. This table shows the descriptive statistics for the TPP subsample we use for our analysis in subsection 4.2: individuals who married between 2004 and 2012. *Share divorced*: share of individuals with observed divorce (data until 2018). *Years married*: length of marriage if divorced. *N*: number of individuals in the subsample.

and omit $r = t - 12$, i.e. the period 12 months before the month of marriage. For estimations with yearly data, we measure t in years and omit $r = t-1$, i.e. the period before the year of marriage. Year-of-age fixed effects η^g_{age} capture potential life-cycle effects that are correlated with age k, and calendar-year fixed effects ζ^g_{cal} account for common time trends. Since we estimate Equation 1 separately for women and men, the fixed effects account for potential gender differences in life-cycle and time trends.

Main Specification: Accounting for the Child Penalty For our main specification, we account for the child penalty by adding a full set of event dummies for event time relative to childbirth to our event study specification:

$$\begin{aligned} Y^g_{it} = &\sum_{j \neq r} \beta^g_j \cdot \mathbb{1}[j = t] \\ &+ \sum_k \gamma^{1,g}_k \cdot \mathbb{1}[k = t^{child1}] + \sum_l \gamma^{2,g}_l \cdot \mathbb{1}[l = t^{child2}] + \sum_m \gamma^{3,g}_m \cdot \mathbb{1}[m = t^{child3}] \\ &+ \zeta^g_{cal} + \eta^g_{age} + \epsilon^g_{it}, \end{aligned} \tag{2}$$

where t^{child1}, t^{child2}, and t^{child3} denotes the event time relative to the birth of the first, second, and third child, respectively. The coefficients $\gamma^{1,g}_k$, $\gamma^{2,g}_l$, and $\gamma^{3,g}_m$ capture

dynamics in Y_{it}^g that correlate with the event time relative to the respective birth.[7] Note that we do not omit any time dummy for $\sum_k \gamma_k^{1,g}$, $\sum_l \gamma_l^{2,g}$, and $\sum_m \gamma_m^{3,g}$. Because not all married individuals have children, the omitted group is individuals who never have a (first, second, third) child. Our coefficient of interest is still β_j^g, now denoting the change in earnings relative to marriage that is not correlated to the event time with respect to having a first, second, or third child. Following the same steps as for Equation 1, we estimate Equation 2 separately for gender $g = \text{w, m}$, denoting women and men, respectively.

Although the first child is often born relatively soon after marriage, there is large variation in the exact timing (Figure A.1), which allows us to isolate the estimates for β_j^g from $\gamma_k^{1,g}$. A unique feature of the SOEPRV data is that it provides monthly data for the exact date of marriage and childbirth for a subsample of married individuals.[8] With monthly data, marriage and childbirth rarely happen in the same period: Only 1% of married women who have children at some point in their life have their first child in the month of marriage. With yearly data, the variation is mechanically smaller, i.e., more women have their first child in the period of marriage. However, we show that the results remain virtually unchanged when using annual data rather than monthly.

Estimating the Marriage Earnings Gap To estimate the marriage earnings gap, we follow the standard approach for child penalties from Kleven, Landais, Posch, Steinhauer, and Zweimüller (2019). We first estimate Equation 1 and 2 separately for women and men. This yields gender-specific absolute estimates $\hat{\beta}_t^g$, where $g = \text{w, m}$ for women and men, respectively.

We then rewrite the absolute estimates in relative terms:

$$MG_t^g \equiv \hat{\beta}_t^g / E[\tilde{Y}_{it}^g | t], \tag{3}$$

where $\tilde{Y}_{it}^g | t$ denotes the counterfactual outcome, i.e., the predicted outcome without the event dummy estimates. The child penalty literature then typically defines the total penalty as $P_t \equiv (\hat{\beta}_t^m - \hat{\beta}_t^w)/E[\tilde{Y}_{it}^w | t]$, which is interpreted as the percentage by which women fall behind men after the event (see, e.g., Kleven, Landais, and Søgaard

[7] Berniell et al. (2022) and Kleven, Landais, and Leite-Mariante (2024) employ a similar approach to estimate child penalties while accounting for event time with respect to marriage.

[8] Those for whom the administrative VSKT data provides marriage info (this is conditional on divorce) and those who married while already in the SOEP sample. We have yearly dates for everyone else, i.e., those without divorce who married before entering the SOEP sample.

2019). This measure does not account for a gender gap in the counterfactual incomes, which is why we only report MG_t^g and do not compute P_t. We provide a detailed explanation in Appendix A.2.

We follow the same procedure for the estimates from Equation 1 and Equation 2. It is important to note that not accounting for the child penalty in Equation 1 not only confounds $\hat{\beta}_t^g$, but also $E[\tilde{Y}_{it}^g|t]$. The marriage gap estimates (MG_t^g) from Equation 1 and Equation 2 are thus not directly comparable. This is because accounting for event time with respect to childbirth affects both the numerator (the coefficient for event time with respect to marriage) and denominator (the counterfactual absent the contribution of these estimates).

3.2 Main Result: There is a Marriage Earnings Gap

We first estimate the marriage earnings gap with monthly event time for the subsample of individuals for whom we know the exact monthly date of their marriage and the birth of their first child (if they have children) from the SOEPRV data. With monthly data, the two events almost never happen in the same period, and there is large variation in the exact timing of both, as shown in Figure A.1. This allows us to isolate both events. We start with the sample with monthly info as a proof of concept before moving to the full sample, where we show that annual event time yields similar results.

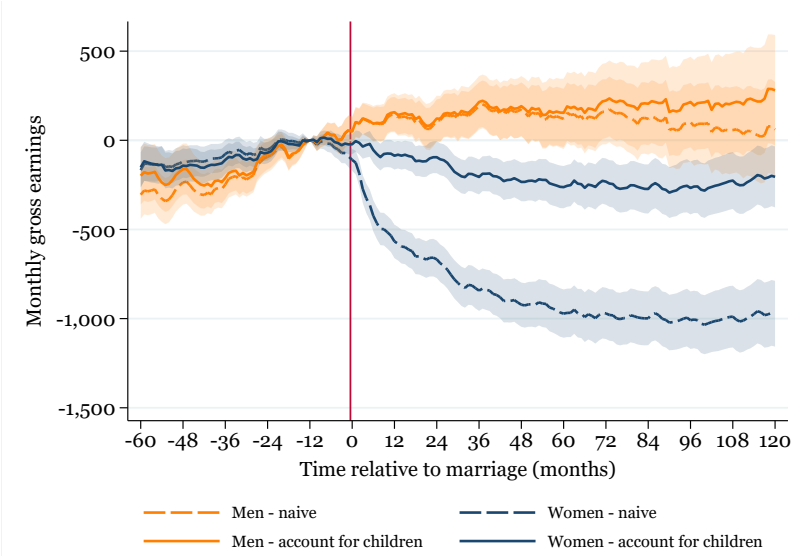
Figure 1 plots the results from estimating Equation 1 and Equation 2 with monthly data. Immediately after marriage, women's monthly gross earnings start decreasing. The gap widens over time before stabilizing at a constant level after some years. Averaging over five to ten years after marriage, the naive estimate $\hat{\beta_w}$ is €1128 (Figure 1a). A large part of this is explained by the child penalty, but the remaining drop is sizable: After accounting for the child penalty, women's monthly gross earnings still decrease by €251 per month, equivalent to €3012 in annual earnings. This absolute estimate translates into a relative marriage earnings gap of $MG_t^w = 20\%$ (Figure 1b).

We then turn to the full SOEPRV sample, where we can estimate the marriage earnings gap with annual event time. Mechanically, it is more common that marriage and childbirth happen in the same period if time is measured in years rather than months: 17% of married women have their first child in the year of marriage (compared to 1% giving birth in the month of marriage). However, there is still substantial variation (Figure A.1c), allowing us to estimate Equation 1 and 2 with annual data.

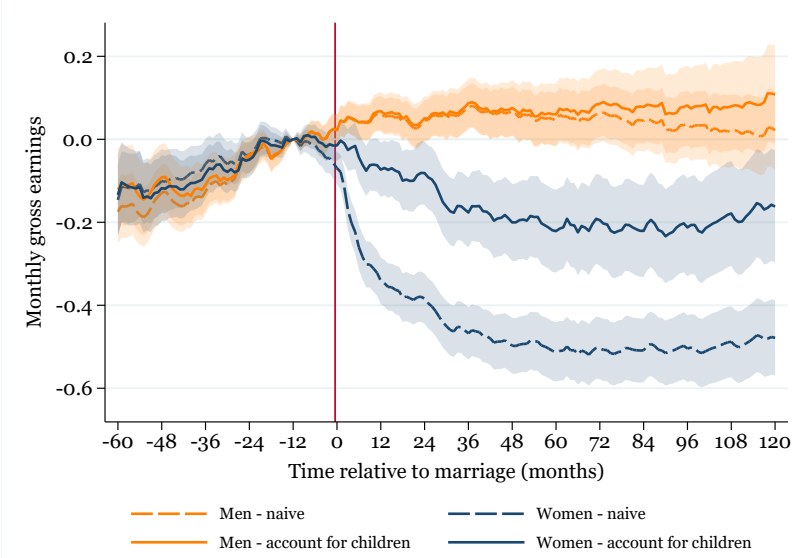
Figure 2a plots the absolute estimates for the marriage earnings gap using annual

Figure 1: Marriage Earnings Gap – Monthly Event Time

(a) Absolute Estimates



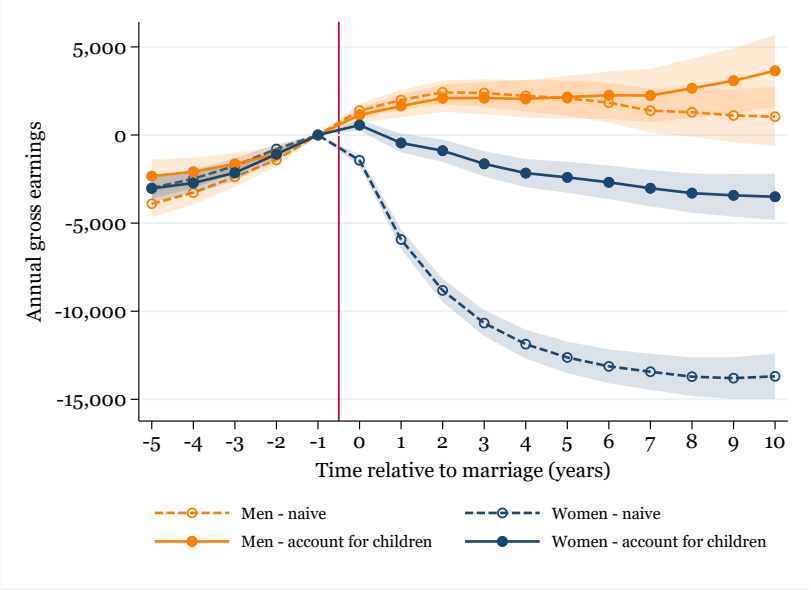
(b) Relative Gap Estimates



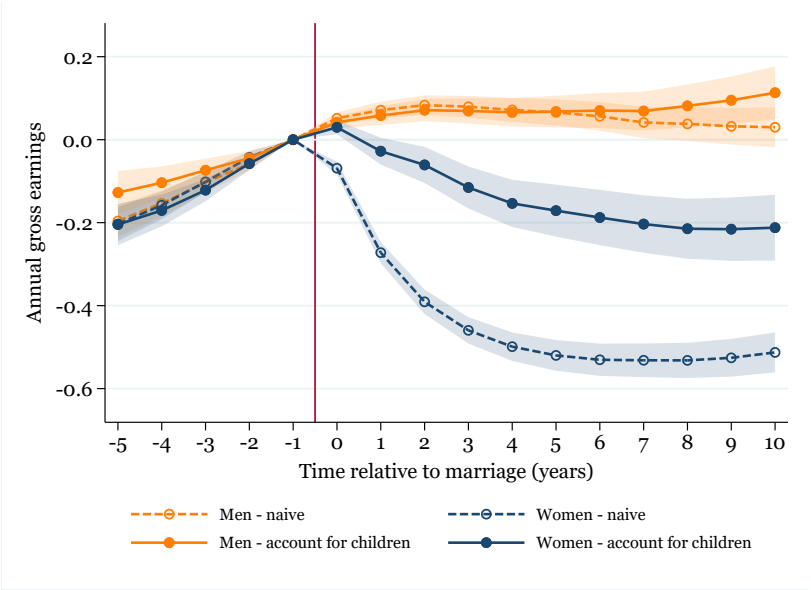
Notes: Monthly data from SOEPRV. This figure shows the event study estimates from estimating Equation 1 (dashed lines) and Equation 2 (solid lines) for monthly gross earnings for the SOEPRV subsample for which we observe the exact month of marriage and childbirth (if an individual has children). Plots show point estimates and 95% confidence intervals. Panel a shows the absolute estimates $\hat{\beta}_t^g$ in 2015 Euros, directly retrieved from estimating Equation 2. Panel b shows relative earnings gap estimates for MG_t^g as defined in Equation 3.

Figure 2: Marriage Earnings Gap – Annual Event Time

(a) Absolute Estimates



(b) Relative Estimates



Notes: Data from SOEPRV, aggregated over years. This figure shows the event study estimates from estimating Equation 1 (dashed lines) and Equation 2 (solid lines) for annual gross earnings for the full SOEPRV sample. Plots show point estimates and 95% confidence intervals. Panel a shows the absolute estimates $\hat{\beta}_t^g$ in 2015 Euros, retrieved directly from estimating Equation 2. Panel b shows relative earnings gap estimates for MG_t^g as defined in Equation 3.

data. Similar to the results with monthly data, women's and men's earnings are on very similar trajectories before marriage.[9] After marriage, a gap opens: Women's earnings decrease, while men's earnings increase slightly. The immediate drop in earnings here is less pronounced than for the monthly estimation. Over time, however, the results are very similar: Averaging over the period five to ten years after marriage, women's annual earnings drop by €3060, compared to €3012 for the subsample with monthly event time (results are also very similar for the naive estimation with €13 404 vs. €13 536 with monthly event time).

We then translate the absolute estimate into a relative estimate for the marriage earnings gap. Figure 2b plots the resulting estimates for MG_t^g for gender g over event time t. Averaging MG_t^w over five to ten years after marriage, women's earnings drop by 20%. This is almost exactly the same result as for the estimation with monthly event time (20%).

The MG_t^g estimates from Equation 1 and Equation 2, i.e., with and without accounting for the child penalty, are not directly comparable. Not including event time for children in the naive estimation not only confounds $\hat{\beta}_t^g$, but also $E[\tilde{Y}_{it}^g|t]$, the denominator of $MG_t^g = \hat{\beta}_t^g / E[\tilde{Y}_{it}^g|t]$. A similar argument can be applied to earnings gaps related to children (child penalties). When comparing absolute estimates, marriage clearly confounds the child penalty estimate: The child penalty estimate for annual earnings $\hat{\beta}_t^w$ is €2261 smaller when accounting for the marriage earnings gap, equivalent to a reduction of 13% (Figure A.3a).[10] When transforming the absolute estimates to the relative child penalty estimates, the difference is much less pronounced (Figure A.3b). This is in line with the results of Berniell et al. (2022), who find no significant difference for child penalties with and without accounting for marriage.

The marriage earnings gap is not explained by a change in the cohabitation status at the event of marriage. Figure A.4 compares the marriage earnings gap for spouses that cohabited already in the year prior to marriage ($t = -1$) and those that did not. The marriage earnings gap remains, and is even larger, for those who are already cohabiting, ruling out that the drop in $t = 0$ is driven by spouses moving in together

[9] We find earnings increase for $t < 0$, a pattern similar to child penalty estimates for German-speaking counties (e.g., Kleven, Landais, Posch, Steinhauer, and Zweimüller 2019; Boelmann, Raute, and Schönberg 2021). Intuitively, this means the timing of marriage is correlated with income growth that age and calendar-year fixed effects do not capture. Figure A.2 shows that accounting for career age, defined as time since entering the labor market, reduces income growth for $t < 0$ without visibly impacting the marriage earnings gap estimates.

[10] The absolute difference is not exactly the same as the difference for the marriage earnings gap in Figure 2a. This can be explained by the fact that the underlying samples differ: For the marriage earnings gap, everyone in the sample is married, and some have children, whereas for the child penalty, everyone has children, and some are married.

in that period.

3.3 Decomposition of the Marriage Earnings Gap

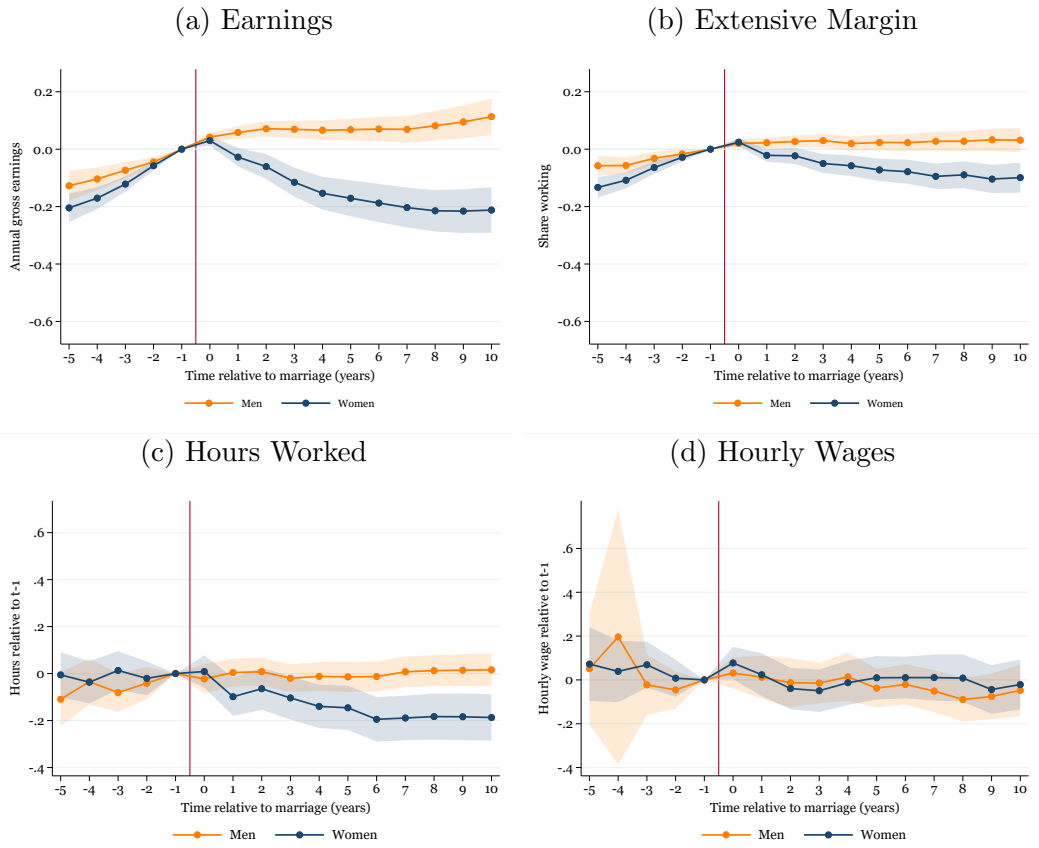
Earnings are defined as the product of the extensive margin (working yes or no), the intensive margin (hours worked), and hourly wages. In this section, we decompose the change in earnings into its three components. Administrative data informs about earnings and the extensive margin. Linking administrative and survey data, we can add information on hours and back out hourly earnings for a subsample of individuals. This information is only available in survey data. We therefore restrict the analysis of hours and wages to those who are already in the survey population in the years around marriage. Since the outcome from survey data is available on the annual level, we estimate all gaps with annual data here. When the outcome is measured monthly (earnings and extensive margin), we average the monthly data over years.

Extensive Margin For the extensive margin marriage gap, we estimate Equation 2, with the outcome being a dummy for whether individual i is working in month t. We define working as being in either regular employment, marginal employment (so-called mini-jobs), or self-employment, as reported in the administrative data. Figure 3b plots the extensive margin marriage gap. We show that women are significantly less likely to work after marriage. Averaging over years five to ten after marriage, women are 10% less likely to be working. In absolute terms, this is equivalent to a reduction of 6 percentage points.

Intensive Margin We define the intensive margin as hours worked, conditional on working. Administrative data rarely informs about hours. We overcome this by using time-use data from the SOEP.[11] To study how the hours worked change after marriage, we restrict the sample to individuals who work in a given t (we take this info from the administrative data). In addition, we now restrict the sample to those already in the SOEP during the period of interest since the outcome of interest is retrieved from the SOEP data. This excludes individuals who married before they entered the survey population. We then estimate Equation 2 with the outcome Y_{it}^{g} being hours. Figure 3c plots the results: Five to ten years after marriage, women reduce their hours by 18% conditional on working. In absolute terms, this is equivalent to a daily reduction of 1.2 hours.

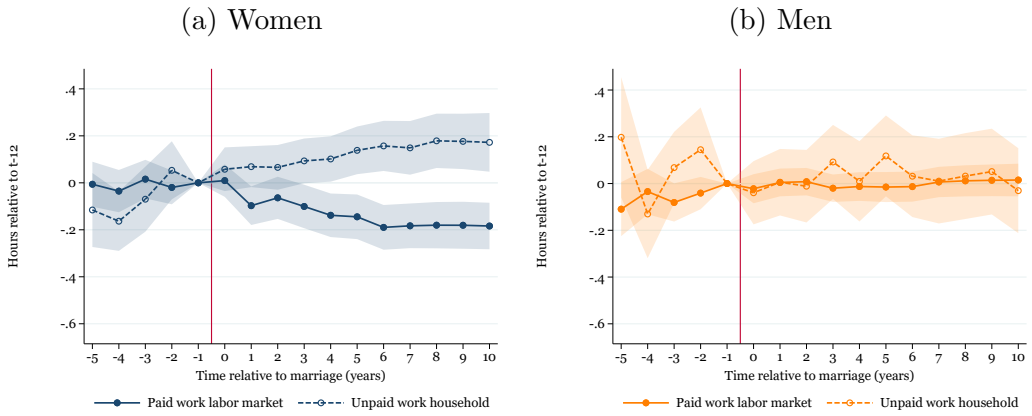
[11] Jessen (2022) validates the SOEP data on time use by comparing it against time use reported in diary data, specifically designed to measure time use.

Figure 3: Decomposition Marriage Earnings Gap



Notes: Data from SOEPRV. This figure decomposes the overall marriage earnings gap (panel a) into extensive margin (panel b), intensive margin (panel c), and hourly wages (panel d). All panels plot MG_t^g estimates for women and men from estimating Equation 2 for different labor market outcomes. For all panels, event time is defined in years. We exclude $t = -1$, i.e., the year before the year of marriage. Panel c plots results for the intensive margin only, conditional on working in a given t. Hourly wages in panel d are defined as earnings divided by hours worked.

Figure 4: Hours Spent in the Labor Market vs. Household



Notes: Annual data from SOEPRV. This figure shows the event study estimates from estimating Equation 2 for time use, i.e., the daily hours spent on a given task. This is intensive margin only, conditional on working in a given t. Time-use data stems from the SOEP questionnaire. We estimate Equation 2 for annual event time and exclude $t = -1$, i.e., the year before the year of marriage.

Analyzing time-use data, we show that the decrease in hours spent on paid work in the labor market for women goes along with an increase in hours spent on unpaid household production (Figure 4a). Five to ten years post marriage, women spend .5% hours more in the household. For men, time-use does not change considerably after marriage (Figure 4b).

Hourly Wages We impute wages by dividing earnings by hours worked. The measure for earnings comes from the administrative data that provides monthly gross earnings (the same measure we use for our main analysis in subsection 3.2). The measure for hours worked comes from survey data that provides information on daily time use (the same measure we use for the intensive margin analysis).

Figure 3d plots the results from estimating Equation 2 for hourly wages. Women's wages do not decrease post marriage. This indicates that the decrease in female earnings after marriage does not stem from decreased productivity. The results are qualitatively the same when using income information from the survey data, i.e., relying on the same data source for both hours and income (see Figure A.5).

Taken together, our decomposition results draw a nuanced picture of how labor market outcomes change after marriage: The marriage earnings gap for women is driven by how much they work and not by how much they are paid. Seeing women shift their hours from paid work in the labor market to unpaid work at home is in line with

predicted specialization behavior from the early literature on household specialization (e.g., Becker 1985, 1991), while finding no wage decrease for women is not. Here, our findings are in line with more recent work, which has questioned the marriage wage penalty for women (e.g., Killewald and Gough 2013; Killewald and Lundberg 2017; Pilossoph and Wee 2021).

3.4 Cross-Country Evidence

The marriage earnings gap is not unique to Germany. To show this, we use cross-country survey data for more than 20 European countries and Israel (SHARE JEP).[12] With the SHARE JEP data, we can estimate Equation 2 for the extensive margin. To validate the SHARE JEP in our context, we first compare the estimates for Germany to our estimates from the administrative German data from our main analysis. Although the data sets are very different and the sample comprises on average earlier marriages than the SOEPRV sample (see Table 1 and Table A.4), the results are virtually the same: the extensive margin marriage gap for Germany is 10% when using the SHARE data (see Figure A.19i) and 10% when using the administrative data from SOEPRV.

The average gap, estimated for women from all countries, is 11% (Figure 5a). This is a lower bound for the total marriage earnings gap, since it only considers the extensive margin. While the average marriage gap across European countries is similar to the marriage gap in Germany, we document substantial heterogeneity between countries (Figure 5b). The point estimates range from 33% in Ireland to no long-term gap in northeast countries such as Latvia, Finland, or Estonia. We provide the event study plots for all countries in Appendix A.3.

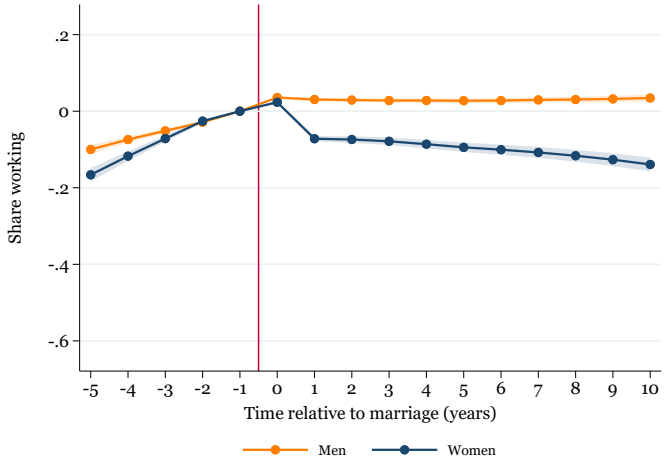
Figure 5b reveals an East-West divide with smaller marriage gaps in the East.[13] On average, women in East-European countries experience a marriage gap of 4% as compared to 18% in the West. This pattern is in line with cultural differences with respect to gender norms in the labor market context. Shaped by communist regimes, East European countries commonly share a dual-earner norm while Western European countries often share a male-breadwinner norm. However, also within East and West, the size of the marriage gap varies between countries. One descriptive pattern is that the gap is larger in countries where marriage is associated with tax benefits (15%) as

[12] We provide more details for the cross-country analysis in Appendix A.3.

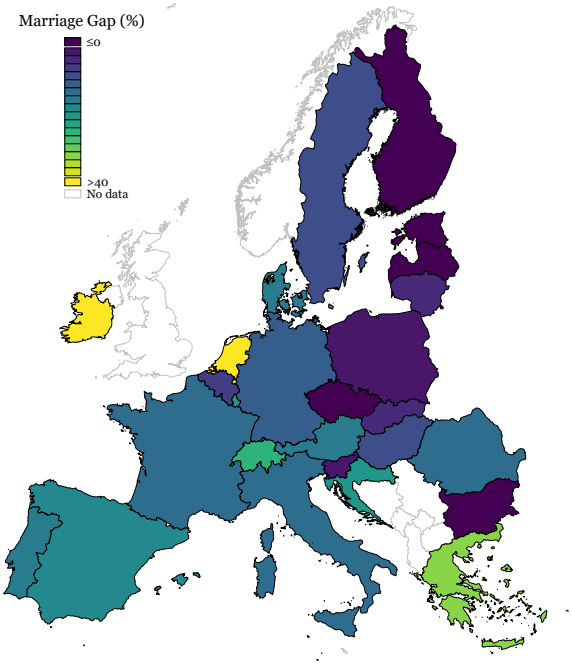
[13] We consider BG, CZ, EE, HR, HU, LT, LV, PL, RO, SI, and SK as Eastern and AT, BE, CH, DE, DK, EL, ES, FI, FR, IE, IL, IT, LU, NL, PT, and SE as Western.

Figure 5: The Extensive Margin Marriage Gap Across Europe

(a) Event Study for All Countries



(b) Long-Term Gap Estimates by Country



Notes: Annual data from SHARE. This figure shows the event study estimates from estimating Equation 2 for the extensive margin (working yes/no) with annual event time. For the event time dummies, we exclude $t = -1$, i.e. the year before the year of marriage. Panel a plots the estimate for MG_t^g when including individuals from all countries in the SHARE data, except Germany. Panel b plots the extensive margin marriage gap estimate for women (MG_t^w) for each country.

compared to countries with no benefit (8%).[14]

4 Mechanisms

The cross-country comparison suggests labor supply incentives from public policy and gender norms as potential mechanisms behind the marriage earnings gap. To study the effect of these mechanisms, we focus on Germany again. Using German administrative data, we explore variation in both financial incentives and norms. To determine how labor (dis-)incentives set by the institutional framework of marriage impact spousal earnings, we analyze the direct impact of joint taxation of spouses, as well as the indirect effect of alimony payments, on spousal earnings. We then study the role of gender norms by exploring quasi-exogenous variation from the German separation and reunification.

4.1 Direct Labor Supply Incentives from Joint Taxation

Mechanism Under joint taxation, the spousal income tax liability is a function of the total income of both partners, i and j. In Germany, spouses can choose between individual taxation (T^{single}) and joint taxation (T^{joint}). The vast majority choose joint taxation. When filing individually, the total tax liability is the sum of the individual liabilities:

$$T^{single}(y_i, y_j) = T^{single}(y_i) + T^{single}(y_j).$$

When filing jointly, spouses are taxed as if both earned exactly 50% of the combined household income:

$$T^{joint}(y_i, y_j) = 2 \times T^{single}\left((y_i + y_j)/2\right).$$

Under the progressive German income tax schedule, there is no tax penalty for jointly filing spouses[15]:

[14] We follow the definition from Christl, De Poli, and Ivaškaitė-Tamošiūnė (2023), who identify a marriage tax benefit as paying less taxes if married, as compared to cohabiting. They document a marriage tax benefit for BE, CZ, DE, GR, HR, IE, IT, LV, LU, NL, PL, SI, SK, and ES, and no benefit for AT, BG, FI, HU, LT, PT, RO, and SE. Since this includes EU countries only, we exclude non-EU countries here.

[15] There are few special cases where individual taxation is beneficial. For example, for certain types of wage-replacement benefits or loss carryforwards, individual taxation can be beneficial under certain circumstances, typically cases where spouses have similar income levels with small or no tax saving potential from income splitting.

$$T^{joint}(y_i, y_j) \leq T^{single}(y_i, y_j) \text{ for all } y_i, y_j.$$

Since the German income tax code is characterized by linearly increasing MTR for large parts of the income distribution (see Figure A.7), joint taxation typically implies a tax benefit as long as $y_i \neq y_j$. $T^{joint}(y_i, y_j) = T^{single}(y_i, y_j)$ for $y_i \neq y_j$ only if both spouses' incomes y_i and y_j fall in the same income tax bracket with a flat MTR. This applies only to very low-income couples, with both y_i and y_j below the basic allowance threshold (MTR = 0), and very high income couples with both y_i and y_j in either of the two top tax brackets (MTR = 0.42 and 0.45, respectively). We illustrate the tax benefit as a function of household income and income shares in Figure A.8.

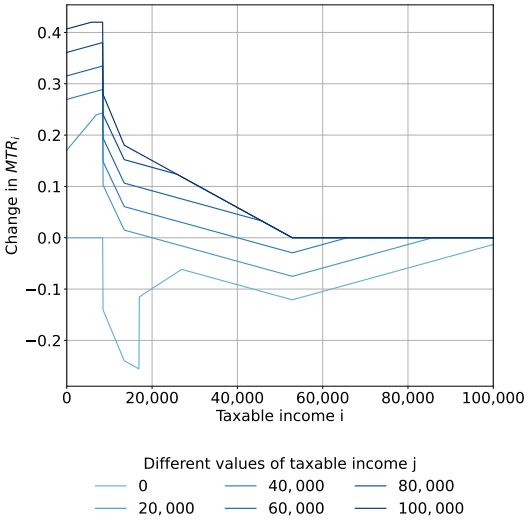
In terms of MTR, both spouses face the same tax rate when filing jointly. This typically decreases the MTR for the primary earner while increasing the MTR for the secondary earner:

$$MTR^{single}(y_i) \leq MTR^{joint}(y_i, y_j) \leq MTR^{single}(y_j) \text{ if } y_i < y_j.$$

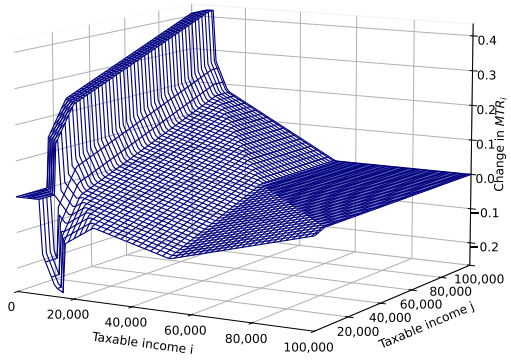
The MTR for the secondary earner remains unchanged only if y_i and $(y_i+y_j)/2$ fall in the same bracket with a flat MTR (true for very low and very high income only) or if both spouses have exactly the same income. For all other cases, secondary earners face an increase in MTR: $\Delta MTR_i = MTR^{joint}(y_i, y_j) - MTR^{single}(y_i) > 0$, implying a disincentive for their individual labor supply. As depicted in Figure 6, the size of the disincentive depends on both spouses' income. In general, ΔMTR_i increases with the relative income difference among spouses, effectively incentivizing spousal specialization. The residual heterogeneity in ΔMTR_i can be attributed to one notch and two kinks between the five different income tax brackets of the German income schedule as shown in Figure A.7. The behavioral response, i.e., the extent to which i reduces their income due to ΔMTR_i, depends on the elasticity of taxable income with respect to the net-of-tax rate $\varepsilon_{y_i,1-MTR_i}$: $\Delta y_i = -\varepsilon_{y_i,1-MTR_i} \cdot y_i \cdot \Delta MTR_i/(1-MTR_i)$.

Figure 6: Heterogeneity in MTR Changes

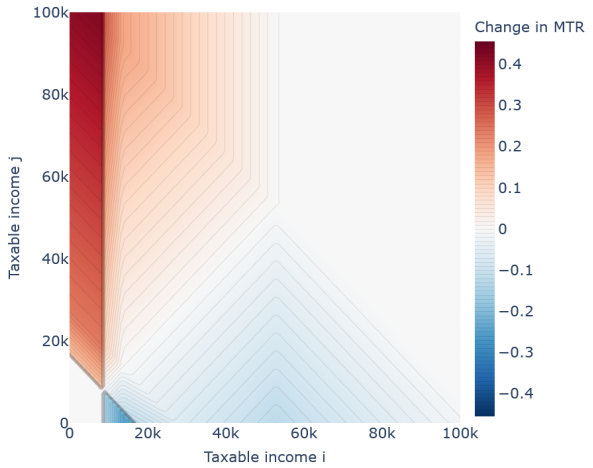
(a) Selected Values of Spousal Income



(b) Full variation (3D view)



(c) Full variation (2D view)



Notes: All three panels of this figure illustrate the heterogeneity in the change in MTR under joint taxation (ΔMTR_i) for spouse i as a function of their own and their spouse's taxable income (y_i and y_j, respectively) for taxable income up to €100,000. Each panel is a different visualization of the same heterogeneity. Panel (a) plots ΔMTR_i over y_i for six different levels of y_j. Panel (b) plots ΔMTR_i for all combinations of y_i and y_j. Intuitively, each line from panel (a) is a slice of panel (b), cut in parallel to the x-axis at the respective y_j value on the y-axis. Panel (c) provides the same information as panel (b) but uses colors instead of a third dimension for plotting ΔMTR_i. Intuitively, panel (c) is a top view of panel (b), where MTR increases are colored in red and decreases in blue, while income combinations with unchanged MTR are gray. The lines mark 1-percentage-points increments of ΔMTR_i.

Testing the Mechanism To analyze the impact of the direct labor disincentives set by joint taxation, we exploit the heterogeneity in ΔMTR_j depicted in Figure 6 to estimate the income elasticity for married women and men. We face two challenges for identification here.

The first challenge is that the change in MTR and the change in income are observed simultaneously, with the latter impacting the former and thus introducing a bias in the elasticity estimate. This is a standard challenge when estimating elasticities based on changes in the tax rate (e.g., after a tax reform). The standard solution to this is to instrument for the observed change in MTR by using a hypothetical mechanical tax rate change as the instrument (for details and a literature overview, see Neisser 2021). Following Gruber and Saez (2002), the "most standard approach" (Neisser 2021, p. 3373) for constructing the instrument is to use the income from the period before the change in MTR. When analyzing a tax reform, this is typically the pre-reform period. In our setting, we take the pre-marriage income in $t = -1$ and simulate the change in MTR spouses would have faced had they already been filing jointly in $t = -1$ ($MTR^{hyp}_{i,-1}$). Put differently, this is the MTR spouses would have faced in $t > -1$ if their income had not changed since $t = -1$. For this, we link spouses in the year before marriage. We describe how we do this in the data section below.

The second challenge is specific to our setting: the change in MTR from joint taxation is not exogenous. Spouses might actively decide to get married because of the lower tax liability they face when filing jointly. In addition, how much the MTR changes under joint taxation is also endogenous as it depends on the individual income and the within-household income share. Income and income share are likely correlated with unobservables that may impact spousal labor supply, such as gender norms or career preferences. We address both challenges by including fixed effects for bins of individual income and within-household income share.[16] The identifying assumption here is that while individual income and within-household income distribution can be endogenous, the exact income and share of income are quasi exogenous. To provide an illustrative example, this means an individual may select into a marriage in which they contribute about two-thirds of the household income. However, they cannot precisely anticipate whether their share of income will be 0.64 or 0.66 in $t = -1$ (and similarly for their absolute income). Here, we exploit the substantial heterogeneity in ΔMTR_i for similar income levels and shares that we show in Figure 6. In addition,

[16] Binning both is required because the exact income and income share perfectly define the MTR under joint taxation and would remove all variation.

having the full taxpayer population provides us with a large enough sample.

For the 2SLS estimation, we estimate the following equation separately for women and men (g = w, m) in a given event year t:

$$ln\left(\frac{y_{it}^g}{y_{i,-1}^g}\right) = \varepsilon_t^g ln\left(\frac{\widehat{1-MTR_{i,t}}}{1-MTR_{i,-1}}\right) + \sum_{t^{child1}} \gamma_{t^{child1}}^g + \theta_{income}^g + \rho_{share}^w + \delta_{age}^w + \zeta_{age}^m + \eta_{cal}^g + \epsilon_{it}^g, \tag{4}$$

with the first stage defined as:

$$ln\left(\frac{1-MTR_{i,t}}{1-MTR_{i,-1}}\right) = \lambda_t^g ln\left(\frac{1-MTR_{i,-1}^{hyp}}{1-MTR_{i,-1}}\right) + \sum_{t^{child1}} \gamma_{t^{child1}}^g + \theta_{income}^g + \rho_{share}^w + \delta_{age}^w + \zeta_{age}^m + \eta_{cal}^g + \epsilon_{it}^g. \tag{5}$$

We denote the gross labor income of individual i in event year t as y_{it}^g, with $y_{i,-1}^g$ referring to the pre-marriage income in $t = -1$.[17] Our parameter of interest is ε_t^g, the elasticity of labor income with respect to the net-of-tax rate for gender g. The net-of-tax rate for i is denoted as $1 - MTR_{i,t}$ for period t, and as $1 - MTR_{i,-1}$ for the pre-marriage base year. $1 - MTR_{i,-1}^{hyp}$ is the hypothetical net-of-tax rate i would have faced under joint taxation in $t = -1$. This is our instrument for $1 - MTR_{i,t}$.

We include fixed effects for women's labor income binned in 1,000-€-bins (θ_{income}^g) and for women's income share binned in 2-%-bins (ρ_{share}^w), both measured pre-marriage in $t = -1$ (we also replicate our results with other bin sizes). In addition, δ_{age}^w and ζ_{age}^m are fixed effects for the age of women and men in year t, respectively, and account for gender-specific life cycle trends. Calendar year fixed effects η_{cal} account for common time trends.

Lastly, we include event time dummies for time relative to the birth of the first child (t^{child1}) to account for the child penalty. Taken together, we combine the standard approach to empirically recover elasticity estimates from changes in MTR (e.g., Gruber and Saez 2002; Saez, Slemrod, and Giertz 2012; Neisser 2021) with our approach for estimating the marriage earnings gap.

[17] The standard approach is to estimate the elasticity of taxable income with respect to the marginal tax rate (e.g., Gruber and Saez 2002; Saez, Slemrod, and Giertz 2012). Our setting does not allow us to use taxable income since, under joint taxation, taxable income is only defined at the household level. To be able to estimate how individual income changes, we use gross labor income, which is observed on the individual level, both for singles and married individuals.

Data To estimate 4 and 5, we use a unique sample of linked spouses from administrative income tax data (TPP). The sample includes the universe of all taxpayers who got married between 2011 and 2016 and file their taxes jointly. The pivotal advantage of the data is that it yields precise information on spouses' incomes and tax liabilities already before marriage, allowing us to determine the exact MTR for singles and married spouses. Our outcome is gross labor earnings, which we define as the sum of gross wages, self-employed income, and business income. We restrict the sample to working age couples (age 18 to 60) for whom both spouses are observed at least three years before and after the year of marriage. We further drop couples if at least one spouse has zero taxable income in $t = -1$ or negative income at any t. Lastly, we require all individuals to have a stable income in the three years preceding their marriage. We drop couples if at least one spouse's income in any $t = -3, -2, -1$ deviates more than 25% from their three-year average income over that period. Having relatively constant income pre-marriage is likely to increase the salience of MTR changes under joint taxation and also mitigates concerns about mean reversion for the elasticity estimation. Table 2 provides descriptive statistics of our sample, and Figure A.9 visualizes the distribution of income, MTR, and MTR changes from joint taxation for the sample.

Results Table 4 reports the 2SLS elasticity estimates from estimating Equation 4 for women (for the first stage results see Table A.2a). Our preferred specification includes fixed effects for absolute female income (θ^w_{income}) and female share of income (ρ^w_{share}) pre-marriage (column 4).

The change in MTR under joint taxation has a significant effect on married women's labor income for all $t > 0$. The higher the increase in MTR, the more women reduce their earnings, which leads to a larger marriage earnings gap. The elasticity estimates range from 0.3 to 1.1 and increase over t. For $t = 0$, we do not find a significant effect in our preferred specification, which is likely explained by measurement error: The income tax data is annual and therefore does not inform about when exactly an individual gets married in $t = 0$. Income in $t = 0$ is thus only partly taxed under joint taxation. For men, the elasticity estimates are statistically insignificant and also economically small. Table A.2 provides the 2SLS results for men.

To quantify the impact of tax incentives for the marriage earnings gap, we provide a back-of-the-envelop calculation for the size of the earnings response to changes in MTR. We compute $\Delta y^g_t = -\varepsilon^g_t \cdot y^g_{-1} \cdot \Delta MTR^g_{-1}/(1 - MTR^g_{-1})$ for $g = women, men$,

Table 4: Joint Taxation – Elasticity Estimates for Women (2SLS)

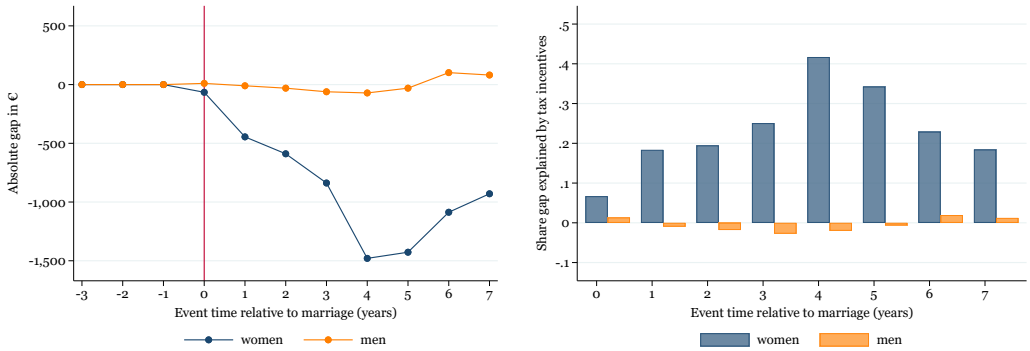
	(1) baseline	(2) income	(3) share	(4) both
t=0	0.18***	0.11***	-0.11*	0.05
	(0.02)	(0.02)	(0.04)	(0.05)
t=1	0.42***	0.49***	0.32***	0.34***
	(0.03)	(0.04)	(0.08)	(0.09)
t=2	0.74***	1.07***	0.72***	0.45***
	(0.04)	(0.05)	(0.10)	(0.11)
t=3	0.94***	1.55***	0.94***	0.64***
	(0.04)	(0.05)	(0.11)	(0.12)
t=4	0.98***	1.76***	1.43***	1.13***
	(0.05)	(0.06)	(0.14)	(0.15)
t=5	1.01***	1.95***	1.37***	1.09***
	(0.07)	(0.08)	(0.17)	(0.19)
t=6	0.98***	1.94***	1.05***	0.83***
	(0.09)	(0.11)	(0.23)	(0.25)
t=7	0.95***	1.73***	0.87**	0.71*
	(0.13)	(0.15)	(0.33)	(0.35)
Income (θ^w_{income})		✓		✓
Income share (ρ^w_{share})			✓	✓
Event time first child	✓	✓	✓	✓
Age (δ^w_{age}, ζ^m_{age})	✓	✓	✓	✓
Year (η_{cal})	✓	✓	✓	✓

Notes: Estimates for the elasticity of gross income with respect to the net-of-tax rate, from our 2SLS estimation of Equation 4 for women. Table A.2a reports first-stage results from estimating Equation 5, while Table A.2 reports the second-stage estimates for men. Each column shows the estimates from a different specification, with differences in the included fixed effects. Column (2) adds fixed effects for female income to the baseline model, binned in 1,000-€-bins, and (3) for the female share of income , using bins of 2 percentage points. Specification (4) is our preferred specification and includes fixed effects for both female income and female share of income. Each row shows the estimates for a different event time, with $t = 0$ denoting the year of marriage.

Significance level: *** 0.01; ** 0.05; * 0.1. Robust standard errors are reported in parentheses.

Figure 7: Pure tax Effect on Earnings

(a) Income Response (b) Share of Marriage Gap Explained



Notes: Annual gross labor income from TPP in 2015 Euros. Panel a shows the average earnings response to tax changes from joint taxation: $\Delta y_t^g = -\varepsilon_t^g \cdot y_{-1}^g \cdot \Delta MTR_{-1}^g/(1 - MTR_{-1}^g)$ computed for every t with parameters from Table 4, Table 2, and Table A.2. We include results for men, although their elasticity estimates are not statistically significant (Table A.2). Panel b shows the share of the total marriage earnings gap that is explained by the response to tax incentives. For this, we divide the earnings response (panel a) by the absolute marriage earnings gap estimated for this sample (see Figure A.10a). *Reading example*: In $t = 3$ women's income response to tax changes from joint taxation is -838€, which explains 25% of the marriage earnings gap in that period.

replacing all parameters on the right hand side with estimates from the data. We take the mean values for $t = -1$ as reported in Table 2 and the elasticity estimates from column 4 of Table 4. Figure 7 plots the absolute annual gross labor income response over event time. Averaging over $t > 0$, women reduce their annual gross labor income by -971€ because of higher MTR from joint taxation. This implies that tax incentives account for 26% of the total marriage earnings gap in this sample (Figure 7b). Our empirical finding is in line with results from structural estimations (Bick and Fuchs-Schündeln 2018; Borella, De Nardi, and Yang 2023), which have highlighted the role of tax incentives for spousal labor supply.

Tax incentives from joint taxation significantly contribute to the marriage earnings gap, but cannot fully explain it. In line with this result, we show that even primary earner women whose MTR decrease under joint taxation face a marriage earnings gap. At the same time, men do not decrease their earnings after marriage, even if they face a higher MTR under joint taxation. This is shown in Figure A.11, which plots the corresponding event study estimates for the two subsamples of spouses with primary and secondary earner women, respectively. Taken together, our results suggests that the responsiveness to tax incentives in the marriage context depends on gender (norms) – a mechanism we exploit in more detail in subsection 4.3.

4.2 Indirect Labor Supply Incentives from Ex-Spousal Alimony

Mechanism With ex-spousal alimony, a reduction in earnings during marriage has a smaller impact on the expected lifetime income of secondary earners. Marriage then acts as an insurance mechanism against the financial risk of reducing earnings. The risk of reducing earnings during marriage and its subsequent consequence of lower post-divorce earnings is mitigated by the future income support provided by the other spouse. In that way, alimony payments are an indirect labor disincentive only coming into effect in the case of divorce.

Testing the Mechanism To test whether spousal earnings respond to this indirect labor supply incentive, we exploit exogenous variation in alimony eligibility. In 2008, the German government introduced a major divorce law reform, cutting the legal entitlement for ex-spousal alimony payments and, thus, financial obligations post-divorce. The reform marked a shift in how court decisions on alimony were made, increasing conditions to meet alimony requirements while shortening payment periods. Besides changing the general understanding of post divorce obligations, the reform also changed specific rules for alimony eligibility. Before the reform, divorced resident parents are eligible for alimony payments until their youngest child turns 15. Post reform, eligibility applies as long as the youngest child is three years old at most. From then on, the resident parent is expected to provide for their living expenses by working full time. The change in the alimony payments is depicted in Figure A.12.

The reform, announced in 2006 (link first draft), enacted in December 2007, and implemented from January 2008, impacted all marriages: future marriages contracted after reform, existing marriages contracted pre reform but not (yet) divorced, and past marriages contracted and divorced pre reform.[18]

The reform introduces exogenous variation in the insurance value of marriages with heterogeneity in the timing of this shock. Determined by the year of marriage, spouses have more or less (or no) leeway to adjust their earnings when learning about the reduced insurance value. We exploit this natural experiment to analyze the causal impact of alimony on spousal earnings during marriage, comparing the marriage earnings gap across different year-of-marriage cohorts. We first estimate our main event study specification from Equation 2 for each cohort separately to compare how the marriage earnings gap evolves around reform.

[18]For already divorced marriages, alimony payments are not adjusted automatically post reform but have to be renegotiated.

After providing graphical evidence, we then estimate a modified version of Equation 2 for all cohorts together, adding a dummy indicating whether an individual is treated by the reform. We allow the alimony regulation to affect individuals in two ways. First, we analyze the effect of the alimony regime in place at the time of marriage ($t = 0$). Treatment is a function of the marriage cohort: individuals are treated by the reform if they get married post-reform. To estimate this effect, we extent Equation 2 by interacting the event time dummies with a dummy treated marriage cohort$_i$, which is 1 for individuals who get married ≥ 2008, and 0 else:

$$\begin{aligned} Y_{it}^{g} = &\sum_{j \neq r} \beta_j^g \cdot \mathbb{1}[j = t] + \sum_{j \neq r} \delta_j^g \cdot \text{treated marriage cohort}_i \times \mathbb{1}[j = t] \\ &+ \sum_k \gamma_k^{1,g} \cdot \mathbb{1}[k = t^{child1}] + \zeta_{cal}^g + \eta_{age}^g + \theta_c^g + \upsilon_{it}^g. \end{aligned} \tag{6}$$

Second, we analyze the effect of the status quo alimony regime in place during a given period on the labor supply in that same period. Treatment is then a function of calendar time: individuals are treated in the calendar years ≥ 2008. To estimate this effect, we extend Equation 2 by interacting the event time dummies with a dummy treated calendar year$_t$, which is 1 if the individual is observed in calendar years ≥ 2008, and 0 else:

$$\begin{aligned} Y_{it}^{g} = &\sum_{j \neq -1} \beta_j^g \cdot \mathbb{1}[j = t] + \sum_{j \neq -1} \delta_j^g \cdot \text{treated calendar year}_t \times \mathbb{1}[j = t] \\ &+ \sum_k \gamma_k^{1,g} \cdot \mathbb{1}[k = t^{child1}] + \zeta_{cal}^g + \eta_{age}^g + \theta_c^g + \upsilon_{it}^g. \end{aligned} \tag{7}$$

The coefficient of interest for Equation 6 and Equation 7 is δ_j^g, which can be interpreted as a difference-in-marriage-gap estimate. δ_j^g captures the difference between event time estimates at event time $t = j$ of individuals treated by the reform versus those who are not, following the respective definition: *married post reform* for Equation 6 and *observed post reform* for Equation 7. Event time t is measured in years relative to the event of marriage, omitting the reference period $r = -1$, i.e., the year prior to the year of marriage. Because of a limited number of available variables, we only account for event time with respect to the first child. We denote η_{age}^g, ζ_{cal}^g, and θ_c^g as fixed effects for individuals' age, calendar year, and year-of-marriage cohort, respectively.[19]

[19]Note that we include treated marriage cohort$_i$ and treated calendar year$_t$ only as interaction

Data As in subsection 4.1, we use administrative income tax data (TPP) for the full population of German taxpayers. Since we want to analyze how the marriage earnings gap changes around the 2008 reform, we restrict the sample to newlyweds who married between 2004 and 2012, i.e., +/- 4 years around the reform. We further restrict the sample to individuals aged 18 or older, who are younger than 40 when they get married. The sample contains individuals who got a divorce, but we drop their post-divorce periods. We proceed the same way with individuals who became widowed. We further restrict the sample to individuals we observe at least three years before or after marriage.[20] Table 3 illustrates the sample descriptives.

In the TPP, women are typically added to their husbands' tax records upon joint taxation. Since German tax authorities introduced individual identifiers only in 2010, the TPP data does not allow for a linkage of women across pre-marriage and marriage years for the years around the reform. We overcome this hurdle by defining single filing women aged 18 to 39 who drop out of the panel data entirely as an indication of marriage. We then define event time relative to the drop-out, with the last observed year defining $t = -1$. Hence, we have a panel of men who are linked over time, and women for whom we observe pre- and post-marriage periods but cannot link both. Note that our event study specification (see Equation 2) does not require an individual panel identifier, allowing us to estimate the marriage earnings gap even when we cannot link individuals across event time. Kleven (2023) shows that child penalties can be estimated on repeated cross-sectional data. In our case, the data is not a repeated cross-section but actual panel data that lacks uninterrupted individual identifiers for women.

We provide descriptive statistics for the sample of women who drop out of the data (what we define as $t = -1$) and the sample of women who appear in $t = 0$ on their husbands' tax records. Table A.3 shows that both samples have similar characteristics with respect to observable characteristics. The samples are not identical for two reasons. First, we do not observe all women in $t = -1$, because not all women pay income taxes in the year before marriage and/or file an income tax return.[21] Second, dropouts are not necessarily an indicator of marriage: Women may stop paying taxes (or stop filing).

terms since they are perfectly collinear with θ_c^g and ζ_{cal}^g, respectively.

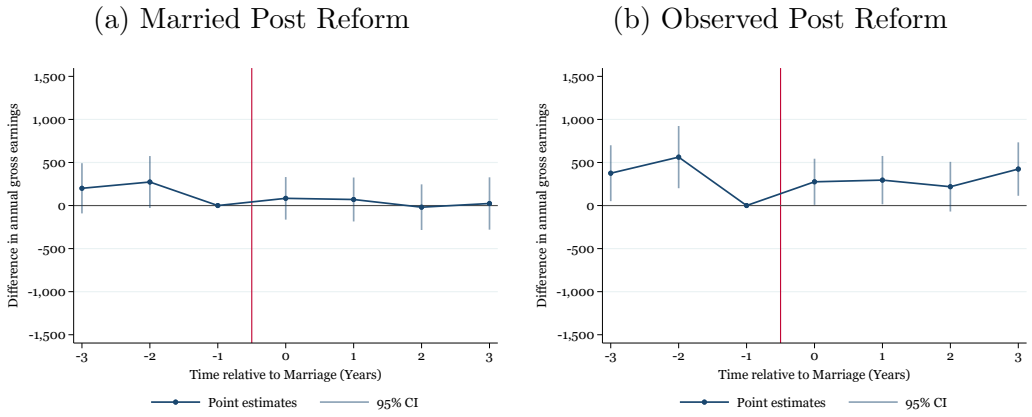
[20]Ideally, we would want to restrict the sample to individuals we can observe three years before *and* after marriage. However, the data structure only allows for linking periods $t < 0$ and $t >= 0$ for men, which is why we implement the weaker restriction of observing individuals at least three years before *or* after marriage. We explain the data structure in more detail in the following paragraph.

[21]Before 2012, non-filers are not included in the TPP data. For most German taxpayers, filing is optional, and many do not choose to file (Hauck and Wallossek 2023).

Results Figure A.13 compares MG_t^w estimates for women across different year-of-marriage cohorts around the reform in 2008. Comparing the MG_t^w cohort-by-cohort allows us to differentiate between two different effects. First, we show that married women's earnings are not affected by the alimony regime in place at the time of marriage, at $t = 0$: There is no difference between the earnings trajectories of women married before the reform (2006 and 2007) and those who married after the reform (2008 and 2009).

Second, we show that women's earnings do not respond to a change in the alimony regime at later points of their marriage, at $t > 0$. Figure A.14 provides pairwise comparisons for all marriage cohorts from 2004 to 2012. Panels A.14a to A.14c compare pre-reform marriage cohorts. If already married women increased their labor supply as a response to the reform, we would expect to see a change in MG_t^w for post reform t, i.e., for $t \geq$ the respective vertical line. For these cohorts, we can also rule that selection into marriage is affected by the reform, simply because they married before the reform is enacted (and even before it is announced for cohorts 2004-2006).

Figure 8: Difference Estimates



Notes: Annual data from TPP. This figure plots the difference estimates for women (δ_y^w) from Equation 6 (panel a) and Equation 7 (panel b). Panel a shows how annual gross labor earnings change at given event time t for women who married post reform, compared to those who married pre reform. Panel b plots the difference for women who, in given period t, are observed in a post-reform calendar year, compared to those observed pre reform.

The regression results from estimating Equation 6 confirm the graphical evidence from Figure A.14. The coefficients for the interaction terms δ_y^w are not statistically different from 0 when comparing women who married post- vs. pre-reform (Figure 8a), suggesting that women do not adjust their earnings in response to the reform. We also find no clear effect when comparing women who are observed post- vs. pre-reform (Equation 7, results are plotted in Figure 8b). Our findings are in line with results

from Bredtmann and Vonnahme (2019), who also do not find an effect of the alimony reform on spousal labor supply, time-use, or divorce probability using survey data from SOEP.

4.3 Gender Norms

Mechanism After studying the role of institutions, we now move to another potential mechanism: gender norms. Norms have shown to be an important driver behind gender gaps in different domains (e.g., Bertrand 2011). They may also explain (part of) the marriage earnings gap when impacting spousal labor supply. We want to test whether the marriage earnings gap is larger under more traditional gender norms.

Testing the Mechanism We first document the correlation between the marriage earnings gap and gender norms. For this descriptive analysis, we use responses to value-based questions from the SOEP to split our sample into individuals with progressive vs. traditional gender norms. We then estimate Equation 2 for each subsample. The obvious caveat of this analysis is that norms are likely to be formed endogenously.

To address this, we apply a second analysis where we exploit a natural experiment with quasi-exogenous variation in gender norms: the division of East and West Germany after World War II. From 1945 until 1990, two German countries existed: the German Democratic Republic (GDR, East Germany) and the Federal Republic of Germany (FRG, West Germany). Under the communist regime in the East, female labor market participation was high, going along with the norm of dual-earner spouses. In the capitalist West, women were less likely to work and the country had a strong male breadwinner norm.[22] East-West differences have been shown to persist even long after reunification. This is true for many dimensions, including preferences and norms (Alesina and Fuchs-Schündeln 2007; Boelmann, Raute, and Schönberg 2021).

To disentangle differences in culture from differences in legal institutions, we compare individuals who were born when Germany was divided but got married after reunification. They grew up in different countries with different norms and cultures, but at the time of marriage, they faced identical legal institutions with identical financial incentives, as discussed in subsection 4.1 and 4.2. We estimate a modified version of Equation 2,:

[22]For details on differences in gender norms and culture in the two countries, see e.g., Boelmann, Raute, and Schönberg (2021).

$$
\begin{aligned}
Y_{it}^{g} = & \sum_{j \neq r} \beta_j^g \cdot \mathbb{1}[j = t] + \sum_{j \neq r} \delta_j^g \cdot \mathbb{1}[j = t] \times east_i \\
& + \sum_{C=1}^{3} \sum_{k} \gamma_k^{C,g} \cdot \mathbb{1}[k = t^{childC}] + \sum_{C=1}^{3} \sum_{k} \lambda_k^{C,g} \cdot \mathbb{1}[k = t^{childC}] \times east_i \\
& + \kappa_{i,-1}^{g} + \phi_{i,-1}^{w} + \zeta_{cal}^{g} + \eta_{age}^{g} + \epsilon_{it}^{g}.
\end{aligned} \tag{8}
$$

The main difference to Equation 2 is that we interact event time dummies with a dummy variable $east_i$ equal to 1 if i was born in East Germany and 0 if i was born in West Germany. We add the interaction for all events, i.e., marriage as well as the birth of the first, second, and third child. This allows us to account for differences in the child penalties between East and West (Boelmann, Raute, and Schönberg 2021; Jessen 2022). In addition, we add fixed effects for pre-marriage income and female income share, measured in the year before marriage ($t = -1$), and denoted by $\kappa_{i,-1}^{g}$ and $\phi_{i,-1}^{w}$, respectively.[23] This accounts for underlying differences in income and gender gaps between East and West.

Data We use the same SOEPRV data we use for our main analysis in section 3. We measure individual gender norms with a set of value-based questions from the SOEP. These questions contain statements on gender roles that respondents are asked to agree or disagree with, based on a 1-7 Likert scale (see subsection A.4 for the questions used). We classify an individual as having traditional gender norms if their average response to these questions is < 4, and as progressive if their average is > 4.

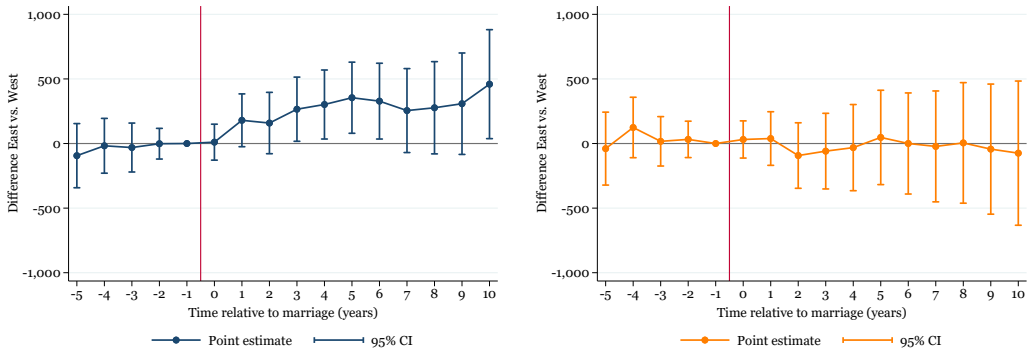
For our East/West analysis, we impose two additional sample restrictions: First, we restrict the sample to individuals who were born in Germany < 1989 and get married > 1990. Since the place of birth is not reported in the administrative data, we rely on information from survey data here: SOEP respondents are asked about their place of residence in 1989. We take this as a proxy for the place of birth. We exclude individuals who were born abroad. For individuals who did not answer the survey question on their place of residence in 1989, we proxy for the place of birth by using the first reported spell in the administrative data, an approach that has been established in the context of German administrative data (see e.g., Boelmann, Raute, and Schönberg 2021; Findeisen, Lee, Porzio, and Dauth 2021). Second, we

[23] For absolute earnings ($\kappa_{i,-1}^{g}$), we bin gross earnings in $t = -1$ in decentiles of the income distribution for each gender g: $x = 0$, $0 < x \leq p10$, $p10 < x \leq p20$, ..., $p90 < x \leq p100$. For the female earnings share ($\phi_{i,-1}^{w}$), we include these bins: $x = 0$, $0 < x \leq 0.1$, $0.1 < x \leq 0.2$, $0.2 < x \leq 0.3$, $0.3 < x \leq 0.45$, $0.45 < x \leq 0.55$, $0.55 < x \leq 0.75$, $0.75 < x < 1$, and $x = 1$.

include only individuals for whom we observe their spouses in the SOEPRV data. This restriction is necessary because we include fixed effects for the within-couple income share in $t = -1$.

Results Splitting the sample into individuals with traditional vs. progressive gender norms, we find a clear correlation between norms and post-marriage earnings. Figure A.16 plots the results from estimating Equation 2. The marriage earnings gap for individuals with progressive gender norms is 10% (Figure A.16a), while the gap is 33% for those with traditional gender norms (Figure A.16b).

Figure 9: The Marriage Earnings Gap – East vs West



(a) Women (b) Men

Notes: Annual data from SOEP–RV. This figure plots the $\hat{\delta}_j^g$ estimates from Equation 8 over event time. This is the difference in earnings for East vs West that is attributed to the given event time $t = j$. Panel a plots the results for women ($\hat{\delta}_j^w$) while panel b plots the results for men ($\hat{\delta}_j^m$). Estimates account for the child penalty and allow child penalty estimates to differ between East an West.

These descriptive findings are supported by the results from leveraging quasi-exogenous variation in gender norms from the German separation. Figure 9 plots the $\hat{\delta}_j^g$ estimates from Equation 8, i.e., the difference in earnings between East and West that is specific to event time $t = j$ and adds to the overall East-West gap. For men, the earnings difference between those born in the East vs West does not change with event time (Figure 9b). For women, the East-West gap is stable for pre-marriage periods as well (Figure 9a). However, for $t > 0$, earnings for women born in the East are relatively higher. Put differently, women born in the West, with more traditional gender norms, reduce their earnings more after marriage than do women born in the East. Figure A.17 plots the marriage earnings gap estimates for East and West. Women born in East Germany face no significant drop in earnings after marriage,

while women born in the West face a marriage earnings gap of 30%. These results strongly support the relevance of gender norms for the marriage earnings gap.

5 Conclusion

In this paper, we show that there is a marriage earnings gap: Women's earnings drop after marriage, and this is not explained by the child penalty. The marriage earnings gap results from the extensive and intensive margin of labor supply: Married women are less likely to work and, conditional on working, work fewer hours, but they are not paid lower wages. Different from children, marriage is a social construct. In modern societies, it is also a legal contract with rules defined by the government. Policy makers can use marriage as a public policy tool to (dis-)incentivize spousal labor supply. We find that labor supply incentives significantly contribute to the marriage earnings gap: A back-of-the envelop calculation suggests that increased marginal tax rates under joint taxation explain 26% of the gap for women. Given that child-related policies have been shown to have only very limited impact on the child penalty (Kleven, Landais, Posch, Steinhauer, and Zweimüller 2023), this is an important insight for policy makers. However, our findings also highlight that not all labor supply incentives from marriage related policies are equally effective. We do not find an effect on earnings from a divorce law reform that aimed at increasing married women's labor supply by reducing their alimony eligibility. Our results also emphasize the relevance of gender norms: Women who grew up in East Germany with more progressive gender norms experience a smaller marriage earnings gap.

References

Alesina, Alberto and Nicola Fuchs-Schündeln (2007). "Good-bye Lenin (or Not?): The Effect of Communism on People's Preferences." *American Economic Review* 97.4, 1507–1528. DOI: 10.1257/aer.97.4.1507.

Andresen, Martin and Emily Nix (2021). "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples." *Journal of Labor Economics* Forthcoming. Publisher: The University of Chicago Press. DOI: 10.1086/718565.

Angelov, Nikolay, Per Johansson, and Erica Lindahl (2016). "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34.3, 545–579. DOI: 10.1086/684851.

Artmann, Elisabeth, Nicola Fuchs-Schündeln, and Giulia Giupponi (2023). *Forward-Looking Labor Supply Responses to Changes in Pension Wealth: Evidence from Germany*. SSRN Scholarly Paper. DOI: 10.2139/ssrn.4447094.

Bartels, Charlotte and Cortnie Shupe (2023). "Drivers of participation elasticities across Europe: gender or earner role within the household?" *International Tax and Public Finance* 30.1, 167–214. DOI: 10.1007/s10797-021-09711-z.

Becker, Gary S. (1985). "Human Capital, Effort, and the Sexual Division of Labor." *Journal of Labor Economics* 3.1, 33–58.

– (1991). *A treatise on the family*. Enl. ed. Cambridge, Mass: Harvard University Press.

Berniell, Inés, Lucila Berniell, Dolores de la Mata, María Edo, Yarine Fawaz, Matilde P. Machado, and Mariana Marchionni (2022). "Motherhood, pregnancy or marriage effects?" *Economics Letters* 214, 110462. DOI: 10.1016/j.econlet.2022.110462.

Berresheim, Ursula and David Koll (2023). *Staying Together Forever? Life-Cycle Effects of Overoptimistic Couples*.

Bertrand, Marianne (2011). "New Perspectives on Gender." *Handbook of Labor Economics*. Vol. 4. Elsevier, 1543–1590. DOI: 10.1016/S0169-7218(11)02415-4.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz (2010). "Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors." *American Economic Journal: Applied Economics* 2.3, 228–255. DOI: 10.1257/app.2.3.228.

Bick, Alexander and Nicola Fuchs-Schündeln (2018). "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis." *The Review of Economic Studies* 85.3, 1543–1576. DOI: 10.1093/restud/rdx057.

Blau, Francine D. and Lawrence M. Kahn (2007). "Changes in the Labor Supply Behavior of Married Women: 1980–2000." *Journal of Labor Economics* 25.3. Publisher: The University of Chicago Press, 393–438. DOI: 10.1086/513416.

– (2017). "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55.3, 789–865. DOI: 10.1257/jel.20160995.

Boelmann, Barbara, Anna Raute, and Uta Schönberg (2021). *Wind of Change? Cultural Determinants of Maternal Labor Supply*.

Boneva, Teodora, Marta Golin, Katja Kaufmann, and Christopher Rauh (2022). *Beliefs About Maternal Labor Supply*. IZA DP No. 15788.

Borella, Margherita, Mariacristina De Nardi, and Fang Yang (2023). "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labour Supply?" *The Review of Economic Studies* 90.1, 102–131. DOI: 10.1093/restud/rdac018.

Bredtmann, Julia and Christina Vonnahme (2019). "Less money after divorce – how the 2008 alimony reform in Germany affected spouses' labor supply, leisure and marital stability." *Review of Economics of the Household* 17.4, 1191–1223. DOI: 10.1007/s11150-019-09448-z.

Brugiavini, Agar, Cristina E. Orso, Mesfin G. Genie, Rinaldo Naci, and Giacomo Pasini (2022). *SHARE Job Episodes Panel*. DOI: 10.6103/SHARE.JEP.800.

Chiappori, Pierre-Andre, Murat Iyigun, Jeanne Lafortune, and Yoram Weiss (2017). "Changing the Rules Midway: The Impact of Granting Alimony Rights on Existing and Newly Formed Partnerships." *The Economic Journal* 127.604, 1874–1905. DOI: 10.1111/ecoj.12385.

Christl, Michael, Silvia De Poli, and Viginta Ivaškaitė-Tamošiūnė (2023). "Does it pay to say 'I do'? Marriage bonuses and penalties across the EU." *Journal of European Social Policy* 33.3, 317–336. DOI: 10.1177/09589287231159492.

Eurostat (2023). *Gender overall earnings gap*.

Findeisen, Sebastian, Sang Yoon (Tim) Lee, Tommaso Porzio, and Wolfgang Dauth (2021). *Transforming Institutions. Labor Reallocation and Wage Growth in a Reunified Germany*.

Foerster, Hanno (2022). *Untying the Knot: How Child Support and Alimony Affect Couples' Dynamic Decisions and Welfare*. Working Paper.

Giommoni, Tommaso and Enrico Rubolino (2022). *The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit*. Dondena Working Paper No. 153.

Goebel, Jan, Markus M. Grabka, Carsten Schröder, Hannah Penz, Tatjana Mika, Daniel Brüggmann, Sebastian Ellert, Katharina Werhan, Forschungsdatenzentrum Der Rentenversicherung, and Deutsches Institut Für Wirtschaftsforschung (DIW Berlin) (2022). *SOEP-RV: Linking German Socio-Economic Panel Data to Pension Records*. DOI: 10.5684/SOEP.V37-RV.VSKT2020.

González, Libertad and Berkay Özcan (2013). "The Risk of Divorce and Household Saving Behavior." *Journal of Human Resources* 48.2, 404–434. DOI: 10.1353/jhr.2013.0013.

Gruber, Jon and Emmanuel Saez (2002). "The elasticity of taxable income: evidence and implications." *Journal of Public Economics* 84.1, 1–32. DOI: 10.1016/S0047-2727(01)00085-8.

Hauck, Tobias and Luisa Wallossek (2023). "Optional (Non-) Filing and Effective Taxation."

Hermle, Johannes and Andreas Peichl (2018). *Jointly Optimal Taxes for Different Types of Income*. DOI: 10.2139/ssrn.3275422.

Hill, Martha S. (1979). "The Wage Effects of Marital Status and Children." *The Journal of Human Resources* 14.4. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], 579–594. DOI: 10.2307/145325.

Jessen, Jonas (2022). "Culture, children and couple gender inequality." *European Economic Review* 150, 104310. DOI: 10.1016/j.euroecorev.2022.104310.

Juhn, Chinhui and Kristin McCue (2016). "Evolution of the Marriage Earnings Gap for Women." *The American Economic Review: Papers & Proceedings* 106.5, 252–256. DOI: 10.1257/aer.p20161120.

– (2017). "Specialization Then and Now: Marriage, Children, and the Gender Earnings Gap across Cohorts." *Journal of Economic Perspectives* 31.1, 183–204. DOI: 10.1257/jep.31.1.183.

Kalíšková, Klára (2014). "Labor supply consequences of family taxation: Evidence from the Czech Republic." *Labour Economics* 30, 234–244. DOI: 10.1016/j.labeco.2014.04.010.

– (2020). "Tax and transfer policies and the female labor supply in the EU." *Empirical Economics* 58.2, 749–775. DOI: 10.1007/s00181-018-1544-2.

Killewald, Alexandra and Margaret Gough (2013). "Does Specialization Explain Marriage Penalties and Premiums?" *American sociological review* 78.3, 477–502. DOI: 10.1177/0003122413484151.

Killewald, Alexandra and Ian Lundberg (2017). "New Evidence Against a Causal Marriage Wage Premium." *Demography* 54.3, 1007–1028. DOI: 10.1007/s13524-017-0566-2.

Kleven, Henrik (2023). *The Geography of Child Penalties and Gender Norms: Evidence from the United States*. NBER Working Paper 30176 (Revised Version, Issue Date: June 2022). DOI: 10.3386/w30176.

Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante (2024). "The Child Penalty Atlas." *Review of Economic Studies* Forthcoming. DOI: 10.3386/w31649.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller (2019). "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109, 122–126. DOI: 10.1257/pandp.20191078.

– (2023). "Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation." *American Economic Journal: Economic Policy* forthcoming.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard (2019). "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11.4, 181–209. DOI: 10.1257/app.20180010.

Koch, Leonie (2024). *Income Taxation of Couples and Gender (In)equality*.

Korenman, Sanders and David Neumark (1991). "Does Marriage Really Make Men More Productive?" *The Journal of Human Resources* 26.2. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System], 282–307. DOI: 10.2307/145924.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington (2018). *The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?* NBER Working Paper 24740. DOI: 10.3386/w24740.

LaLumia, Sara (2008). "The effects of joint taxation of married couples on labor supply and non-wage income." *Journal of Public Economics* 92.7, 1698–1719. DOI: 10.1016/j.jpubeco.2008.01.009.

Lüthen, Holger, Carsten Schröder, Markus M. Grabka, Jan Goebel, Tatjana Mika, Daniel Brüggmann, Sebastian Ellert, and Hannah Penz (2021). "SOEP-RV: Linking German Socio-Economic Panel Data to Pension Records." *Jahrbücher für Nation-*

alökonomie und Statistik. Publisher: De Gruyter Oldenbourg. DOI: 10.1515/jbnst-2021-0020.

Neisser, Carina (2021). "The Elasticity of Taxable Income: A Meta-Regression Analysis." *The Economic Journal*. DOI: 10.1093/ej/ueab038.

Pilossoph, Laura and Shu Lin Wee (2021). "Household Search and the Marital Wage Premium." *American Economic Journal: Macroeconomics* 13.4, 55–109. DOI: 10.1257/mac.20180092.

Rangel, Marcos A. (2006). "Alimony Rights and Intrahousehold Allocation of Resources: Evidence from Brazil." *The Economic Journal* 116.513. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0297.2006.01104.x, 627–658. DOI: 10.1111/j.1468-0297.2006.01104.x.

Research Data Centre of the Federal Statistical Office and the statistical offices of the Länder (RDC) (2022). *Taxpayer-Panel 2001-2018, On-Site, Version 1*. DOI: 10.21242/73111.2018.00.01.1.1.1.

Research Data Centre of the German Pension Insurance (FDZ-RV) (2022). *Codeplan Scientific Use File SOEP-RV VSKT 2020*.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz (2012). "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review." *Journal of Economic Literature* 50.1, 3–50. DOI: 10.1257/jel.50.1.3.

Selin, Håkan (2014). "The rise in female employment and the role of tax incentives. An empirical analysis of the Swedish individual tax reform of 1971." *International Tax and Public Finance* 21.5, 894–922. DOI: 10.1007/s10797-013-9283-y.

Siminski, Peter and Rhiannon Yetsenga (2022). "Specialization, Comparative Advantage, and the Sexual Division of Labor." *Journal of Labor Economics* 40.4, 851–887. DOI: 10.1086/718430.

Stevenson, Betsey (2007). "The Impact of Divorce Laws on Marriage-Specific Capital." *Journal of Labor Economics* 25.1, 75–94. DOI: 10.1086/508732.

– (2008). "Divorce Law and Women's Labor Supply." *Journal of Empirical Legal Studies* 5.4, 853–873. DOI: 10.1111/j.1740-1461.2008.00143.x.

United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Marriage Data 2019*.

Voena, Alessandra (2015). "Yours, Mine, and Ours: Do Divorce Laws Affect the Intertemporal Behavior of Married Couples?" *American Economic Review* 105.8, 2295–2332. DOI: 10.1257/aer.20120234.

A Appendix

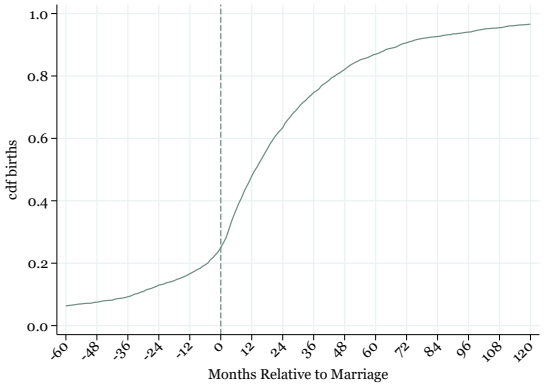
A.1 Additional Figures and Tables

Figure A.1: Timing Birth First Child Relative to Marriage

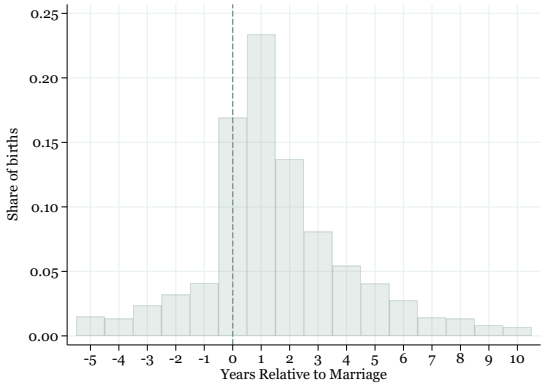
(a) Share of Births – Monthly Event Time



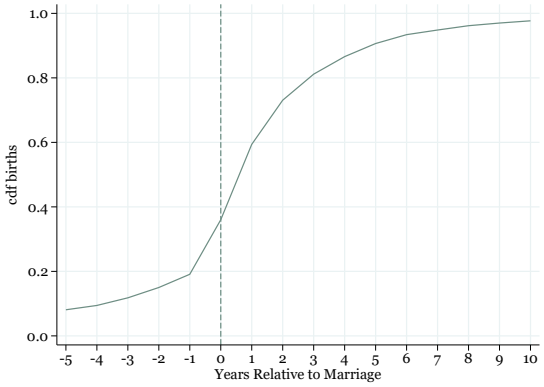
(b) CDF of Births – Monthly Event Time



(c) Share of Births – Yearly Event Time



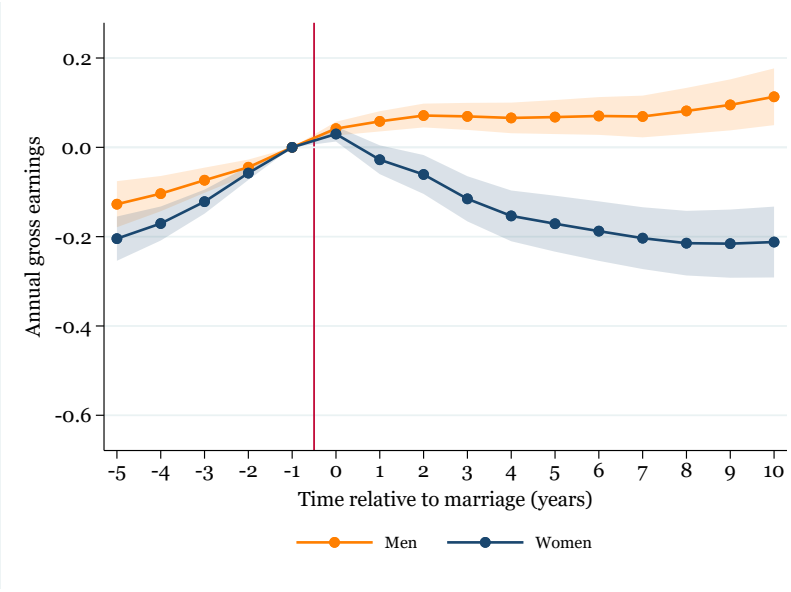
(d) CDF of Births – Yearly Event Time



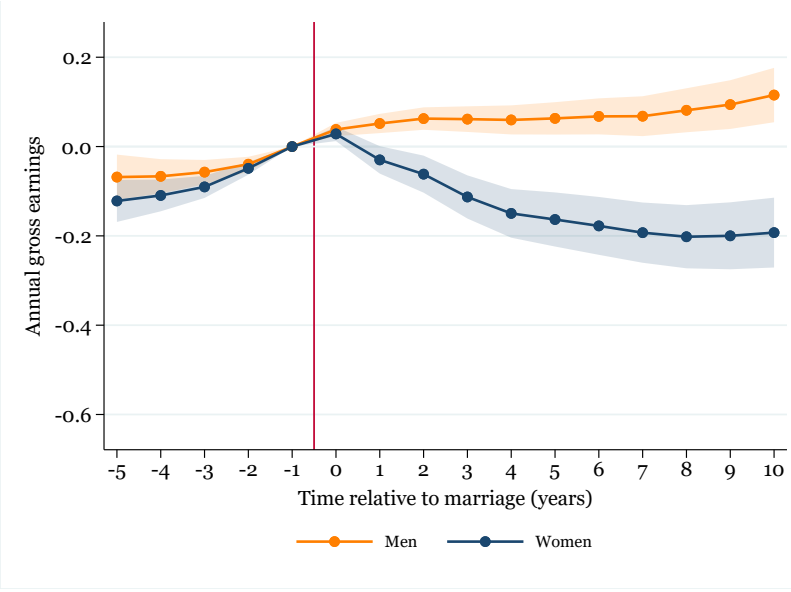
Notes: Data from SOEPRV. All panels visualize the distribution of the timing between birth of the first child and marriage for women who get married and have at least one child at some point in their (observed) life. The x-axis denotes time relative to marriage. Women who have their first child before marriage are on the left side of the dashed line, women who have their first child after marriage are on the right side. Panel a and c plot the share of births in a given period and panel b and d plot the corresponding cumulative density function (CDF). Panel a and b show monthly data for the subsample where we observe the exact month of marriage and childbirth. Panel c and d show yearly data for the full sample. *Reading example*: Panel a shows that the most common month (mode) to give birth to the first child is 4 months after marriage, implying that these women get married at about 5 months pregnant. Panel b shows that 24% of women have their first child before the month of marriage.

Figure A.2: Pre-Marriage Income Growth and Labor Market Experience

(a) Main Specification



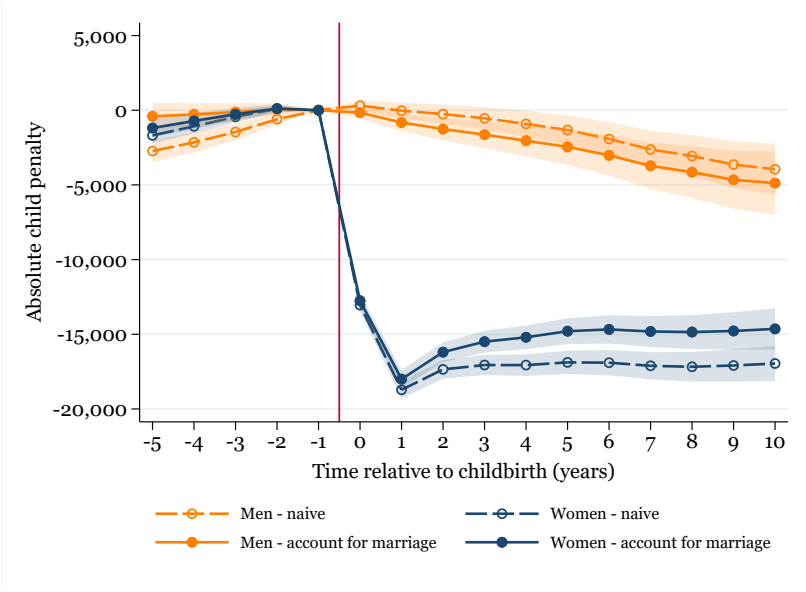
(b) Account for Career Age



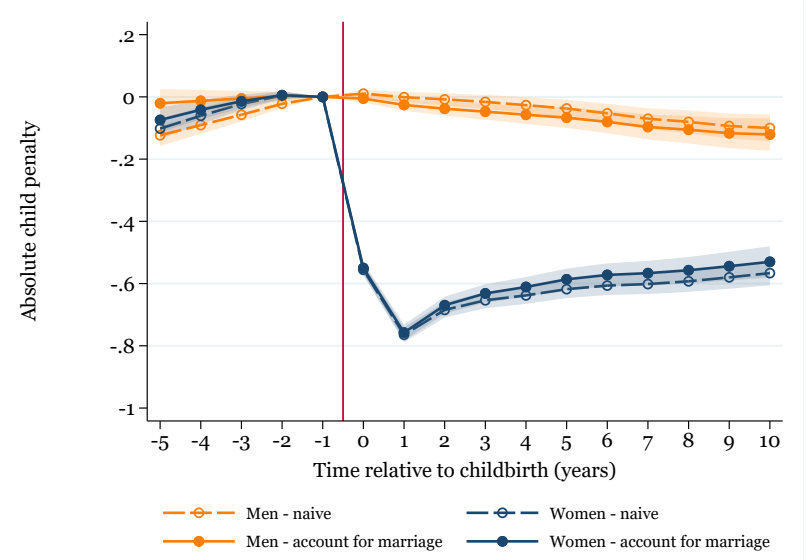
Notes: Annual data from SOEPRV. Figure plots MG_t^g for gender g over event time t from estimating Equation 2 with 95% confidence intervals. Panel a shows results for the main specification. Panel b shows results from estimating a modified version of 2 that include fixed effects for the number of years since the first employment of individual i (including regular employment and self-employment).

Figure A.3: Child Penalties Accounting for Marriage

(a) Absolute Estimates

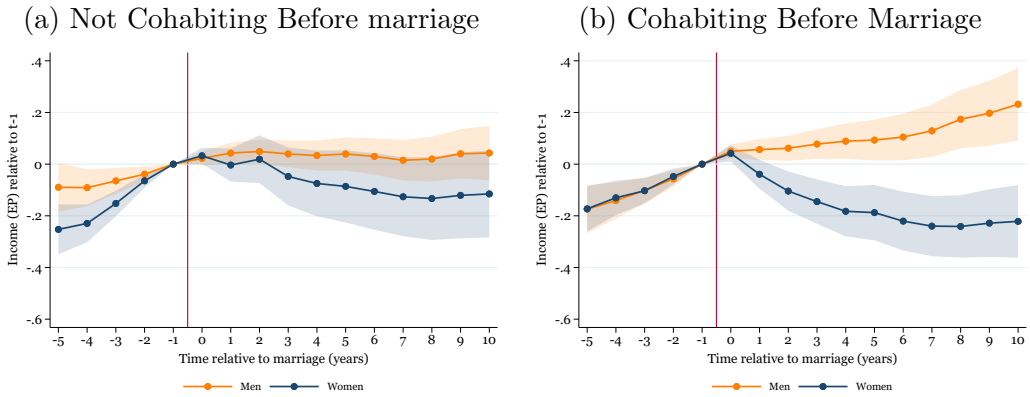


(b) Relative Gap Estimates



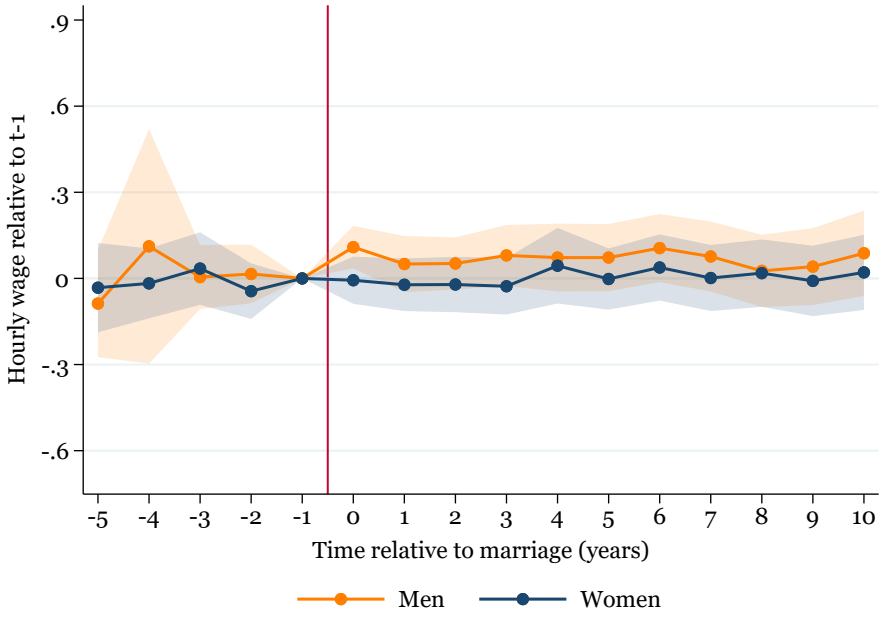
Notes: Annual data from SOEPRV. This figure shows the event study estimates from estimating the child gap (or child penalty) in monthly gross earnings. Plots show point estimates and 95% confidence intervals. Dashed lines plot naive estimations from estimating a modified version of Equation 1 but over event time relative to the birth of the first child. Solid lines plot estimations from estimating a modified version of Equation 2 but over event time relative to the birth of the first child while also including event time dummies for marriage. Panel a shows the absolute estimates $\hat{\beta}_t^g$ in 2015 Euros. Panel b shows relative earnings gap estimates for CG_t^g. The sample is not the same as for our main analysis: every individual in the sample here is a parent, but not all a married. Whereas in our main sample it's the opposite: everyone is married, but not all are parents.

Figure A.4: Marriage Earnings Gap by Cohabitation Status Before Marriage



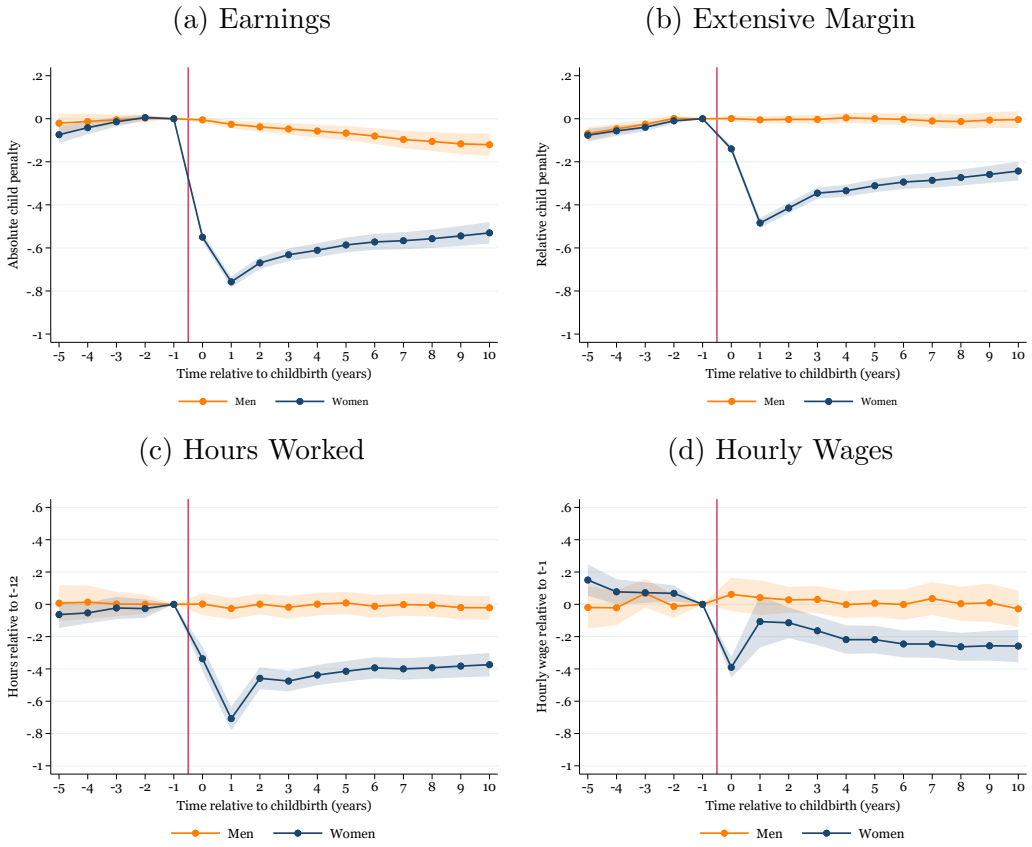
Notes: Annual data from SOEPRV. This figure shows the event study estimates from estimating the marriage gap in yearly gross earnings. Plots show point estimates and 95% confidence intervals. Lines plot estimations from estimating Equation 2. Panel a shows the result for spouses who lived together one year before marriage and Panel b for spouses who lived separately. The sample is not the same as for our main analysis: every individual in the sample here has survey information on the relationship status prior to marriage.

Figure A.5: Hourly Wages – Survey Data Only



Notes: Annual data from SOEP. This figure shows the event study estimates from estimating Equation 2 for hourly wages. Hourly wages defined as gross income divided by time spent on paid work in the labor market. Both income and time-use data stem from survey responses in the SOEP data. This is the same analysis as shown in Figure 3d but with income from the survey data (same source as hours) rather than from administrative data. The sample is also the same, except for individuals for whom income is reported in the administrative data but not in the survey data (or vice versa).

Figure A.6: Decomposition Child Penalty



Notes: Data from SOEPRV. This figure decomposes the overall child penalty (panel a) into extensive margin (panel b), intensive margin (panel c), and hourly wages (panel d). All panels plot MG_t^g estimates for women and men from estimating a modified version of Equation 2 but over event time relative to the birth of the first child, while also including event time dummies for marriage. For all panels, event time is defined in years. We exclude $t = -1$, i.e., the year before the year of child birth. Panel c plots results for the intensive margin only, conditional on working in a given t. Hourly wages in panel d are defined as earnings divided by hours worked.

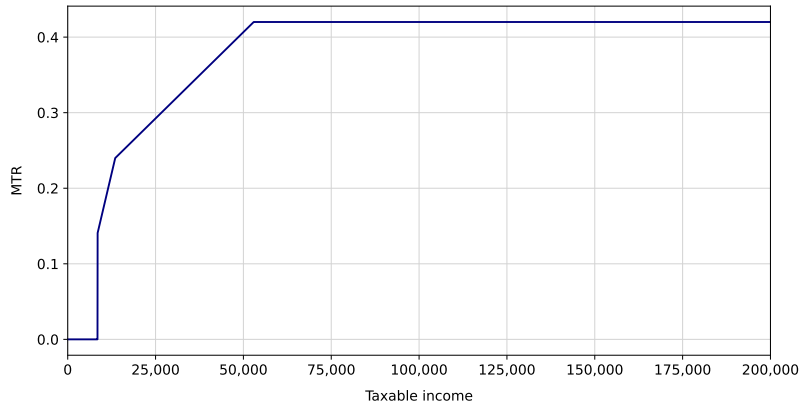
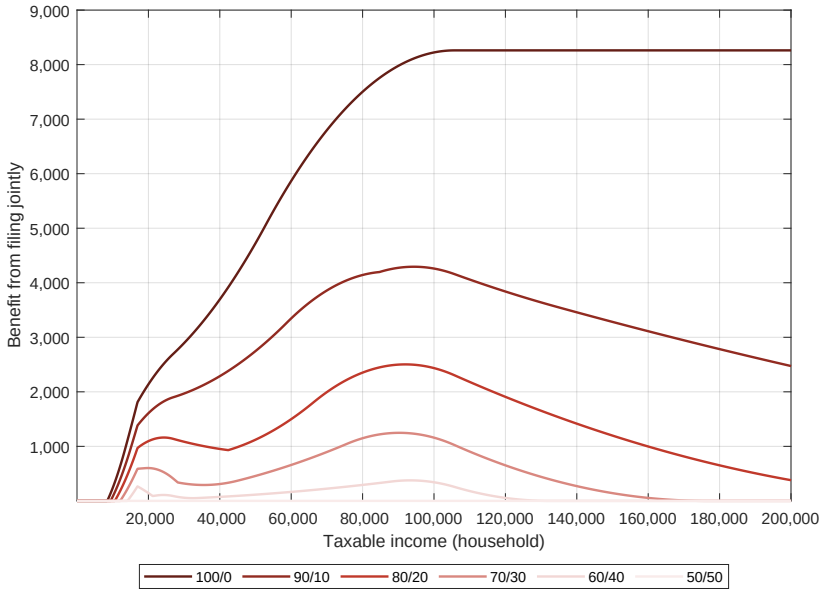


Figure A.7: Marginal Tax Rate

Notes: This figure plots the MTR over taxable income for individual taxation in 2015. Income is measured in Euro.

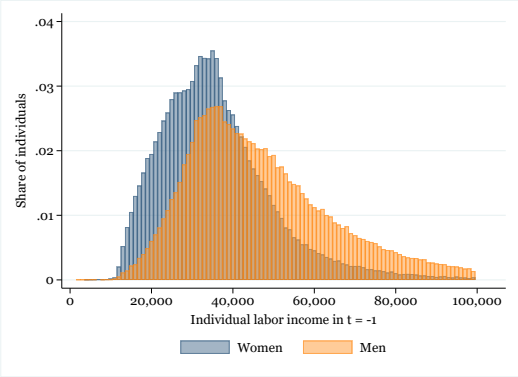
Figure A.8: Monetary Benefits from Joint Taxation



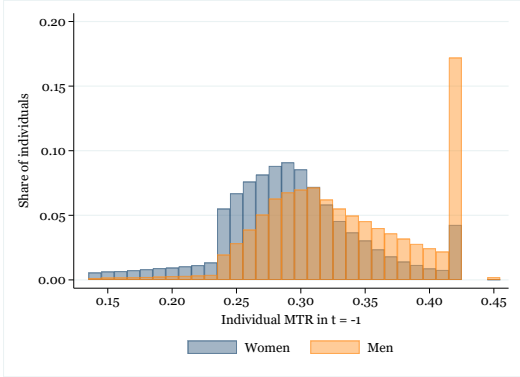
Notes: This figure shows the annual monetary benefit from joint taxation (as compared to individual taxation) for the tax year 2015 as a function of the total taxable household income for 5 different within-household income shares. 100/0 refers to the extreme case of a single-earner couple, with one spouse earning 100% of the joint income. 90/10, 80/20, 70/30, and 60/40 refer to cases where both spouses have positive, but different income, with the primary earner earning 90%, 80%, 70%, and 60%, respectively. 50/50 refers to the case of spouses earning exactly the same. In this case, joint taxation has zero monetary benefits, since both spouses face the same ATR when filing individually. All values are in Euro.

Figure A.9: Descriptive Statistics MTR Sample

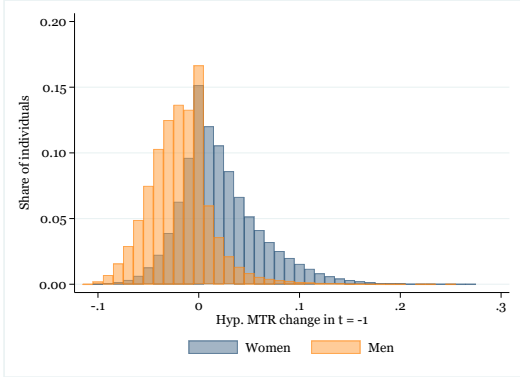
(a) Gross Labor Income ($y_{i,-1}$)



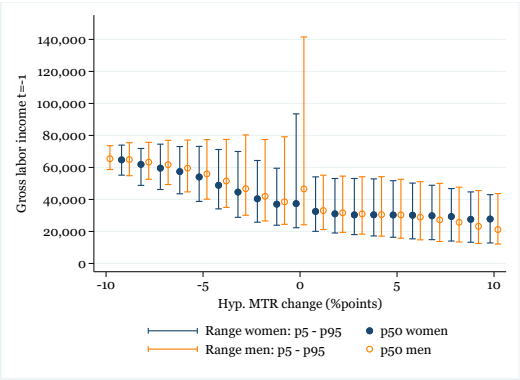
(b) $MTR_{i,-1}$



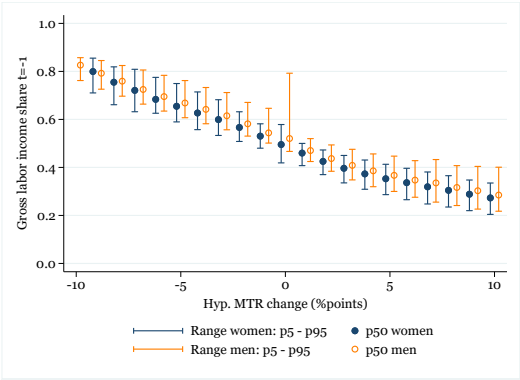
(c) $\Delta MTR_{i,-1}$



(d) Income by $\Delta MTR_{i,-1}$

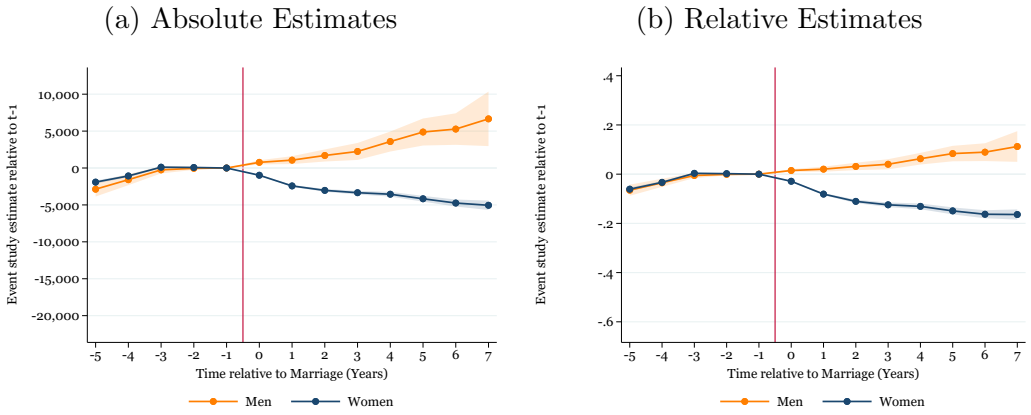


(e) Income Share by $\Delta MTR_{i,-1}$



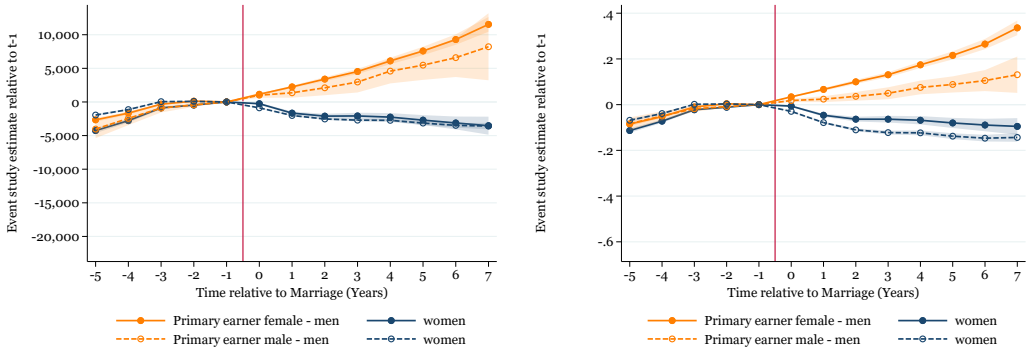
Notes: This figure plots descriptive statistics for the MTR sample that we use to estimate Equation 4 and 5. All panels show results for women and men for the pre-marriage year ($t = -1$). Income in 2015 Euros. Figures a, b, and c show histograms for the distribution of individual gross labor income ($y^g_{i,-1}$), the marginal tax rate ($MTR_{i,-1}$) and the absolute difference in marginal tax rate i would have faced if already filing jointly in $t = -1$ ($\Delta MTR_{i,-1} = MTR_{i,-1} - MTR^{hyp}_{i,-1}$), respectively. Panel d and e depict the variation in absolute income and within couple income share over Diff MTR. The vertical lines show the range for the medium 90% of the distribution (p5 to p95) and the dots show the median value (p50). Outliers in the p95 values at Diff MTR 0 because (very) high income taxpayers are less likely to face a change in MTR under joint taxation: they are in the top MTR brackets with a flat MTR.

Figure A.10: Marriage Earnings Gap (MTR Sample)



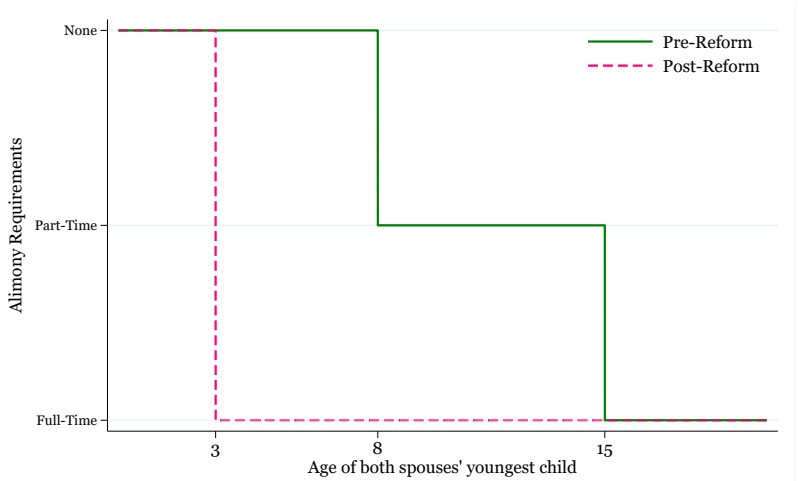
Notes: Annual data from TPP. This figure shows the event study estimates from estimating a modified version of Equation 2 for the MTR sample (this is the equivalent of Figure 2). Due to the TPP data structure, event time is measured in years and we include event time dummies for the first child only. Plots show point estimates and 95% confidence intervals. Panel a shows the absolute estimates $\hat{\beta}_t^g$ in 2015 Euros, directly retrieved from estimating Equation 2. Panel b shows relative earnings gap estimates for MG_t^g as defined in Equation 3.

Figure A.11: Marriage Earnings Gap Primary vs. Secondary Earners (MTR sample)



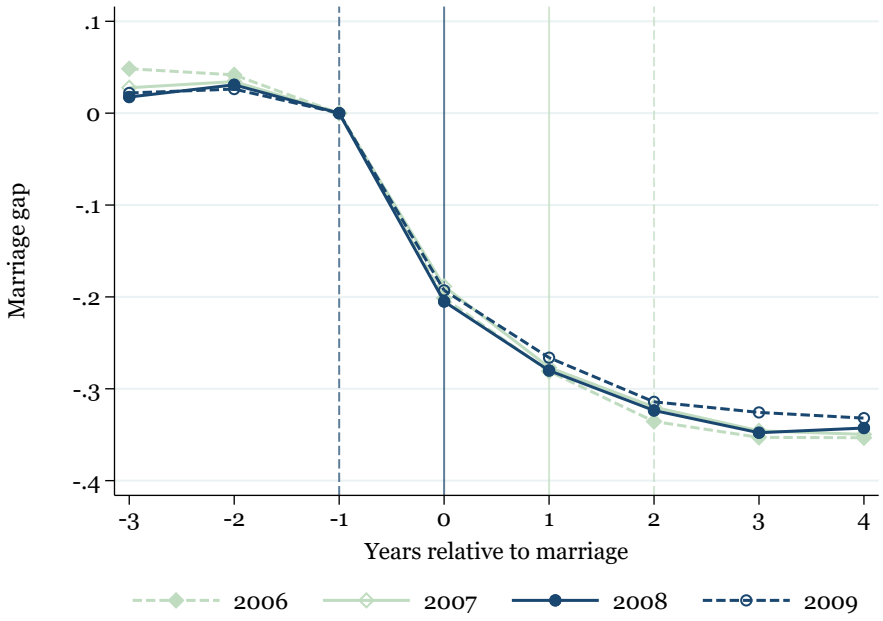
Notes: Annual data from TPP. This figure plots the marriage earnings gap estimates for two subsamples of the full sample from Figure A.10: couples with female primary earners (solid lines) and couples with male primary earners (dashed lines). Based on their income in $t = -1$, primary earners face a labor supply incentive from a lower MTR under joint taxation, while secondary earners face a labor supply disincentive from a higher MTR. We exclude couples where at least one spouse has no change in MTR. These are typically (very) high-income couples, where the MTR remains in the flat top-MTR brackets with joint taxation. Plots show point estimates and 95% confidence intervals. Panel a shows the absolute estimates $\hat{\beta}_t^g$ in 2015 Euros. Panel b shows relative earnings gap estimates for MG_t^g.

Figure A.12: Illustration 2008 Reform



Notes: Illustration of the alimony payments as a function of child age pre-reform (green line) and post-reform (pink dashed line). Pre-reform, divorced resident parents are fully entitled to alimony payments from their ex-spouse until the youngest child turns 8. For children 8 to 14, the parent is expected to work part-time to be eligible. After that, the parent is no longer eligible. Post reform, eligibility ends as soon as the youngest child turns 3, expecting the caregiving parent to work full-time from this point.

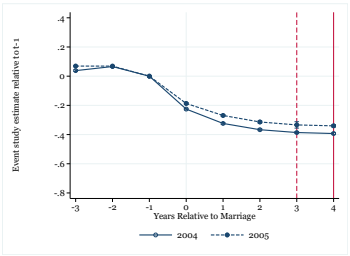
Figure A.13: The Marriage Earnings Gap Around the Reform



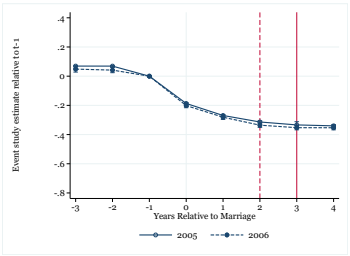
Notes: Annual data from TPP. This figure shows the event study estimates MG_t^w from estimating Equation 2 for women of different year-of-marriage cohorts. 2006 and 2007 are the last two cohorts that got married under the old regime, 2008 and 2009 are the first two cohorts that got married under the new regime. Vertical lines indicate the event time t at which a cohort is hit by the 2008 reform. For the 2008 cohort, this is $t = 0$, for the 2007 cohort, this is $t = 1$, and so on. Figure A.14 provides pairwise comparisons of the marriage earnings gap estimates for a larger set of year-of-marriage cohorts (2004 to 2011).

Figure A.14: The Marriage Earnings Gap Around the Reform – Many Years

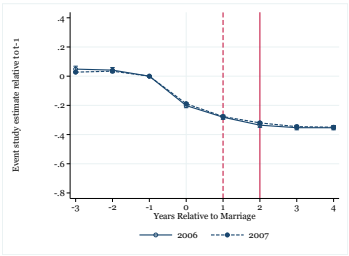
(a) 2004 vs 2005



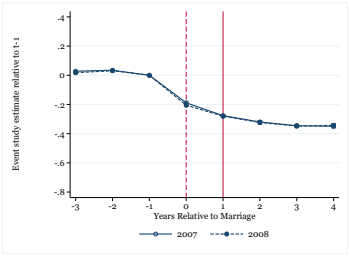
(b) 2005 vs 2006



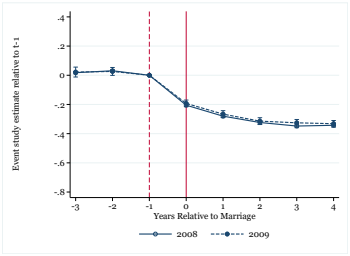
(c) 2006 vs 2007



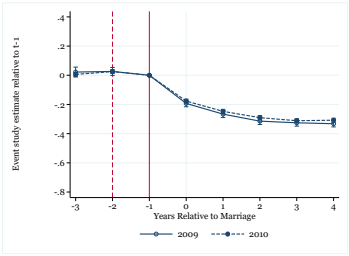
(d) 2007 vs 2008



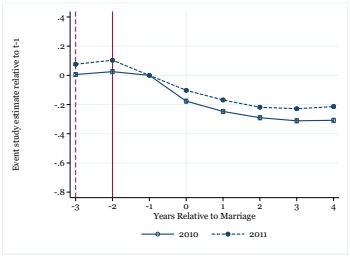
(e) 2008 vs 2009



(f) 2009 vs 2010



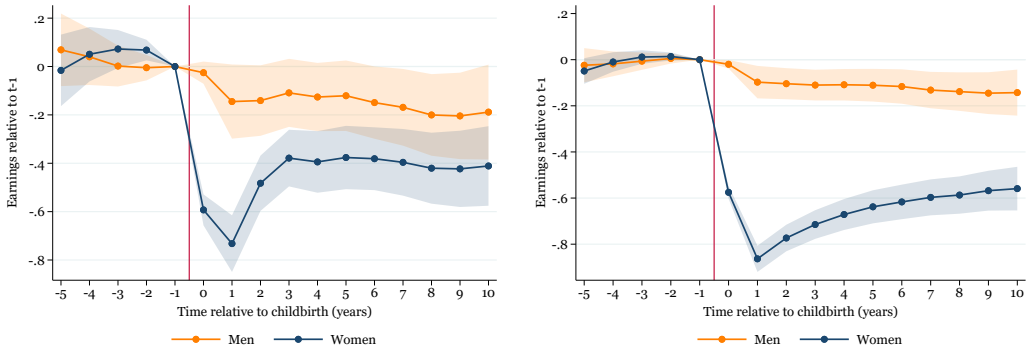
(g) 2010 vs 2011



Notes: Annual data from TPP. This figure shows the event study estimates from estimating Equation 2 for different year-of-marriage cohorts. Each panel compares the marriage earnings gap for two marriage cohorts that got married in subsequent years. Vertical lines indicate the event time t in the year of the reform (2008) for the respective year-of-marriage cohort.

Figure A.15: Child Penalties - East vs. West German Culture

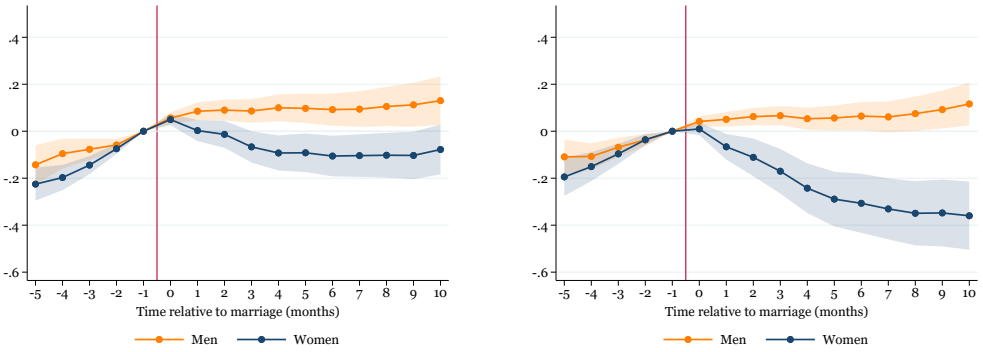
(a) East (b) West



Notes: Annual data from SOEPRV. This figure shows the event study estimates from estimating the child earnings gap (or child penalty) in relative earnings for individuals born before the German reunification (< 1989) with childbirth after reunification (> 1990). We split the sample into individuals born in the German Democratic Republic (East, panel a) and individuals born in the Federal Republic of Germany (West, panel b).

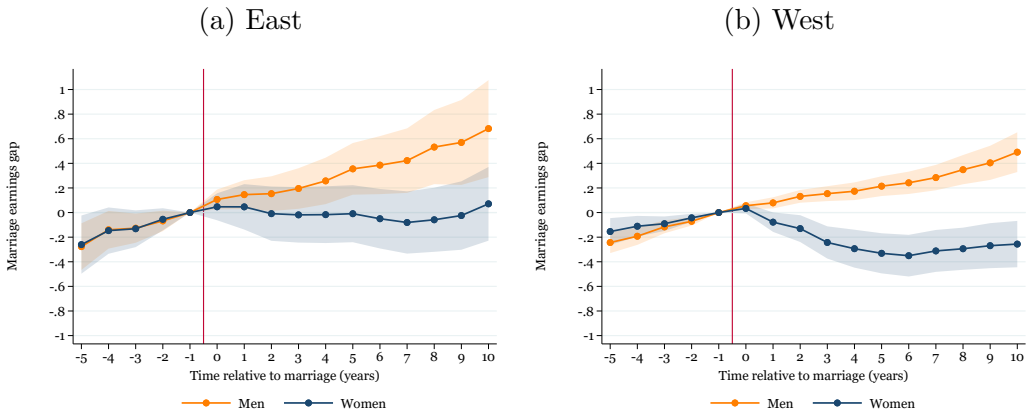
Figure A.16: The Marriage Earnings Gap – Gender Norms

(a) Progressive Norms (b) Traditional Norms



Notes: Annual data from SOEP–RV. This figure shows the event study estimates from estimating Equation 2 for individuals who stated their views on gender norms in the SOEP questionnaire. We split the sample into individuals with traditional (panel a) vs progressive (panel b) views on gender roles. See subsection A.4 for the underlying survey questions.

Figure A.17: The Marriage Earnings Gap – East vs West German Culture



Notes: Annual data from SOEP–RV. This figure shows the event study estimates from estimating Equation 8. The figure plots the marriage earnings gap estimates for individuals who were born before 1989 in East Germany (panel a) vs. in West Germany (panel b) and got married after 1990, in the reunified country. We define the marriage earnings gap as $(\hat{\beta}_j^g + \hat{\delta}_j^g)/E[\tilde{Y}_{it}^g|t]$ for the East, and as $\hat{\beta}_j^g/E[\tilde{Y}_{it}^g|t]$ for the West. Estimates account for the child penalty and allow child penalty estimates to differ between East an West.

Table A.1: 2SLS first stage

(a) Women

	(1) baseline	(2) income	(3) share	(4) both
t=0	0.70***	0.66***	0.35***	0.32***
	(0.00)	(0.00)	(0.00)	(0.00)
t=1	0.55***	0.50***	0.21***	0.19***
	(0.00)	(0.00)	(0.00)	(0.00)
t=2	0.51***	0.45***	0.17***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)
t=3	0.48***	0.42***	0.16***	0.14***
	(0.00)	(0.00)	(0.00)	(0.00)
t=4	0.47***	0.41***	0.15***	0.13***
	(0.00)	(0.00)	(0.00)	(0.00)
t=5	0.45***	0.40***	0.14***	0.12***
	(0.00)	(0.00)	(0.00)	(0.00)
t=6	0.45***	0.39***	0.13***	0.12***
	(0.00)	(0.00)	(0.00)	(0.00)
t=7	0.43***	0.37***	0.12***	0.11***
	(0.01)	(0.01)	(0.00)	(0.00)
Income (θ^w_{income})		✓		✓
Income share (ρ^w_{share})			✓	✓
Event time first child	✓	✓	✓	✓
Age (δ^w_{age}, ζ^m_{age})	✓	✓	✓	✓
Year (η_{cal})	✓	✓	✓	✓

(b) Men

	(1) baseline	(2) income	(3) share	(4) both
t=0	0.60***	0.51***	0.32***	0.30***
	(0.00)	(0.00)	(0.00)	(0.00)
t=1	0.42***	0.34***	0.19***	0.18***
	(0.00)	(0.00)	(0.00)	(0.00)
t=2	0.37***	0.30***	0.16***	0.15***
	(0.00)	(0.00)	(0.00)	(0.00)
t=3	0.35***	0.28***	0.14***	0.14***
	(0.00)	(0.00)	(0.00)	(0.00)
t=4	0.34***	0.28***	0.13***	0.13***
	(0.00)	(0.00)	(0.00)	(0.00)
t=5	0.32***	0.27***	0.13***	0.12***
	(0.00)	(0.00)	(0.00)	(0.00)
t=6	0.32***	0.26***	0.12***	0.12***
	(0.00)	(0.00)	(0.00)	(0.00)
t=7	0.31***	0.25***	0.11***	0.11***
	(0.01)	(0.00)	(0.00)	(0.00)
Income (θ^w_{income})		✓		✓
Income share (ρ^w_{share})			✓	✓
Event time first child	✓	✓	✓	✓
Age (δ^w_{age}, ζ^m_{age})	✓	✓	✓	✓
Year (η_{cal})	✓	✓	✓	✓

Notes: First stage results from estimating Equation 5 for women (panel a) and for men (panel b). Second stage results are reported in Table 4 and Table A.2, respectively. For details on the specification, see table notes of Table 4 and A.2.
Significance level: *** 0.01; ** 0.05; * 0.1. Robust standard errors are reported in parentheses.

Table A.2: Joint taxation: elasticity estimates for men (2SLS)

	(1) baseline	(2) income	(3) share	(4) both
t=0	0.04**	0.03	0.02	0.01
	(0.02)	(0.02)	(0.02)	(0.02)
t=1	-0.01	-0.07**	0.03	-0.01
	(0.02)	(0.02)	(0.03)	(0.03)
t=2	-0.07**	-0.16***	0.03	-0.03
	(0.02)	(0.03)	(0.04)	(0.03)
t=3	-0.16***	-0.26***	0.01	-0.06
	(0.02)	(0.03)	(0.04)	(0.04)
t=4	-0.17***	-0.23***	-0.03	-0.07
	(0.03)	(0.03)	(0.05)	(0.05)
t=5	-0.20***	-0.22***	-0.01	-0.03
	(0.04)	(0.04)	(0.06)	(0.06)
t=6	-0.20***	-0.22***	0.17	0.10
	(0.05)	(0.05)	(0.09)	(0.09)
t=7	-0.18*	-0.14	0.07	0.08
	(0.08)	(0.09)	(0.15)	(0.15)
Income (θ^w_{income})		✓		✓
Income share (ρ^w_{share})			✓	✓
Event time first child	✓	✓	✓	✓
Age (δ^w_{age}, ζ^m_{age})	✓	✓	✓	✓
Year (η_{cal})	✓	✓	✓	✓

Notes: Estimates for the elasticity of gross income with respect to the net-of-tax rate, from estimating Equation 4 for men. This is the same analysis as reported in Table 4 for women. For more details see table notes of Table 4.
Significance level: *** 0.01; ** 0.05; * 0.1. Robust standard errors are reported in parentheses.

Table A.3: Sample characteristics women: dropout vs. observed

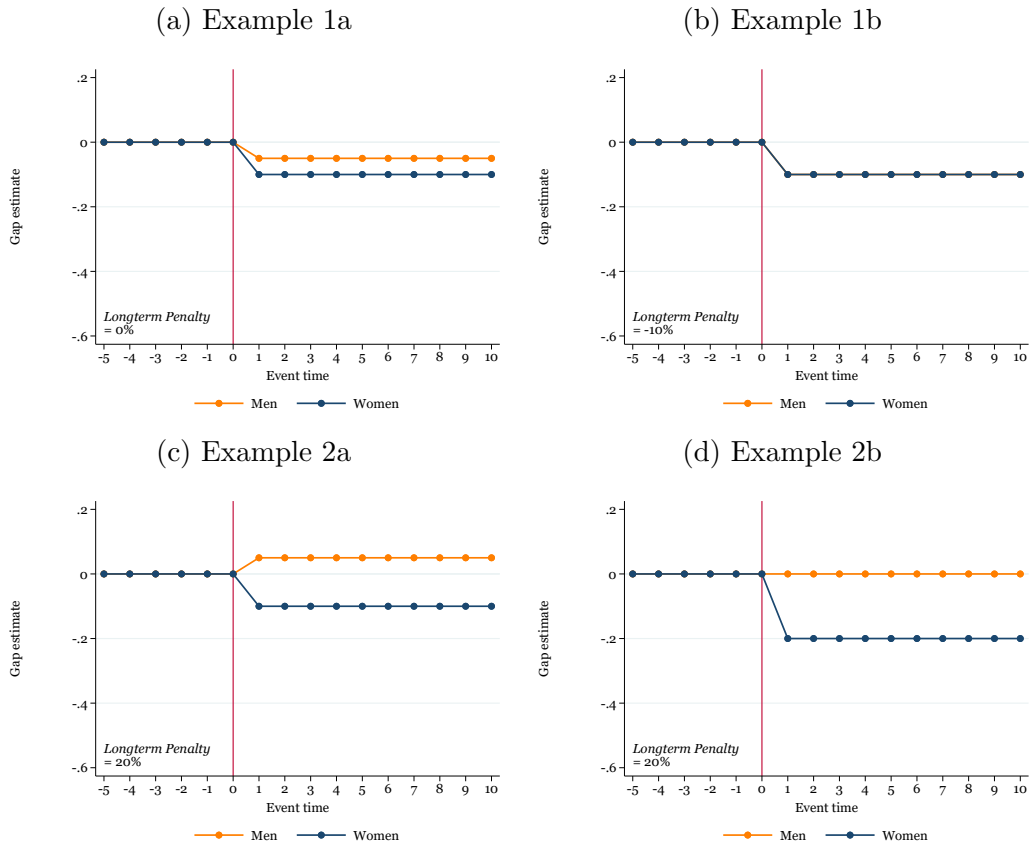
	$t = -1$ (drop out)	$t = 0$ (joint filing)
Share children (pre marriage)	0.22	0.34
Share East Germany	0.18	0.16
Share no religion	0.34	0.36
Share protestant	0.27	0.25
Share catholic	0.32	0.32
Share other	0.07	0.06
N	1,929,430	1,726,221

Notes: TPP data. The table compares characteristics of women that drop out of the tax data (what we define as $t = -1$) and those of women that show up on their husbands tax records in $t = 0$. All values are means. Samples are not identical as described in subsection 4.2.

A.2 Methodology Penalty Estimates

The child penalty literature typically plots the relative gap $G_t^g \equiv \hat{\beta}_t^g / E[\tilde{Y}_{it}^g|t]$ [24] but states a total penalty point estimate as the definite outcome. This total penalty is defined as $P_t \equiv (\hat{\beta}_t^m - \hat{\beta}_t^w)/E[\tilde{Y}_{it}^w|t]$, which is interpreted as the percentage by which women fall behind men after the event (see, e.g., Kleven, Landais, and Søgaard 2019). This measure compares the difference in the absolute income change of women and men in the numerator $(\hat{\beta}_t^m - \hat{\beta}_t^w)$ relative to women's counterfactual earnings in the denominator $(E[\tilde{Y}_{it}^w|t])$. P_t is thus unaffected by men's counterfactual earnings $E[\tilde{Y}_{it}^m|t]$ and does not account for differences in female and male counterfactual earnings.

Figure A.18: Illustration Gap Estimates



Notes: This figure illustrates the estimates for G_t^w (orange) and G_t^m (blue) for the 4 illustrative examples described in the text. For all panels, $E[\tilde{Y}_{it}^m|t] = 2{,}000$ Euro and $E[\tilde{Y}_{it}^w|t] = 1{,}000$ Euro for $t \leq 0$.

As a result, the interpretation of P_t is not straightforward, particularly if the

[24] We denote the gap by G_t^g here, while the child penalty literature typically denotes this by P_t^g (see e.g., Kleven, Landais, Posch, Steinhauer, and Zweimüller 2019). This is equivalent to our MG_t^g definition for marriage.

absolute female and male counterfactual earnings in t are different. We illustrate this with four simple examples, assuming constant, but different incomes for women and men that are only affected by t for $t > 0$, with counterfactual incomes $E[\tilde{Y}_{it}^{m}|t] = 2,000$ and $E[\tilde{Y}_{it}^{w}|t] = 1,000$ for all t.

Since P_t is a relative measure based on absolute differences in event-time estimates of women and men that does not consider differences in the counterfactual income, differences in G_t^m and G_t^w do not directly translate into P_t – and vice versa. This can be illustrated by considering two extreme cases. First, assume that women and men face the same absolute drop in earnings in $t = 1$, e.g., $\hat{\beta}_t^w = \hat{\beta}_t^m = -100$ for $t > 0$. In this case, women's earnings are 10 percent lower than their counterfactual earnings in $t > 0$, while the relative gap is 5 percent for men, as illustrated in Figure A.18a. This scenario results in a penalty of $P_t = (-100 - (-100))/1,000 = 0$. Put differently, the point estimate for the penalty here is zero, although, in relative terms, women reduce their earnings twice as much as men with $G_t^w = -.1$ and $G_t^m = -.05$ for all $t > 0$.

Now consider the other extreme case where women and men face the same relative drop in earnings in $t = 1$, e.g., $\hat{\beta}_t^w = -100$ and $\hat{\beta}_t^m = -200$ for $t > 0$, resulting in a 10% drop for both genders as shown in Figure A.18b. While the relative gap is the same, this scenario results in a penalty of $P_t = (-200 - (-100))/1,000 = -10\%$, indicating that female earnings increase relative to male earnings (or decrease less). These examples show that since male earnings enter the numerator but not the denominator, P_t is not necessarily informative of the relative gap $G_t^g = \hat{\beta}_t^g / E[\tilde{Y}_{it}^{g}|t]$.

In addition, the P_t estimate expresses changes in female earnings always relative to changes in male earnings. For given $E[\tilde{Y}_{it}^{w}|t]$ the P_t estimates are the same as long as the absolute difference $\hat{\beta}_t^m - \hat{\beta}_t^w$ is the same. As a result, P_t is not informative about which gender experiences a change in earnings and of what size. To illustrate this, assume first that women's absolute earnings decrease by 100 for $t > 0$, while men's earnings increase by 100 ($\hat{\beta}_t^w = -100$ & $\hat{\beta}_t^m = 100$). In this case, women's earnings are 10 percent lower than their counterfactual earnings, while men's earnings increase by 5 percent post event (Figure A.18c). This results in a penalty of $P_t = (100 - (-100))/1,000 = 0.2$. Now assume that men's earnings remain unchanged while women's earnings decrease by 200 ($\hat{\beta}_t^w = -200$ & $\hat{\beta}_t^m = 0$). This again results in a penalty of $P_t = (0 - (-200))/1,000 = 0.2$, while the relative gap estimates are very different as illustrated in Figure A.18d. Specifically, the relative gap for women is now twice as large, while men face no change in earnings.

Since this paper focuses on relative earnings changes post marriage, we decided to not use the P_t estimates.

A.3 International Comparison

In this section, we reproduce our results for 25 European countries plus Israel.

Data For this cross-country analysis, we use survey data from the SHARE Job Episodes Panel (SHARE JEP, Brugiavini, Orso, Genie, Naci, and Pasini 2022). The SHARE JEP is a retrospective panel, covering the entire biography of respondents from the Survey of Health, Ageing and Retirement in Europe (SHARE). SHARE JEP focuses on older individuals (respondents are aged 50+), but provides information on their labor market status, marriage status, and number of children for every year in a respondent's life. This allows us to estimate Equation 2 for the extensive margin. The survey is run in a total of 26 countries: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland.

Table A.4: Descriptive Statistics Germany SHARE

	All	Women	Men
Marriage Sample			
Age at marriage	25	23	26
Year of marriage	1971	1971	1972
Share divorced	0.17	0.17	0.17
Years married	14	13	15
Share children (ever)	0.91	0.92	0.91
N	4,498	2,417	2,081

Data: SHARE, raw data (unweighted), mean values. *Marriage Sample* Germany: Individuals that are married during the sample period, aged 18 to 45 at marriage and observable at minimum 4 years before and 8 years after marriage.

Marriage Gap We closely follow our main analysis and estimate Equation 1 separately for men and women. Our outcome Y_{it}^g is as a binary dummy of labor market participation that is 1 if respondent i worked in event year t and 0 else, i.e., the extensive margin (working yes/no). Event time is measured in years relative to marriage.

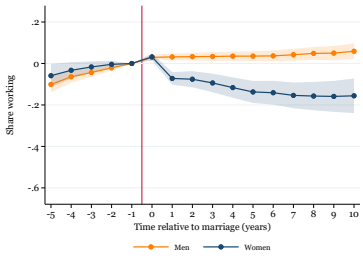
Table A.5: Sample Size SHARE

	All		Never		Before		During		After	
	w	m	w	m	w	m	w	m	w	m
AT	1,987	1,390	147	120	322	194	1,488	1,046	30	30
BG	1,104	790	86	68	26	19	989	699	3	4
BR	3,207	2,525	326	277	89	81	2,758	2,131	34	36
CH	1,479	1,203	186	160	63	52	1,218	969	12	22
CY	703	481	23	13	9	5	671	461	0	2
CZ	2,917	1,941	109	84	142	76	2,642	1,752	24	29
DE	2,488	2,129	195	213	214	151	2,036	1,712	43	53
DK	2,055	1,703	171	168	399	332	1,456	1,170	29	33
EE	2,836	1,753	143	97	292	149	2,362	1,483	39	24
ES	2,918	2,290	257	212	79	71	2,577	1,995	5	12
FI	964	813	51	57	64	67	836	682	13	7
FR	2,459	1,761	170	133	153	115	2,095	1,478	41	35
GR	2,280	1,694	257	188	30	26	1,992	1,473	1	7
HR	1,283	994	66	50	50	39	1,159	898	8	7
HU	903	571	51	32	18	14	828	520	6	5
IE	441	327	27	20	9	10	403	295	2	2
IL	1,098	819	155	131	42	31	893	650	8	7
IT	2,785	2,328	220	185	61	46	2,498	2,093	6	4
LT	1,230	695	72	45	48	19	1,094	621	16	10
LU	642	528	51	45	32	31	548	441	11	11
LV	1,011	569	54	30	69	34	877	490	11	15
MT	581	458	110	91	2	4	467	363	2	0
NL	1,124	937	135	101	25	17	956	807	8	12
PL	2,928	2,294	127	114	146	88	2,646	2,081	9	11
PO	643	512	46	37	25	19	570	455	2	1
RO	1,166	848	52	41	52	44	1,046	749	16	14
SE	1,972	1,666	118	137	350	308	1,469	1,194	35	27
SI	1,972	1,381	86	73	208	118	1,671	1,184	7	6
SK	1,067	877	97	105	41	23	928	746	1	3

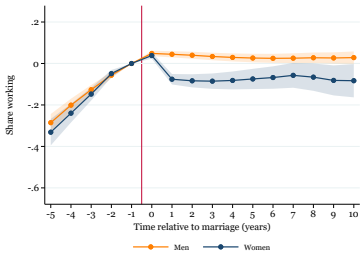
Notes: Number of married individuals (w = women, m = men) for each country in the SHARE data. *All*: all individuals who get married at some point in their life with $18 \leq$ age ≤ 45 at their first marriage. The other columns refer to subsamples based on whether and when the individual has children. We differentiate between individuals who never have children (*Never*), those who have their first child before their first marriage (*Before*), during that marriage (*During*), or after divorce (*After*).

Figure A.19: Extensive Margin Marriage Gap – International Comparison

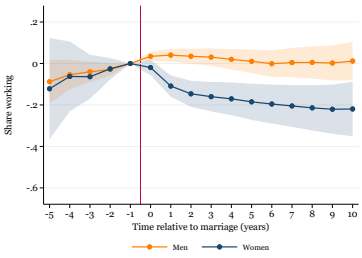
(a) Austria



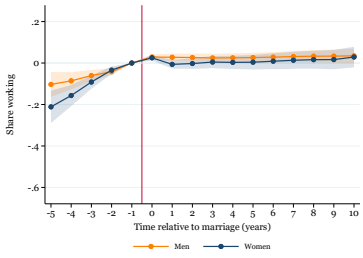
(b) Belgium



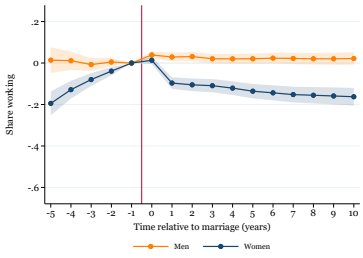
(c) Croatia



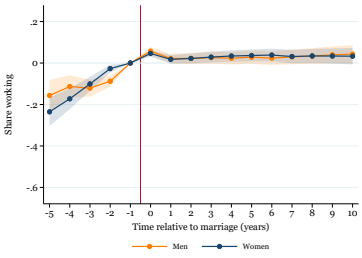
(d) Czech Republic



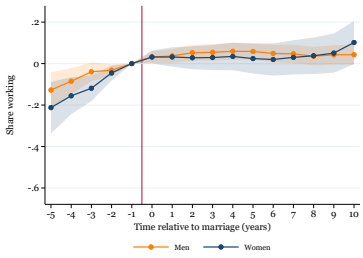
(e) Denmark



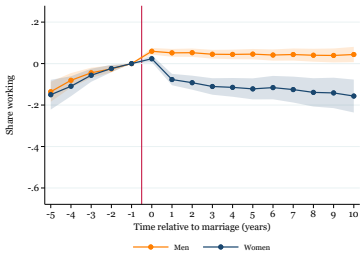
(f) Estonia



(g) Finland



(h) France



(i) Germany

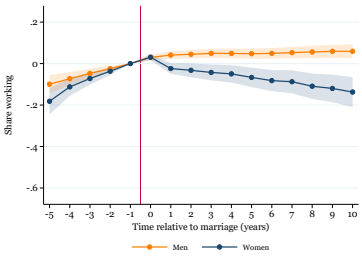
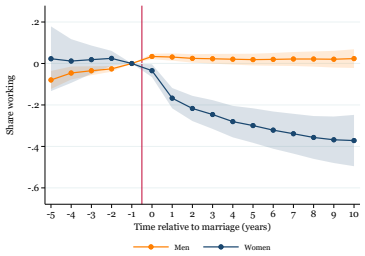
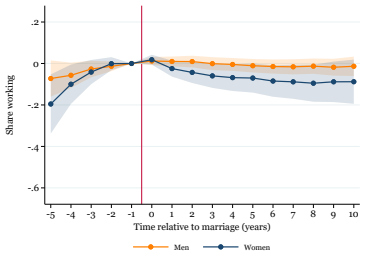


Figure A.19: Extensive margin marriage penalties – international comparison

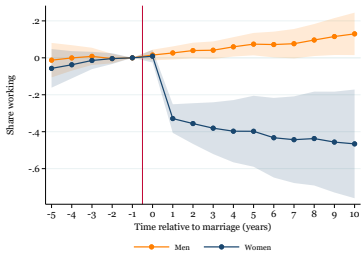
(j) Greece



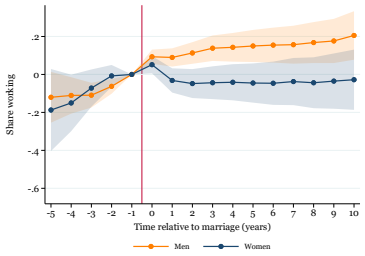
(k) Hungary



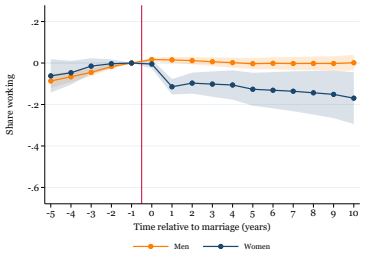
(l) Ireland



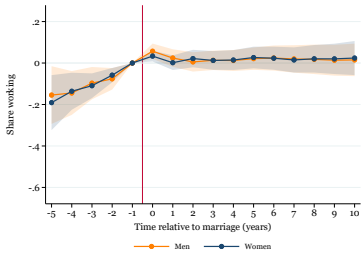
(m) Israel



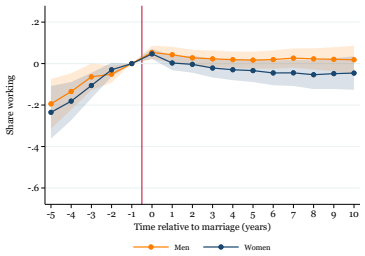
(n) Italy



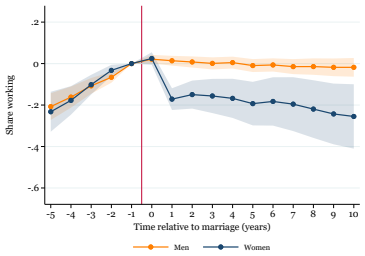
(o) Latvia



(p) Lithuania



(q) Luxembourg



(r) Netherlands

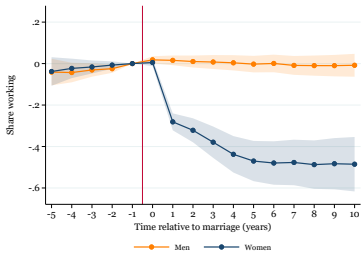
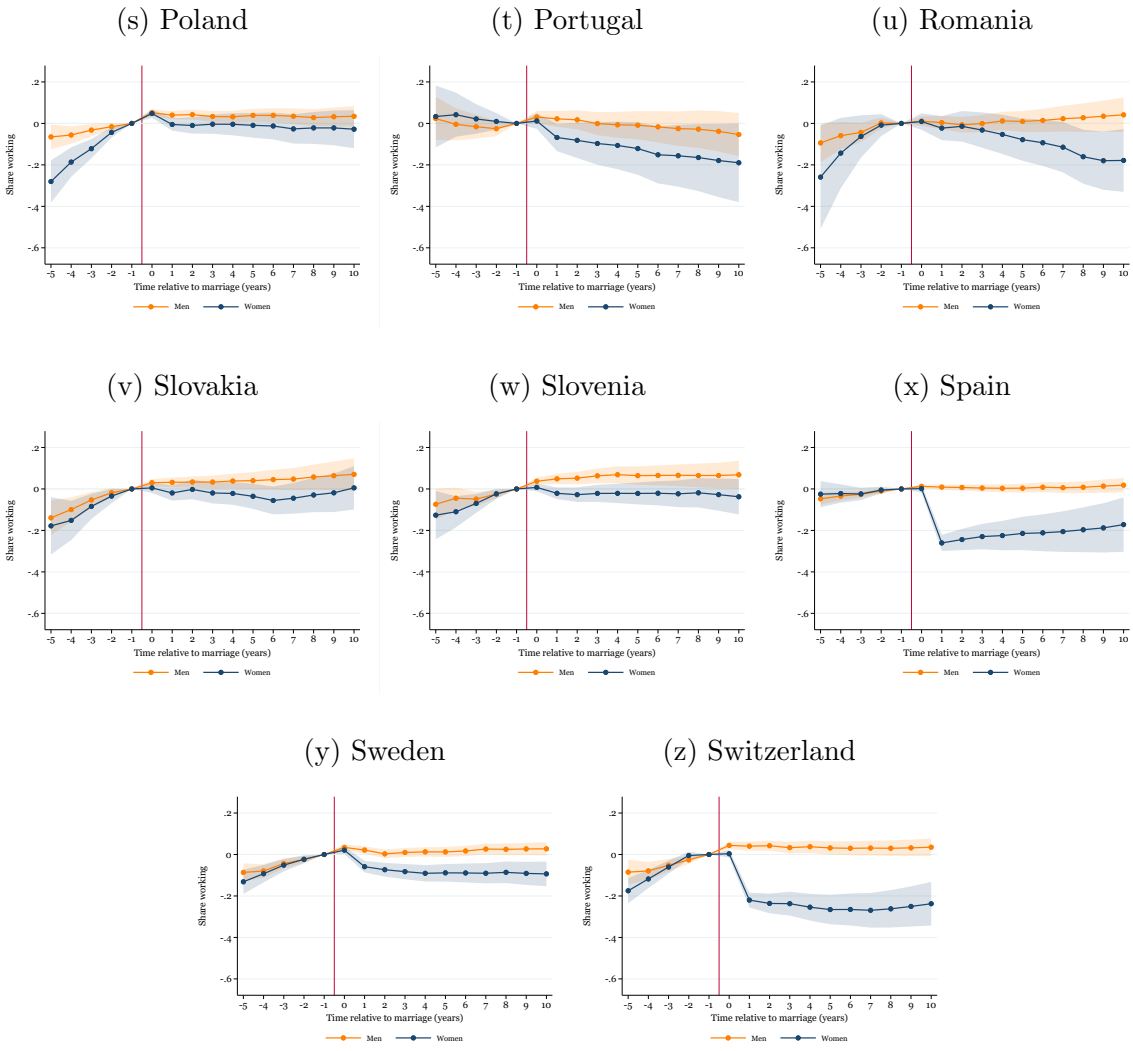


Figure A.19: Extensive margin marriage penalties – international comparison



Notes: Extensive margin marriage earnings gap (outcome = working yes/no). Results from estimating Equation 2 with the SHARE JEP data. All panels plot the marriage earnings gap estimates $\widehat{MG_t^g}$ for the extensive margin. *Time relative to marriage*: years before / after the first marriage.

A.4 Gender Norms Questions SOEP

Below are the SOEP questions used for our analysis on the gender norms of the survey participant in subsection 4.3:

1. **Gender Norm Question 1:** Child under 6 years suffers when mother works
2. **Gender Norm Question 2:** Child under 3 years suffers when mother works
3. **Gender Norm Question 3:** Women Should Rather Care About Family Than Career
4. **Gender Norm Question 4:** Best if man and woman work the same amount so they can share the responsibility