

A Service of

ZBU

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Giommoni, Tommaso; Rubolino, Enrico

Working Paper The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit

CESifo Working Paper, No. 11311

Provided in Cooperation with: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Giommoni, Tommaso; Rubolino, Enrico (2024) : The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit, CESifo Working Paper, No. 11311, CESifo GmbH, Munich

This Version is available at: https://hdl.handle.net/10419/305553

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit

Tommaso Giommoni, Enrico Rubolino



Impressum:

CESifo Working Papers ISSN 2364-1428 (electronic version) Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute Poschingerstr. 5, 81679 Munich, Germany Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de Editor: Clemens Fuest https://www.cesifo.org/en/wp An electronic version of the paper may be downloaded • from the SSRN website: www.SSRN.com

- from the RePEc website: <u>www.RePEc.org</u>
- from the CESifo website: <u>https://www.cesifo.org/en/wp</u>

The Cost of Gender Identity Norms: Evidence from a Spouse Tax Credit

Abstract

This paper studies the impact of tax incentives on economic behavior within the household. We focus on an Italian tax policy that grants a large tax credit to main earners if their spouses, designated as "dependent spouses" by the tax law, report income below a certain threshold. Combining a novel administrative dataset with a bunching approach, we find that second-earner women adjust their income to benefit from the tax credit, while second-earner men do not. Second-earner women holding more conservative gender norms are the ones who mostly reduce their income. This suggests that tax policies can exacerbate economic inequalities among families and depress female labor market outcomes when they interact with entrenched gender norms.

JEL-Codes: H240, H310, J160, J120.

Keywords: spouse tax credit, income taxation, gender norms, bunching, household behavior, female labor supply.

Tommaso Giommoni ETH Zurich, Department of Management, Technology, and Economics Leonhardstrasse 21 Switzerland – 8092 Zurich giommoni@ethz.ch Enrico Rubolino University of Lausanne, Faculty of Business and Economics Quartier de Chambronne, Internef 261 Switzerland – 1015 Lausanne enrico.rubolino@unil.ch

August 23, 2024

We would like to thank Sonia Bhalotra, Marius Brülhart, Maria De Paola, Francesco Fasani, Libertad González, Marko Köthenbürger, Rafael Lalive, Salvatore Lattanzio, Giorgia Menta, and Martin Fernandez Sanchez for useful discussions, and seminar participants at the University of Essex, LISER (University of Luxembourg), University of Milan, University of Naples (Federico II), Freie Universität Berlin, University of Zürich, and AIEL Conference for comments. Rubolino gratefully acknowledges financial support from the Swiss National Science Foundation (grant number 192546).

1 Introduction

In many countries, tax policies are designed to alleviate the financial strain on families by offering generous tax credits to primary earners, often when their spouses earn below a specified threshold.¹ Such policies raise critical questions about how tax incentives influence economic behavior within the household. Would secondary earners deliberately limit their income to secure this benefit for their partner? Would the gender of the secondary earner influence this decision?

Answering these questions would open a window into understanding how tax incentives shape household behavior. In standard economic models, it is usually assumed that individuals are rational utility maximizers, and individual behavior is characterized by finding conditions under which utility is maximized given available resources. When studying households, an additional layer of complexity arises: while resources may be shared, preferences within households may not be perfectly aligned (see Chiappori and Donni (2011); Chiappori and Mazzocco (2017); Almås et al. (2023) for surveys). The previous literature has emphasized how household decision-making involves not only economic factors, but also social and psychological dimensions (see, e.g., Akerlof and Kranton 2000, 2010, Guiso et al. 2006, Fernández 2007, Fernández and Fogli 2009, Bertrand et al. 2015, Field et al. 2021). However, the existing literature has largely overlooked how tax policy might interact with deep-rooted gender norms, potentially exacerbating economic inequalities between men and women (Alesina et al. 2011; Albanesi et al. 2023; Bursztyn et al. 2024). While previous studies have explored the effects of tax policy on the labor supply of (typically female) secondary earners (see, e.g., Kleven et al. 2009; Keane 2011; Bick and Fuchs-Schündeln 2017, 2018; Borella et al. 2023) and intra-couple time allocation (Ichino et al. 2024), empirical evidence specifically addressing the role of gender norms in responses to tax incentives remains sparse.

To break new ground on these questions, we focus on the Italian spouse tax credit. This policy entitles the primary earner in a couple to receive a significant tax credit if the second earner reports a gross annual income below 2,840.51 euros. The policy aims to provide substantial tax relief to low- and middle-income families. On average, the tax credit enables primary earners to reduce their tax burden by approximately 25 percent. This tax credit is salient and popular, with over a third of households receiving it in the past decade. Once eligible for the spouse tax credit, the tax law designates the secondary earner as the "dependent spouse" (coniuge a carico).

The spouse tax credit offers a propitious testing ground for evaluating how tax incentives shape household economic behavior. Since the Italian tax system is individually assessed, the policy creates a substantial "notch" in the budget constraint of families: a discontinu-

¹For example, in the UK, the marriage allowance permits the main earner in a couple to significantly reduce her or his tax burden if the secondary earner has an income below £12,570 (see Gov.uk/marriage-allowance for details). In Japan, a tax allowance of up to JPY 380,000 is given to a taxpayer who lives with a spouse whose income does not exceed JPY 480,000. Spanish households whose second earner has a net labor income below 8,000 euros may reduce their taxable income up to a maximum of 1,000 euros if the principal earner contributes to a pension fund for the spouse (see Taxing Wages (2024) - Tax and Gender through the Lens of the Second Earner for details and other examples).

ity in the choice set of second earner gross income versus family net income. Under the standard household model, the spouse tax credit notch should induce second earners, who would otherwise report more income, to "bunch" at the tax credit cutoff. However, in a society where traditional gender roles hold sway, the decision to bunch would be based on a cost-benefit analysis, where the benefit of having a larger family net income would be weighed against the cost of violating the adopted gender norm. In this context, a social expectation that positions men as the primary breadwinners and women as secondary earners would make men reluctant to be referred to as the "dependent spouse." This could lead men to refrain from reporting income below the tax credit cutoff. Conversely, women might not face the same level of societal pressure to avoid the label, as traditional norms might expect them to be more likely to be in a dependent role.

Using a novel administrative dataset, we study second-earners bunching responses to the spouse tax credit notch. Our results provide striking evidence of gender differences in responding to the spouse tax credit. We find sizable bunching at the tax credit cutoff from second-earner women, but no response from second-earner men.² Our bunching estimate suggests that the female marginal buncher reduces her income by about 261 euros to let her husband enjoy a tax credit of nearly 689 euros, thus increasing the family's net income by around 428 euros. The lack of any bunching response from second-earner hold around 428 euros less than comparable couples where the wife is the second earner.

Because women and men typically sort into different occupations, one possible explanation for our finding is structural differences by gender in the scope for bunching. For instance, women can sort into occupations that allow them more flexibility in adjusting their labor supply (see, e.g., Goldin and Katz 2011; Goldin 2014; Olivetti and Petrongolo 2016). To rule out this hypothesis, we follow several steps. First, we focus on self-employees, where the scope for responding to the policy (either through labor supply or tax evasion) is fairly similar across male and female jobs.³ We also show that other possible factors that could explain why women are more responsive at the spouse tax credit notch, such as a larger economic incentive or a weaker labor market attachment, do not vary systematically by gender at the tax notch. If any, second-earner men should have a larger economic incentive to respond compared to second-earner women, since the spouse tax credit accounts, on average, for a larger portion of their family income. To assess an alternative interpretation that women are simply more likely to respond to tax incentives than men, we test for gender differences in other tax kinks or notches that create a discontinuity in an *individual* (rather than family) budget constraint. Although factors that could explain a gender difference in the scope for bunching are still present, gender norms would be less relevant in this case. We find no evidence of gender differences, suggesting that gender norms can be a suitable explanation for our finding.

²We flexibly control for the distribution of men's income and the distribution of women's income. Hence, we are not simply picking up the fact that women are more likely to report income close to the tax credit cutoff.

³The focus on self-employees does not represent a significant threat to the generalizability of our estimates. Self-employment has traditionally been an important element of the Italian economy, accounting for 22 percent of employment in 2021 (OECD (2023), Self-employment rate indicator).

Our result seems to reflect the influence of gender norms even when looking within genders. Second-earner women who hold conservative gender views are more likely to bunch at the spouse tax credit notch. We use several proxies to gauge gender norms. Following Fernández (2007), we conduct an epidemiological study on foreign-born immigrants. Adjusting for selection in their current municipality of residence and controlling for other potential (non-gender norms-related) determinants of bunching, we show that responses are more prominent among immigrant women coming from countries with more traditional gender norms. Extending to native women, we find larger responses in municipalities where gender norms are more traditional and among older women, who tend to report more conservative views on questions related to gender norms in survey data.

In the final part of our analysis, as a complimentary effect, we provide evidence that the policy has tangible costs for women: they *persistently* report income below the tax notch and miss job opportunities. However, we also show that couples where the wife is a "dependent spouse" are more likely to remain together. These results suggest that the policy creates a trade-off in women's choice: prioritizing family income and marital stability comes at the expense of their career advancement.

This paper contributes to several strands of the existing literature on gender inequality, gender norms, and tax policy. We show that gender norms significantly shape individual behavior. This finding is consistent with a growing literature showing that culture affects economic outcomes (see, e.g., Fernández et al. 2004; Guiso et al. 2006; Fernández 2007; Fernández and Fogli 2009; Bursztyn et al. 2017; Bursztyn et al. 2020; Giuliano 2020; Bursztyn et al. 2024). Previous works have focused on the non-monetary costs associated with violating gender identity norms and how norms affect within-household time allocation. Bertrand et al. (2015) show that couples where the wife outearns the husband are more likely to divorce and report lower marriage satisfaction. Ichino et al. (2024) focus on the impact of gender norms on household's time allocation decisions, showing how norms can lead to misallocation of women's time. Using a novel dataset and setting, our analysis focuses on an untested, perhaps more compelling, outcome: adopting behaviors that maximize family income. To our knowledge, we offer the first evidence of monetary costs that couples face when men are secondary earners.⁴

Our results emphasize how policy-making can backfire in the presence of traditional gender norms. We show that the spouse tax credit limits women's income by incentivizing second-earner women to report income below a small threshold. Our results relate to studies showing how public policy can hold back female employment, generating efficiency losses. In most countries, taxes and benefits depend on one's marital status and tend to reduce the labor supply of the secondary earner (Taxing Wages (2024) - Tax and Gender through the Lens of the Second Earner). For instance, the work of Bick and Fuchs-Schündeln (2017, 2018) has pointed to the relevance of tax policy for understanding cross-country variation in

⁴Our results challenge the predictions of the standard unitary labour supply model. In particular, we reject the assumption of income pooling, where the source of income does not matter for household behavior (Chiappori 1988; Chiappori 1992; Apps and Rees 1996; Blundell et al. 2007; Donni 2008). Our result is consistent with Bobonis (2009), which shows female partners having greater sensitivity to own income changes, consistent with norms that oblige women to devote their earnings to meet collective consumption needs.

married women's labor supply. Borella et al. (2023) show that eliminating marriage-related provisions in the US would significantly increase married women's labor market participation. Our contribution to this literature is to show how such tax policies can exacerbate inequalities among families and between genders in the presence of deep-seated gender norms.

We also contribute to a growing literature documenting that optimization frictions dampen responses to tax policy (Chetty et al. 2011; Chetty 2012; Kleven and Waseem 2013; Gelber et al. 2020). The existing literature has presented several sources of frictions, including imperfect knowledge (Chetty et al. 2013a), search costs and hours constraints (Chetty et al. 2011), complexity (Bhargava and Manoli 2015), and salience (Chetty et al. 2009), among the others. We propose a new source of friction: gender identity norms. Incorporating these sources of optimization frictions into traditional labor supply models can help to reconcile some long-standing empirical puzzles in public and labor economics, such as the divergence between micro and macro elasticity estimates (Chetty et al. 2013b). Our results should be taken into account in the design of a more efficient and equitable tax system.⁵

The rest of the paper proceeds as follows. In section 2, we describe the background and data. Section 3 illustrates a conceptual framework that helps to contextualize our findings and presents our empirical approach. Section 4 shows our main result and studies the role of gender norms and other alternative explanations. Section 5 presents the impact of the policy on women's labor market choices and marital stability. Finally, section 6 concludes.

2 Institutional Background and Data

2.1 Income Taxation and the Spouse Tax Credit

All Italian residents are subject to personal income taxation (IRPEF, *Imposta sul Reddito delle PErsone Fisiche*). The tax base depends on individual income, computed by subtracting deductions from gross income. All sources of income, such as labor (including self-employed work), business and capital income enter the tax base. The tax schedule is progressive: it is composed of five income brackets with tax rates ranging from 23 to 43 percent (see Table A1). Tax rates and income bracket cutoffs have not been modified during the period we study.

The final tax burden is calculated net of tax credits. A spouse tax credit was introduced by law 917/1986 (see *Decreto del Presidente Della Repubblica*, 22 December 1986, n. 917, article 12). This policy grants the main earner in a couple to receive a tax credit if the second earner reports gross annual income below 2,840.51 euros (including if she or he is out of

⁵Differentiating income tax rates by gender could be a solution to internalize the costs associated with gender identity norms. The implications of gender-based taxation have been studied by Alesina et al. (2011). Using a collective household model in which labor supply elasticities arise endogenously, they find that the optimal tax scheme would present higher marginal tax rates on men when gender-specific lump sump transfers are available. Empirical evidence on the labor market impacts from gender-based taxes has been recently studied by Rubolino (2022). Exploiting a recent policy change in Italy, he shows that lower payroll taxes on female hires stimulated female employment.

the labor force).⁶ Eligibility is granted to married couples and can be claimed even after a legal separation if the main earner pays alimony.⁷ Once eligible for the spouse tax credit, the tax law titles the second earner as the "dependent spouse" (*coniuge a carico*; see Appendix Figure Figure A1 for a sample of the tax return).

Appendix Table A2 illustrates the main features of the spouse tax credit. The table shows that the size of the tax credit is a negative function of the main earner's gross income. For instance, a main earner with an income lower than 10,000 euros would get a tax credit of 726.7 euros, which would almost halve her tax burden. On average, main earners with less than 15,000 euros get a tax credit that accounts for 43.2 percent of their tax burden, corresponding to around 10 percent of their gross income. The relevance of the tax credit then monotonically decreases over the main earner's income distribution, accounting for less than one-tenth of the tax burden for incomes above 29,000 euros. Consistent with the main goal of compensating disadvantaged families, the tax credit is eventually phased out for main earners reporting more than 80,000 euros.

2.2 Gender Norms in Italy

Similarly to other Southern European countries, Italy is a conservative society with traditional gender norms. Most families embrace the male breadwinner model, where wives are mostly out of the labor force or enter the labor market as second earners. According to the OECD Family Database, Italy ranks in the lowest position regarding female labor market outcomes: in 2018, the full-time equivalent employment share of women was 40.3 percent, and the gender employment gap was 26.5 percentage points. Casarico and Lattanzio (2023) show that the gender pay gap has declined steadily over the last two decades.

Italian society is also characterized by strong imbalances in family chores allocation. According to a nationwide survey conducted by the Italian Institute of Statistics (ISTAT) (see *Indagine sulle discriminazioni in base al genere, all'orientamento sessuale, all'appartenenza etnica*), more than one-fifth of married working women report "to feel overwhelmed by family chores."⁸ Moreover, around half of married women agree that "husbands are the main responsible for the provision of family needs". These figures portray Italy as a gender-conservative environment. It thus provides a useful setting to study whether pervasive gender norms can shape household behaviors.

2.3 Data and Descriptive Evidence

We use administrative data provided by the Veneto tax administration, based on the universe of personal income tax returns for residents in the Veneto region. Veneto is an important and large Italian region: in 2020 it was the third richest region in Italy (Italian Institute

⁶During the period Italy had its own currency, the cutoff was 3 million lire. It was then converted to 2,840.51 euros. A similar policy was also in place during the early postwar period, but based on different criteria (see *Decreto del Presidente Della Repubblica*, 29 January 1958, n. 645; *Supplemento Ordinario alla Gazzetta Ufficiale della Repubblica Italiana*, 7 July 1958, n. 162).

⁷In our dataset, we find that 95 percent of spouse tax credit recipients are married individuals.

⁸Alesina and Ichino (2009) discuss the implications of unpaid family work for labor supply decision of Italian women.

	Average value (1)	Standard deviation (2)	Count (3)		
	A. Men				
Married	0.635	0.481	1,892,849		
Age	45.73	10.77	1,892,849		
Foreign	0.068	0.253	1,892,849		
Gross income	34,695.34	54,866.94	1,892,849		
Taxable income	29,045.2	52,209.76	1,892,849		
Income tax	7,682.23	21,475.43	1,892,849		
		B. Women			
Married	0.648	0.477	822,493		
Age	45.87	11.16	822,493		
Foreign	0.071	0.257	822,493		
Gross income	25,049	34,918.3	822,493		
Taxable income	21,185.89	33,079.84	822,493		
Income tax	5,184.49	13,022.9	822,493		

Table 1: Descriptive Statistics of Self-Employed Workers

Note: This table displays the descriptive statistics. The variables *Gross income*, *Taxable income*, and *Income tax* are expressed in Euros. Data from tax returns on the universe of self-employed taxpayers from Veneto.

of Statistics, ISTAT).⁹ As income taxes in Italy are filled individually, the unit of observation is the individual. The dataset contains taxable income data (divided by income sources) and basic socio-demographic characteristics, such as gender, marital status, date of birth, municipality of residence, and nationality. We also observe all sources of tax deductions and credits. Data are available for the 2007-2014 period.

When studying second-earner responses, we focus on self-employees for two main reasons. First, our focus on self-employees implies that we study gender differences in bunching responses among individuals working in similar jobs, where the scope for responding to the spouse tax credit (either through labor supply or evasion) is similar. Second, since income is self-reported, self-employees can more easily adjust their taxable income. By contrast, since employees' wages are third-party reported, employees have limited room to respond to the spouse tax credit by bunching at the cutoff.¹⁰ We will confirm this stylized fact below, where we show that the responses from wage earners are substantially attenuated.

Table 1 reports the summary statistics on our sample of self-employees, which includes

⁹In Appendix Figure A2, we show that Veneto scores slightly below the Italian average on the traditional gender norms index, while Veneto's female employment is relatively higher than the Italian average. Therefore, our estimates should provide a lower bound effect on the economic impacts of gender identity norms for Italian society.

¹⁰For instance, Saez (2010) finds that bunching responses on kink points of the U.S. Earnings Income Tax Credit (EITC) are exclusively concentrated among self-employees. EITC recipients with only wage earnings show no evidence of bunching. See Di Marzio et al. (2023) for evidence of bunching responses of Italian self-employees to tax notches.

around 2.7 million taxpayers. Panel A focuses on male taxpayers, who represent 69.7 percent of the sample; panel B on female taxpayers. Men's average gross income is 34,695 euros, while women's average gross income is 25,049 euros. This gender income gap is not surprising, given the existence of gender income inequality in Italy. The socio-demographic characteristics of the taxpayers in our sample present a fairly homogeneous distribution by gender. The average age is 45 for both men and women, and around 7 percent of the sample is composed of immigrants.¹¹

In Figure A4, we depict the spouse tax credit's take-up (as a share of total married taxpayers). We distinguish by gender and we plot the evolution in the take-up rate by bins of the value of the tax credit (as a share of the total tax burden). The figure shows a gender gap in the spouse tax credit's take-up rate. This is not surprising: husbands are more likely to earn more than their wives. Therefore, it is more likely, ceteris paribus, that wives' income is reported below the cutoff, making husbands eligible for the tax credit.

3 Conceptual Framework and Empirical Strategy

3.1 Conceptual Framework

This section sets out a simple conceptual framework that incorporates gender identity norms in the standard model of household behavior. We study behavioral responses to the spouse tax credit notch building from the seminal contributions of Saez (2010), Chetty et al. (2011), and Kleven and Waseem (2013).

Benchmark model. The spouse tax credit creates a "notch" in the budget constraint of families, that is a discontinuity in the choice set of the second-earner gross income versus family net income. Under the conventional (unitary) model of household behavior, this notch should induce second earners, who would otherwise report more income, to instead bunch right at the tax notch.¹²

Figure 1 offers a simple illustration of how second earners would respond to the spouse tax credit notch. Panel A presents a budget set diagram; panel B the density distributions. Before couple formation, individuals report gross income, y, that maximizes their utility subject to a budget constraint. Income is distributed according to a smooth density distribution h(y), and any heterogeneity is due to preferences or idiosyncratic shocks. When individuals marry, and a household is created, the second earner in the couple will face a tax notch at income level y^* . The notch generates a region of strictly dominated choice in the second earner's income interval $(y^*, y^* + \Delta y^D]$, where she or he can increase both leisure and consumption (family net income) by moving to the notch point y^* . At this income level,

¹¹Appendix Figure A3 presents summary statistics of potential spouse tax credit "bunchers", defined as those reporting income between 2,840.51 and 4,840.51 euros before marriage. The values of several socio-economic and demographic characteristics present a fairly homogeneous distribution by gender.

¹²As the spouse tax credit accrues to the main earner in the couple, second earners will face a notch in the family budget constraint as long as family income, or a portion of it, is pooled together. The standard (unitary) labor supply model, where the source of income does not matter for household behavior (symmetry of the Slutsky matrix), has been challenged in several contexts (see Chiappori and Donni (2011); Chiappori and Mazzocco (2017); Almås et al. (2023) for surveys).

second earners maximize their family net income by letting their spouse claim a spouse tax credit *C*. All second earners located in the income interval $(y^*, y^* + \Delta y^*]$, where the bunching region is larger than the area of strictly dominated choice, $\Delta y^* > \Delta y^D$, will respond to the spouse tax credit tax notch by bunching.



A. Budget Sets

Figure 1: Behavioral Responses to a Spouse Tax Credit Notch

The figure offers an example of behavioral responses from two types of "bunchers". We

C. Optimization Frictions



define the second earner L as the one with the lowest income, y^* ; the second earner H as the one with the highest income, $y^* + \Delta y^*$. When a household is created, the second earner L will continue to choose income y^* , making his or her spouse eligible for the tax credit *C*. The second earner H will also bunch at the tax notch because is exactly indifferent between

the notch point y^* and the interior point y^I . Second earners L and H represent the two extreme cases: each second earner between L and H will bunch at the spouse tax credit notch after couple formation. Therefore, because no one is willing to locate between the spouse tax credit notch y^* and the interior point y^I , this model would predict a *density hole* in the segment (y^*, y^I) and *excess bunching* at the spouse tax credit notch y^* .

Assuming that the counterfactual density $h_0(y)$ is roughly constant on the bunching segment $(y^* + \Delta y^*)$, we can denote excess bunching at the spouse tax credit notch as:

$$B = \int_{y^*}^{y + \Delta y^*} h_0(y) dy \approx h_0(y^*) \Delta y^*.$$
 (1)

Gender identity norms and other frictions. The predictions of the benchmark model can be questioned by optimization frictions, such as adjustment costs or inattention. Panel C of Figure 1 incorporates optimization frictions (depicted by the gray shaded area) into the model. The key implication is that frictions prevent individuals from bunching, generating a significant density mass in the (otherwise empty) strictly dominated region.

In addition to these (standard) optimization frictions presented above, we introduce a new source of frictions: gender identity norms. Why and how gender identity norms would affect behavioral responses to the spouse tax credit? Bringing insights from social psychology into economics, Akerlof and Kranton (2000) propose a model where one's identity directly enters the utility function. Identity norms can influence economic outcomes because deviating from the behavior that is expected for one's social category is assumed to decrease utility. Relating the identity model to the concept of gender identity, the two relevant social categories are "man" and "woman", each associated with specific behavioral prescriptions which, if violated, will decrease utility. For instance, norms prescribing that "the husband is the main responsible for the provision of family needs" would be at odds with the fact that the husband is the second earner in the couple. Being titled as the "dependent spouse" would thus negatively enter in the men's utility function, preventing them from bunching. By contrast, norms prescribing women that "being a housewife is just as fulfilling as working" would not prevent women from bunching.¹³

To formalize these predictions, we can incorporate gender identity norms in the model, which can differently influence the income reporting behavior of husbands and wives who are secondary earners. They will suffer a utility loss λ - the disutility associated with being denoted as the "dependent spouse" - if he or she keeps income below the notch. The utility function for the household is therefore:

$$U(y_{main} + y_{secondary} + C) - \lambda, \tag{2}$$

where $U(y_{main} + y_{secondary} + C)$ is the utility from total family income (including the spouse tax credit). A secondary earner will adjust her or his income and report income $y^* < y$ just below the notch if

¹³The previous literature has emphasized that traditional gender roles and women's more prominent role in non-market work may negatively affect women's labor market outcomes (see, e.g., reviews in Bertrand 2011, Goldin 2014, Blau and Kahn 2017, and Bertrand 2020).

$$U(y_{main} + y *_{secondary} + C) - \lambda > U(y_{main} + y_{secondary}).$$
(3)

In a society with traditional gender norms (i.e., λ is high for male secondary earners and low for female secondary earners), male secondary earners will prefer to keep income above the threshold, while the wife might choose to reduce her income to qualify for the tax credit, assuming that the tax credit is large enough to offset the income reduction and frictions are small. By contrast, in a society with weaker traditional norms, secondary earners are willing to reduce their income to maximize the household's utility, as long as the tax credit more than compensates for the income drop and friction costs.

The implications of gender identity norms in a traditional society are presented in panel D of Figure 1, where frictions due to gender identity norms are represented by the green triangular area. The presence of gender identity norms implies that second-earner men may be unresponsive and stay above the notch, while women would continue to bunch at the tax notch. This result suggests that the density distribution of second earner *men* would be smooth around the spouse tax credit notch (as depicted by the red solid line in the graph), while the density distribution of second earner *women* would still present very sharp bunching. A key assumption is that optimization frictions (as well as other factors that could affect the scope for bunching) do not systematically differ by gender. We will inspect (and validate) this assumption below.

Dynamics and career concerns. As a benchmark, we have considered a static model. However, if career concerns are important, the spouse tax credit affects not only current income, but also income reported in the future. Bunching responses in a multi-period decision context would include intertemporal substitution effects and might dampen the static (annual) bunching response. To evaluate this implication, consider a dynamic framework where current income affects future income (due, e.g., to savings or through effects on earnings from learning by doing, job promotions, etc.). If there is a positive and continuous correlation between current and future income, then the dominated range would be smaller and the "optimal" bunching response lower. At the spouse tax credit notch, *current* net family income is discretely higher than at the point corresponding to the end of the strictly dominated choice region, but *future* net family income is only infinitesimally larger.

These arguments suggest that dynamic behavioral responses to the spouse tax credit notch can have important implications. Under a "static" perspective, bunching responses can be systematically smaller or even absent among second-earner women whose utility returns from complying with gender identity norms do not offset the costs of career deterioration and future income loss. Therefore, women's bunching response can be substantially dampened if they care about their future income prospects and they embrace a less traditional gender identity model.

3.2 Identification Strategy

We study whether the second earner's gender affects family income maximization choices by examining bunching responses at the spouse tax credit notch. Following previous studies (Saez 2010; Chetty et al. 2011; Kleven and Waseem 2013), we start by grouping taxpayers in *j* bins of gross income and calculate the number of taxpayers in each bin, n_j . To account for the fact that the density distribution around the tax notch determining tax credit eligibility might differ by gender (due, e.g., to gender income gaps), we estimate gender-specific counterfactual distributions. We define an excluded range around the tax credit cutoff $[m_L, m_U]$, such that $m_L < 0 < m_U$, and we then run regressions as the following:

$$n_{j} = \sum_{i=0}^{p} \beta_{i} \cdot (m_{j})^{i} + \sum_{i=L}^{U} \gamma_{i} \cdot 1(m_{j} = i) + u_{j},$$
(4)

where the first term on the right-hand side is a p-th degree polynomial that accounts for potential curvature in the counterfactual density; the second term is an indicator function for bins located in the excluded range. Following Chetty et al. (2011), our baseline approach uses a seventh-degree polynomial (p = 7). To determine the excluded range, we follow the procedure proposed by Kleven and Waseem (2013): the lower bound is determined by visual inspections, determined as the point where excess bunching starts to emerge; the upper bound is computed such that excess bunching below the notch equals the missing mass above the notch.

We can then calculate counterfactual bin counts as the predicted values from equation (4) omitting the contribution of dummies in the excluded range:

$$\widehat{n_j} = \sum_{i=0}^p \beta_i \cdot (m_j)^i.$$
(5)

We estimate excess bunching by comparing the observed and counterfactual gross income distributions:

$$\widehat{B} = \sum_{j=L}^{0} (n_j - \widehat{n_j}).$$
(6)

The excess bunching estimate, \hat{B} , computes the difference between the observed density of taxpayers located in the excluded range and the counterfactual distribution. For instance, a $\hat{B} = 1$ would suggest that the excess mass around the tax notch is 100 percent of the average height of the counterfactual distribution within the dominated area range.

Following Chetty et al. (2011), we compute the standard error of \hat{B} by using a parametric bootstrap procedure in which a large amount of gross income distributions are generated by random resampling the error term u_j . This procedure generates a new set of counts that can be used to calculate new \hat{B} estimates. We can then define the standard error of \hat{B} as the standard deviation of the distribution of \hat{B} that we obtain through this iterative procedure.

4 Gender Differences in Bunching Responses

This section presents our main results: gender differences in bunching responses to the spouse tax credit notch. We then investigate the role of gender norms as the main expla-

nation for this result, and we explore alternative explanations.

4.1 Bunching Responses at the Spouse Tax Credit Notch

Baseline results. We present our main empirical results in Figure 2, which plots the empirical distribution of gross income by gender. To construct this figure, we first group taxpayers in 150 euro bins of gross income, and then we calculate the fraction of taxpayers in each bin around the tax credit cutoff (demarcated by the dashed vertical line). We plot the taxpayers' distribution up to an income level of 25,000 euros.¹⁴ Each graph also compares the observed distributions (blue dots) with the counterfactual distribution (red solid line).





Notes: The figure presents density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). The left-hand side graph focuses on married male taxpayers; the right-hand side graph focuses on married female taxpayers. In each graph, we report the number of taxpayers by 150 euro gross income bin. The graphs also report gender-specific counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2.

The figure shows a clear spike in the number of female taxpayers just below the tax credit cutoff, relative to the counterfactual distribution (right panel). By contrast, the distribution of male income is smooth and does not present any visible spike at the tax credit cutoff (left panel). Although the shape of the income distribution differs by gender (due, e.g., to the gender income gap), we do not detect any visible spikes at other points of the gross income distribution. We estimate excess bunching to be 1.740 times the height of the counterfactual distribution of women, while it is 0.356 for men. The standard error associated with our excess bunching estimate is 0.267 for women and 0.353 for men. The null hypothesis that there is no excess mass at the tax notch relative to the counterfactual distribution is rejected for the female distribution (t-statistics of 6.51), while it is not rejected for the male distribution.

¹⁴In these graphs, we do not make any restrictions on our sample of self-employed taxpayers. In Appendix Figure B1, we show that our results are remarkably similar when we remove spouse tax credit recipients. Their inclusion does not affect our bunching estimate since they are mostly located at an income level well above the tax credit cutoff, thus not affecting the counterfactual distribution. We also show that our results are robust to different polynomial orders (Table B1) and extensive margin responses (Figure B2 and Figure B3).

tion (t-statistics of 1.01). This result does not reflect a lack of power for identifying bunching responses by men: 11,756 married men and 13,084 married women are located in a 500 euros bandwidth just around the spouse tax credit notch.

The bunching estimate suggests that the density of second-earner women located in an income range strictly below the tax credit cutoff is 1.740 times larger than the density that we would have observed in the absence of the policy. This estimate implies that the female marginal buncher reduces her taxable income by 1.740 bins, which corresponds to around 261 euros. It thus suggests that the female marginal buncher reduces her reported income by around 261 euros to let her husband enjoy a tax credit larger than 689 euros, thus increasing the family's net income by around 428 euros. The absence of any bunching response from second-earner men suggests that, ceteris paribus, couples where the husband is the second earner hold around 428 euros less than comparable couples where the wife is the second earner. For the representative family in our sample, this income loss corresponds to around 2.5 percent of the annual net family income.¹⁵

Gender bunching differences around marriage. The previous result indicates a significant gender disparity in the way second-earners respond to the spouse tax credit notch. This difference may not be due to gender norms, but rather to structural differences between genders in various aspects. To offer a first scrutiny on the role of these alternative explanations, we study how the likelihood of reporting income below the spouse tax credit notch depends on the number of years elapsed from marriage. This analysis is useful for two main reasons. First, it helps us understand assortative mating: if couples form assortatively, we would expect to see a gender disparity in the likelihood of reporting income below the spouse credit cutoff even before marriage. In this case, our finding would fail to reflect a behavioral response from second earners' wives to maximize their family income.

Second, the timing of the response can provide suggestive evidence of the reasons underlying our result. If our finding reflects the fact that second-earner wives are more likely to significantly reduce their income when facing economic shocks or other career disruptions, then a gender difference would not necessarily emerge exactly at marriage. Conversely, if the response emerges exactly at marriage, it suggests, though not conclusively, that marriage "activates" the gender identity norm. Gender differences can arise if the husband and wife contribute to separate household public goods ("separate spheres"), with exclusive responsibility assigned by socially prescribed gender roles (Lundberg and Pollak 1993). Women may then spend more time on non-market activities, reducing their effort in market jobs and increasing the likelihood of reporting income just below the spouse tax credit cutoff.

To test these hypotheses, we implement an event-study approach, comparing the probability of reporting income just below the spouse tax credit cutoff between female and male taxpayers by event time, defined as the number of years elapsed from marriage. We define a dichotomous variable, $Bunch_{i,t}$, that is equal to 1 in the first year *t* when individual *i* reports

¹⁵Appendix Figure B4 reports the income distribution of *wage earners* around the spouse tax credit notch. We also find a positive bunching estimate, but it is small and not statistically significant. This is in line with previous evidence that responses from wage earners are substantially dampened (Saez 2010).

gross income in the income interval [2,500-2,840.51].¹⁶

Namely, we run specifications like the following:

$$Bunch_{i,t} = \sum_{k=-4,k\neq-1}^{3} \alpha_k \cdot D_{i,t}^k + \sum_{k=-4,k\neq-1}^{3} \beta_k \cdot F_i \cdot D_{i,t}^k + \gamma_i + \delta_{m(i),t} + u_{i,t}$$
(7)

where F_i indicates female taxpayers and $D_{i,t}^k$ is a dummy variable for k years before and after the marriage. The interaction between a dummy for female taxpayers and years, $F_i \cdot D_{i,t}^k$, omits the year before marriage (denoted by k = -1), so that the event-study coefficient β_k can be interpreted as the gender difference in the probability of bunching in year k, relative to the year before marriage. In the absence of differential pre-existing gender differences in bunching probability, $\beta_k = 0 \forall k < 0$. The inclusion of municipality-by-year fixed effects, $\delta_{m(i),t}$, allows us to compare gender differences in the outcome variable within a given municipality. γ_i are individual fixed effects. Finally, $u_{i,t}$ is an error term. We cluster the standard errors at the individual level.

Figure 3 plots the β_k coefficient estimates and the 95 percent confidence intervals. The figure shows that there is no significant gender difference in bunching during the years leading to marriage: the probability that a woman started to report income just below the spouse tax credit cutoff is not significantly different from the men's probability. This result suggests that couples are not formed in a way that would predict eligibility for the spouse tax credit. It is consistent with studies showing that the impact of taxation on marriage is modest (see, e.g., Alm and Whittington 1995, 1997).¹⁷ As shown in Appendix Figure B5, this result follows the fact that neither men nor women were bunching at the notch over the pre-marriage years. Moreover, the figure shows that gender differences emerge exactly after marriage, suggesting that gender norms might be an explanation for this result.

Summary. Both our bunching and event study results suggest that women are more likely to adjust their own income to maximize their family income. There are several possible reasons for this result. First, women may have lower income potential than men, so adjusting their income can bring greater gains compared to potential career costs. Another possibility is that even if women's predicted income is not lower than men's, they may face more frictions in adjusting their income. All these explanations indicate that couples are making rational decisions to maximize household income. Another possible explanation is that couples abide by a social expectation that positions men as the primary breadwinners and women as secondary earners. This gender norm would make men more hesitant to be referred to as the "dependent spouse", leading them to avoid reporting income below the tax credit cutoff. Conversely, women might not face the same level of societal pressure to avoid the label, as traditional norms might expect them to be more likely to be in a dependent role.

¹⁶In this analysis, our sample is limited to second earners that we can observe both before and after marriage. The sample is composed of 300,792 observations, which is around 11 percent of our full sample.

¹⁷Note, moreover, that our results are conservative if gender identity norms influence assortative mating (Bertrand et al. 2015; Bertrand et al. 2021). Men who are second-earners might hold more progressive gender views than those who are not. Therefore, our sample of second-earners may consist of men who, on average, hold more progressive views than the typical male.

Figure 3: Gender Differences Emerge After Marriage



Notes: This figure presents gender differences in the likelihood of reporting income between 2,500 and 2,840.51 (the spouse tax credit cutoff) on the number of years elapsed since marriage. The figure reports the β_k estimates and 95 percent confidence intervals, computed from equation (7). The model includes individual fixed effects and municipality-year fixed effects.

In the next sections, we attempt to distinguish the role of gender norms from alternative explanations, including the possibility of structural differences by second earners in several characteristics, arguing that the results are largely driven by the former explanation.

4.2 The Role of Gender Norms

In this section, we test the hypothesis that gender identity norms are the main responsible for our results. We present several exercises. We start by focusing on foreign-born immigrants to conduct an epidemiological approach, testing whether gender differences in bunching are relatively larger for couples coming from more traditional societies. We then extend to natives, and we employ two different proxies for gender norms. First, we use cross-cohort differences in the progressivity of gender views. Finally, we use the share of female politicians elected in town councils as a proxy for gender norms. These analyses allow us to evaluate bunching differences not only between genders, but also *within* genders. They also allow us to account for several plausible non-gender norms-related confounding factors.

Epidemiological approach. We propose an epidemiological study to examine gender norms among foreign-born immigrants. Building on the concept of "portability" of cultural factors introduced by Fernández (2007), we aim to explore how individuals may carry aspects of their culture when they emigrate. This approach can help isolate the role of gender norms from other economic and institutional factors. Immigrants face the same markets

and institutions, but they potentially differ in their cultural heritage based on their country of origin. Drawing on Fernández and Fogli (2009), we use past female labor force participation from the immigrant's country of origin as a proxy for gender norms. We retrieve information on female labor force participation from the World Bank database.¹⁸

We start by plotting the origin country's female employment rate by female income. If the "portability" hypothesis has a meaningful impact on bunching responses, we should observe a relatively lower origin country's female employment rate below the cutoff and a shift upward at the cutoff. The left-hand graph in the top panel of Figure 4 yields evidence in favor of this hypothesis (details are reported in Appendix Table B2). Women coming from more conservative societies are those who are more likely to report income just below the spouse tax credit notch.

Although suggestive, this relationship would not allow us to separate gender norms from other competing factors, such as economic factors and institutions of the destination municipality. For instance, it is possible that immigrants self-select into municipalities that are more comparable to those of their home country (Borjas 1987). To assuage this concern, in the following analysis we control for municipality fixed effects, which adjust for selection in the current municipality of residence. In this way, we compare individuals living in the same municipality, but with different gender norms based on their origin country.

The middle graph in the top panel of Figure 4 depicts the relationship between the probability of reporting income in the income interval [2,500-2,840.5] and the origin country's female employment rate. In addition to municipality fixed effects, we also control for other potential confounding factors, such as age and marital status, and we include year fixed effects to gain precision. The figure shows that the likelihood of reporting income below the spouse tax credit cutoff strongly relates to the origin country's female labor force participation. Conditional on their municipality of residence, women born in high-female employment countries are less likely to report incomes just below the spouse tax credit cutoff, compared to women born in low-female employment countries. This relationship is statistically significant (see Appendix Table B3, panel A, column 1).

Despite controlling for some non-gender norms-related factors, our results might still be biased by other unobservable factors that tend to vary between individuals coming from different countries. For instance, we might have captured just different probabilities of being low-income, or a different willingness to misreport income across people from various countries, such as in Fisman and Miguel (2007). To account for this concern, the right-hand side graph in the top panel of Figure 4 depicts the same relationship as before, but demeaning the women's probability of reporting income below the cutoff by the men's probability. This strategy allows us to control for systematic (non-gender-specific) differences between individuals with a different home country. Our result remains qualitatively similar and statistically significant (see Table B3, panel A, column 2).¹⁹

¹⁸Following the classification proposed by the International Labour Organization, female labor force participation is computed relative to the share of the female population older than 15. We use the female employment share observed in 2000. The final sample of our analysis includes information on 189,007 foreign-born taxpayers (9% of which are women), coming from 159 different countries.

¹⁹Appendix Figure B6, Figure B7, and Figure B8 show no evidence of an association between our proxies for

Figure 4: The Role of Gender Norms



Panel A. Epidemiological approach

Notes: This figure presents three approaches to studying the role of gender norms. Panel A presents the epidemiological approach, where we focus on immigrants and compare the probability of reporting income below the spouse tax credit notch by the origin country's female labor force participation. Panel B offers a cross-cohort approach, which rests on leveraging differences in the progressiveness of gender norms across different generations of women. Panel C proposes a cross-municipality approach, which relies on leveraging heterogeneity in the share of female politicians who are elected in the town council in a certain electoral term, over the period covered in our analysis. For each panel, the left-hand side graphs ("unadjusted") plot the variable of interest by the second earner's income bin. The middle graphs ("adjusted") present a binscatter comparing the proportion of "buncher" second earners (i.e., those reporting income between 2,500 and 2,840.5 euros) with the variable of interest. In plotting this relationship, we control for municipality fixed effects, year fixed effects, and individual-specific controls (age and marital status in panels A and C; marital status in panel B). The right-hand side graphs ("adjusted - gender difference") propose a binscatter where the outcome variable of interest is demeaned by the men's outcome and uses the same controls as in the middle graphs. The sample includes taxpayers who declared less than 10,000 euros in the first year of the dataset. See Appendix B.5 for details on the construction of these graphs.

Cross-cohort approach. Our second approach rests on leveraging differences in the progressiveness of gender norms across different generations of women. As shown in the Appendix Figure A5, attitudes towards gender norms have become significantly more progres-

gender norms and bunching responses for men.

sive among younger generations.²⁰ The share of women strongly agreeing with statements that reflect conservative gender views, such as "men are the main responsible for the provision of family needs" or "men should have more right to a job than women", has declined by about 20 percentage points over the past 50 years. Specifically, while around one-third of women born in the early post-war period were likely to agree with these statements, just one-tenth of women born in the late 1990s and early 2000s share this view.

Does this shift away from traditional gender views result in a decreased willingness to be referred to as a dependent spouse? Following our previous approach, we initially plot the average age by female income. If "bunchers" mainly consist of older female second earners, who are more likely to hold traditional gender views, we would expect to see a relatively higher average age below the cutoff and a downward shift at the cutoff. The left-hand side graph in the middle panel of Figure 4 provides evidence consistent with our hypothesis. On average, female bunchers are about 0.4 years older than those located just above the notch (see Appendix Table B2 for numerical estimates). This relationship consistently holds when we look at the relationship between the probability of reporting income below the cutoff and age (middle graph in the middle panel of Figure 4). Controlling for municipality fixed effects, year fixed effects and marital status fixed effects, the graph depicts a clear positive relationship, which is statistically significant (as shown in Appendix Table B3, panel B, column 1). This result is not due to a spurious correlation between gender norms and lifecycle labor market effects (e.g., older generations are both more likely to hold traditional gender norms and to report income just below the spouse tax credit cutoff). As shown in the right-hand side graph of the middle panel of Figure 4, the relationship is still visible when demeaning the outcome variable by the men's outcome (see Appendix Table B3, panel B, column 2 for details).

Cross-municipality approach. Our third approach to approximate gender norms relies on leveraging heterogeneity in the share of female politicians who are elected in the town council in a certain electoral term, over the period covered in our analysis. Our view is that gender norms should be, all else equal, more progressive in municipalities with a higher share of female politicians. There are two (non-mutually exclusive) reasons for using female politicians as an indicator of gender norms. First, traditional gender norms may discourage women from entering politics, a traditionally masculine sector. Therefore, we expect to see more female politicians elected in places with more progressive gender norms. Second, the presence of women in town councils can help challenge traditional gender norms and lead to policies and behaviors that empower women (see, e.g., Miller 2008; Baskaran and Hessami 2023), reducing the likelihood that they limit their income and act as the dependent spouse.

The bottom panel of Figure 4 corroborates our hypothesis that the likelihood of reporting income below the spouse tax credit cutoff is lower in places where more female politicians are elected. The left-hand side graph shows a dip in the share of female politicians just below the cutoff, followed by an upward shift above (numerical estimates are shown in Appendix

²⁰To construct an index of gender norms that varies across cohorts, we use a nationwide survey conducted by the Italian Institute of Statistics (ISTAT) called *Indagine sulle discriminazioni in base al genere, all'orientamento* sessuale, all'appartenenza etnica.

Table B2). The middle graph, which accounts for municipal fixed effects, year fixed effects, age and marital status fixed effects, depicts a significant negative association between the probability of reporting income below the cutoff and the share of female politicians in the town council (see Appendix Table B3, panel C, column 1). This result does not seem to reflect the possibility that these municipalities are different in other factors (e.g., in the share of low-income individuals), as the relationship remains strong when we adjust for the probability of bunching by men (right-hand side graph; see Appendix Table B3, panel C, column 3 for details).

Taken together, the results emerging from these analyses yield converging evidence: gender differences in bunching are mostly driven by larger responses from women who are more likely to share traditional gender norms.

4.3 Other Potential Explanations

Whether gender norms accurately represent gender differences in responses to the spouse tax credit notch depends on the possibility of ruling out alternative explanations. In this section, we will examine several alternative explanations and demonstrate that none of them are supported by the data.

Gender differences in the scope for bunching. In principle, our results could be masking gender differences in scope for bunching at the spouse tax credit notch. However, our focus on self-employees should substantially limit this issue: we compare second earners that reasonably face a similar scope for responding to the policy. We propose an additional exercise to test this assumption. We study whether gender bunching differences emerge at other tax notches or kinks where a discontinuity is present in *own* gross versus net income. At these discontinuities, the factors that could explain a gender difference in scope for bunching are still present, but gender norms would not matter: if responsive, second earners will not be titled as the "dependent spouse". If unobservable gender differences in scope for adjusting income are driving our results, these unobservable characteristics should lead us to observe larger bunching by women at these points of the income distribution as well. Namely, we look at the tax notch created by the tax exemption cutoff and tax kinks created by marginal tax rate changes. Appendix Figure B9 and Figure B10 show the results of this test: in both cases, they provide no clear evidence of gender differences in bunching. If any, we find a larger response from men at the 15,000 euros tax kink. This result suggests that gender differences in scope for bunching are not a plausible explanation for our findings.

Gender differences in the economic incentive to bunch. Another possible interpretation of our findings is the presence of systematic differences by gender in the economic incentive to respond to the policy. Because the spouse tax credit is a negative function of the main earner's income, there are relatively stronger bunching incentives for second earners with a low-income spouse. Therefore, our result could be explained by the fact that "bunchers" are mostly composed of second-earner women whose spouse reports relatively lower income than comparable spouses who do not bunch. A transparent way to evaluate this possibility is to plot the distribution of the main earner's income by the second earner's income. If





Notes: Each graph shows the average value of a variable (vertical axes) against the second earner's income (horizontal axes) for men (left-hand side graphs) and women (right-hand side graphs). We depict the average value of the following variables: spouse's gross income (panels A and B); home ownership (panels C and D); probability of claiming some tax credits, apart from the spouse tax credit (panels E and F); probability of not filling the tax returns at least one year (panel G and H). The red vertical line indicates the spouse tax credit cutoff. The black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals for each bin average. The coefficient estimates of a regression discontinuity analysis are reported in Appendix Table B4.

"bunchers" are second-earner women with a low-income spouse, then there should be a substantial drop in the average main earner income below the spouse tax credit cutoff and a shift upward at the cutoff. We test this hypothesis in Figure 5 (panels A and B), which plots the average main earner income by second earner income's bin.²¹ The figure provides no evidence of a discontinuous change in the main earner's income at the cutoff. This result implies that gender differences in bunching are not consistent with the fact that second-earner women are more likely to bunch because it is relatively more advantageous for their family's net income (see panel A of Appendix Table B4 for numerical estimates). If any, the figure suggests that *male* second-earners should have a larger economic incentive. Given that the tax credit is a negative function of the main earner's income, male second-earners have a stronger economic incentive because their spouse's income is lower than the spouse's income of female second-earners.

Gender differences in homeownership. In our context, knowing who is the homeowner in the couple can be particularly interesting. If the main earners tend to be the homeowners of the family, then they can exert higher pressure on their spouses to respond to the policy. If homeownership systematically differs by the second earner's gender, our results can be explained under this interpretation. Using information on property income from our tax returns data, we test the presence of a discontinuity in home ownership (defined as the probability of declaring positive property income) at the tax notch. Panels C and D in Figure 5 show no evidence of a discontinuity: women reporting income just below the spouse tax credit cutoff are not less likely to be homeowners compared to those just above the cutoff. A similar result also emerges for second-earner men.

Gender differences in information frictions and use of tax benefits. While we already showed that the scope for bunching responses does not vary by second earner's gender, it remains possible that the salience of the policy does. For example, if female second earners tend to be more informed about the tax system, then they would be more likely to claim tax benefits than second-earner men, ceteris paribus. If this is the case, then an alternative interpretation for our results might be that there are gender differences in information frictions. We test this hypothesis by plotting the average probability of claiming any type of tax credit (apart from the spouse tax credit) by second-earner's income. If there are gender differences over this dimension, we should find an increase in the proportion of women claiming tax credits just below the cutoff, and a shift downward above the cutoff. Panels E and F in Figure 5 depict a similar distribution by gender, with no clear effect at the spouse tax credit.²²

²¹To conduct this test, we predict the spouse's income based on a probabilistic match that uses information on the municipality of residence, age, gender, marital status, and income information derived from the spouse tax credit. Moreover, to corroborate these results, we employ an alternative measure of spouse's income which is obtained by applying a machine learning model on a large dataset that includes ground truth data on household incomes. Also with this approach, there is no evidence of a discontinuous change in the main earner's income at the cutoff. The procedure we adopt to predict the spouse's income, as well as the applied machine learning procedure, are described in Appendix C.

²²A regression discontinuity analysis reveals a positive coefficient for female second-earners (see Appendix Table B4, panel C, column 1). However, although statistically significant at usual confidence intervals, the

Gender differences in labor market attachment. Our result could also be consistent with the notion that second-earner women are more likely (or willing) to sacrifice their own income because they have weaker labor market attachment than men. Although this explanation cannot be fully tested, we propose the following exercise. We proxy labor market attachment by the probability of not filing a tax return in a given year. We then plot this probability by second-earner income, separately by gender. If there are gender differences in labor market attachment, we should find an increase just below the cutoff in the proportion of women who happened to not file a tax return, and a shift downward above the cutoff. Panel G and H show no evidence in support of this hypothesis, thus suggesting that gender differences in labor market attachment among second earners are not likely to explain our main findings.

5 The Impact of the Spouse Tax Credit on Women's Choices: A Family-Career Trade-off?

The original purpose of the spouse tax credit was to offer tax relief to low- and middleincome families. It was introduced at a time when women were largely not part of the workforce, with the intention of providing support to families where one person (usually the husband) was the sole earner. However, the policy may have ended up being a barrier to women's careers by encouraging them to report income below a small threshold. This means that the policy may hold back female employment and depress women's careers, generating efficiency losses and gender disparities. In other words, the policy could have created a trade-off in women's choice: the benefits of maximizing family income might come at the expense of their own career prospects.

This section attempts to shed light on how the spouse tax credit distorts women's outcomes. We provide evidence that the policy has real costs for women in terms of persistently reporting low income and missing career opportunities (Section 5.1 and 5.2). However, we also show that couples with a "dependent spouse" are more likely to remain together (Section 5.3). These results suggest that the policy created a family-career trade-off for women.

5.1 Bunching Responses Are Persistent

Although the economic rationale of the spouse tax credit is to offer insurance against labor market shocks, the policy can *persistently* affect the work and income reporting incentives of second-earners. Typically, economic downturns last about 3 to 4 years (see, e.g., Von Wachter 2020). If the bunching response reflects a cyclical income reduction due to an economic downturn, we would expect to see a similar time frame for the bunching response. On the other hand, if being labeled as a "dependent spouse" triggers gender norms suggesting that "men should work in the labor force and women mostly work in the home", we might observe responses lasting longer than the typical duration of a recession.

To study the nature of the bunching response, we test how the probability of reporting in-

magnitude is small and not economically meaningful.

come below the spouse tax credit cutoff evolves over time. Focusing on the first individualspecific episode observed in the data of reporting income below the spouse tax credit cutoff (i.e., in the income interval [2,500-2,840.51]), we estimate the probability that the same woman is still reporting income in that income interval in the *k*-th year following the first episode. In calculating these probabilities, we control for municipality-year fixed effects, age, and marital status, which allow us to filter out municipality-specific shocks and lifecycle effects.

Figure 6 shows that the spouse tax credit persistently affects the income reporting incentives of second-earner females. We find that the probability of reporting income below the cutoff in the year following the first episode of bunching is around 41 percent. This gap survives for several years, well after the duration of a typical recession. After 7 years, about 30 percent of the initial bunchers (and 73 percent of women that bunch for at least two consecutive years) are still reporting income just below the spouse tax credit notch. This result suggests that the policy leads a significant share of second-earner wives to limit their income persistently.



Figure 6: Bunching Responses Are Persistent

Notes: Focusing on the first individual-specific episode observed in the data of reporting income below the spouse tax credit cutoff (i.e., between 2,500 and 2,840.51 euros), the figure reports the probability that the same individual is still reporting income below the cutoff in the years following the first episode. The figure only focuses on female taxpayers, and controls for municipality-year, age, and marital status fixed effects. 95 percent confidence intervals from standard errors clustered at the individual level are also reported.

5.2 Labor Supply Choices

Second earners can respond along two main margins: changes in labor supply or tax evasion and avoidance responses. Our result holds regardless of what margins underlie changes in second earners' reported gross income. Intuitively, whether the second earner's response is through real responses, such as changes in hours worked, or underreporting of true income, will make the main earner eligible for the tax credit anyway.²³ Therefore, distinguishing evasion responses from labor supply responses is not critical to our conclusions. However, since evasion responses and labor supply responses have different normative implications (see, e.g., Chetty 2009), it would still be useful to distinguish between these two margins of behavioral responses. This section attempts to provide suggestive evidence on whether the spouse tax credit holds back the female labor supply.

We access an administrative dataset provided by the Italian Social Security Institute on self-employed taxpayers working as "collaborators" for Italian firms.²⁴ To help understand the nature of this dataset, consider the example of a self-employed domestic worker, who mostly does cleaning duties for some private households (job A). Suppose he or she also does some related work for a private firm (job B), classified as an employee job, with third-party reported income. In our main tax returns data, we observe his or her total reported taxable income, summing up earnings from job A and job B. In the social security data, we will only observe earnings from job B. Although the social security dataset is imperfect in terms of data coverage, we believe it provides us a unique opportunity to study the decision of self-employed workers to take jobs with earnings that are third-party reported, and thus difficult to underreport. Therefore, we expect evasion to be inherently limited in these data, and any response likely reflects "real" labor supply responses.

We study how the spouse tax credit influences labor supply responses by examining the earnings distribution of "collaborator" workers. If the spouse tax credit limits their labor supply, we would expect a *missing mass* in the distribution below the spouse tax credit cutoff: fewer workers are likely to increase their earnings when approaching the spouse tax credit cutoff.²⁵ Since we do not observe their total income, note that this missing mass is likely to be *fuzzy* (instead of sharp, as in our main tax return data where total income is observed).

Figure 7 depicts the distribution of wage recipients by earnings. The graphs clearly show that there is a missing mass in the female distribution (panel B). As explained before, the missing mass is not located just below the threshold, but it presents some fuzziness. By contrast, the male distribution is smooth (panel A). Our interpretation of this pattern is that the spouse tax credit discourages female second earners from increasing their labor supply. The threat of crossing the spouse tax credit cutoff, due to other (unobserved) self-employed income, leads women to limit their labor supply (either through fewer hours of work or by accepting fewer "collaboration" jobs). This implies that part of the observed bunching response documented before reflects a real response.

Does this result also reflect the influence of gender norms? In the spirit of the analysis presented in Section 4.2, we investigate this possibility by examining whether there is a missing mass among workers who are more inclined to hold traditional gender views. Us-

²³This also extends to the possibility of within-couple income shifting (Zinovyeva and Tverdostup 2021).

²⁴These data, described in Appendix A.5, cover a 7 percent random sample of the Italian population. It covers a panel of 862,199 "collaborator" workers observed over the 1998-2021 period.

²⁵Note that we expect a *missing mass* (instead of *bunching*) because these self-employed workers also receive other (unobserved) sources of income. Therefore, as long as they have other positive income, they would be likely to cross the spouse tax credit cutoff if increasing their earnings. As a result, the spouse tax credit notch might limit workers whose earnings are close enough to the spouse tax credit cutoff.

Figure 7: Distribution of Wage Recipients Among Self-Employees



Notes: The figure presents density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). The left-hand side graph focuses on male workers; the right-hand side graph focuses on female workers. In each graph, we report the number of workers by 50 euro gross income bin. The sample is composed of a 7 percent random sample of "collaborator" workers, provided by the Italian Social Security Institute on self-employed taxpayers working as "collaborators" for Italian firms.

ing the social security data, where we do not have information on nationality but have data on the region of residence and education level, we employ two other indicators for gender norms. First, we capitalize on long-standing cultural differences between North and South Italy, with the South traditionally holding more conservative views on gender-related matters. Second, we assume that more educated individuals are likely to have more progressive gender norms. Both these two proxies on the distribution of traditional gender norms by region and education level are supported by responses in survey data.²⁶ Appendix Figure B11 provides suggestive evidence that gender norms influence the distribution of female collaborator workers, but not the distribution of men. We find a substantial decline in the number of low-educated and Southern female workers with earnings below the spouse tax credit cutoff. By contrast, the same distributions of male workers are smooth.

The results presented here suggest the policy ended up being a barrier to women's careers, depressing their labor market outcomes. The policy thus has some real costs for women, which are likely to generate aggregate efficiency losses (Hsieh et al. 2019).

5.3 Marital Stability

This section presents the impact of the spouse tax credit on marital stability. Some previous studies have shown that marital stability depends on women's earnings (see Becker et al. (1977) and Becker (1981) for seminal works; Bertrand et al. (2015), Bursztyn et al. (2017), and Folke and Rickne (2020) for recent compelling evidence). In our context, the effect of the policy on marital stability is not a priori obvious. On the one hand, the policy might

²⁶According to the European Values Survey (version 3.0.1), 31 percent of respondents in Southern Italy strongly agree with the statement "when jobs are scarce, priority should be for men", while in the North the share is 23 percent. Among highly-educated workers, only 9 percent agree with that statement, compared to about 30 percent among the low-educated.

increase the economic dependence of women on their spouses. This dependency can reduce the likelihood of divorce, as women may feel less financially secure to leave the marriage (Oppenheimer 1997). If the tax credit substantially improves the overall financial situation of the family, it can reduce financial stress and potentially improve marital stability. On the other hand, women who are "dependent spouses" may also experience reduced bargaining power within the marriage (Lundberg and Pollak 1996). This can lead to dissatisfaction and, potentially, higher conflict, which could increase the likelihood of divorce. Similarly, if the second earner's lower income results in financial stress or unmet personal career goals, it could also lead to marital dissatisfaction and increased chances of divorce.



Figure 8: The Impact of the Spouse Tax Credit on Marital Stability

Notes: The vertical axes show the share of divorced or legally separated men (left-hand side graph) and divorced or legally separated women (right-hand side graph) against reported gross income (horizontal axes). The red solid vertical line indicates the spouse tax credit threshold. The black circles denote bin averages (by 150 euros); the gray dashed lines show a quadratic fit and 95 percent confidence intervals for each bin average. The coefficient of a regression discontinuity analysis with the divorce rate as the outcome variable and gross income as the running variable is 0.016, with robust standard error equal to 0.004.

To investigate these effects, we plot the divorce rate by female second earners' income. If the spouse tax credit has an impact on marital stability, we should see a sudden change in the likelihood of divorce at the spouse tax credit cutoff. If the policy strengthens (or weakens) marital stability, we would expect to see an increase (or decrease) in this likelihood at the cutoff.

Figure 8 focuses on second-earner women and depicts the probability of divorcing by income bin, along with 95 percent confidence intervals. The figure provides clear evidence that the policy significantly influences marital stability. Second-earner women reporting

income just below the spouse tax credit cutoff are 1.6 percentage points less likely to divorce, compared to second-earner women who report income above the cutoff. This is a quite significant effect, corresponding to about one-fifth of the average divorce rate in the Italian population.

In sum, our findings from this section support the idea that the spouse tax credit helped strengthen the marriages of women who chose to be "dependent spouses." This policy thus turns out to be beneficial for families who adhere to traditional roles, resulting in higher overall net-of-tax family income and longer-lasting marriages. However, this comes at a cost for the women, as those who report income below the spouse tax credit cutoff tend to limit their labor supply.

6 Conclusions

In many countries, tax policies provide financial incentives aimed at reducing the economic burden on families. These incentives often benefit primary earners when their spouses have income below a certain threshold (see Taxing Wages (2024) - Tax and Gender through the Lens of the Second Earner for several examples). Such policies raise relevant questions about the impact of tax incentives on household economic behavior. Do secondary earners intentionally limit their income to let their partners qualify for these benefits? Does the gender of the secondary earner play a role in this decision?

This paper studies these questions by examining the Italian spouse tax credit, which grants a generous tax credit to primary earners if their spouse, designed as the "dependent spouse", reports income below a certain threshold. The policy thus creates an incentive for secondary earners to bunch at the spouse tax credit cutoff. Using a novel administrative dataset, we show large bunching responses from second-earner women, but no response from second-earner men. This result suggests that household decisions are not Pareto-efficient when men are secondary earners.

We interpret our findings as the consequence of traditional gender identity norms. In a society with traditional gender roles, men weigh the benefit of a higher family income against the cost of violating gender norms, which typically see them as primary breadwinners. This social expectation makes men less likely to report income below the tax credit cutoff to avoid being labeled as a "dependent spouse." In contrast, women, often expected to be in dependent roles, may feel less pressure to avoid this label and are more likely to adjust their income accordingly. We consistently find that bunching responses are more concentrated among women who are more likely to hold conservative gender norms. We provide evidence that the policy imposes real costs on women: they consistently report income below the threshold and miss out on job opportunities. However, we also find that couples where the wife is a "dependent spouse" are more likely to stay together. This suggests that the policy forces women to choose between prioritizing family income and marital stability versus pursuing career advancement.

The policy implications of these findings are critical. Our results suggest that tax policies, such as the spouse tax credit, may unintentionally perpetuate traditional gender roles and

exacerbate economic inequalities between men and women. Gender identity norms should be taken into account in the design of a more efficient and equitable tax system. Policymakers are encouraged to rethink the structure of tax incentives to ensure they do not discourage female labor force participation or limit women's economic opportunities.

References

- Akerlof, G. A. and Kranton, R. (2000). Economics and Identity. *Quarterly Journal of Economics*, 115(3):715–753. 1, 10
- Akerlof, G. A. and Kranton, R. (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*. Princeton University Press, Princeton and Oxford. 1
- Albanesi, S., Olivetti, C., and Petrongolo, B. (2023). Families, Labor Market, and Policy. In Lundberg, S. and Voena, A., editors, *Handbook of the Economics of the Family*, chapter 5, pages 255–326. Elvesier. 1
- Alesina, A. and Ichino, A. (2009). L'Italia Fatta in Casa. Indagine sulla Vera Ricchezza degli Italiani. Mondadori. 5
- Alesina, A., Ichino, A., and Karabarbounis, L. (2011). Gender-Based Taxation and the Division of Family Chores. *American Economic Journal: Economic Policy*, 3(2):1–40. 1, 4
- Alm, J. and Whittington, L. A. (1995). Does the Income Tax Affect Marital Decisions? National Tax Journal, 48(4):565–572. 15
- Alm, J. and Whittington, L. A. (1997). Income Taxes and the Timing of Marital Decisions. *Journal of Public Economics*, 64(2):219–240. 15
- Almås, I., Attanasio, O., and Carneiro, P. (2023). Household Decisions and Intra-Household Distributions. In Lundberg, S. and Voena, A., editors, *Handbook of the Economics of the Family, Volume 1*, chapter 3, pages 111–149. Elvesier. 1, 7
- Apps, P. F. and Rees, R. (1996). Labour Supply, Household Production and Intra-Family Welfare Distribution. *Journal of Public Economics*, 60(2):199–219. 3
- Baskaran, T. and Hessami, Z. (2023). Women in Political Bodies as Policymakers. *Review of Economics and Statistics*, (forthcoming). 19
- Becker, G. S. (1981). A Treatise on the Family. Harvard University Press, Cambridge, MA. 26
- Becker, G. S., Landes, E. M., and Michael, R. T. (1977). An Economic Analysis of Marital Instability. *Journal of Political Economy*, 85(6):1141–1187. 26
- Bertrand, M. (2011). New Perspectives on Gender. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, chapter 17, pages 1543–1590. North Holland: Elsevier Science, vol.4b edition. 10
- Bertrand, M. (2020). Gender in the Twenty-First Century. *AEA Papers and Proceedings*, 110:1–24. 10
- Bertrand, M., Cortes, P., Olivetti, C., and Pan, J. (2021). Social Norms, Labour Market Opportunities, and the Marriage Gap Between Skilled and Unskilled Women. *The Review of Economic Studies*, 88(4):1936–1978. 15
- Bertrand, M., Kamenica, E., and Pan, J. (2015). Gender Identity and Relative Income Within Households. *Quarterly Journal of Economics*, 130(2):571–614. 1, 3, 15, 26
- Bhargava, S. and Manoli, D. (2015). Psychological Frictions and the Incomplete Take-up of Social benefits: Evidence from an IRS Field Experiment. *American Economic Review*, 105(11):3489–3529. 4

- Bick, A. and Fuchs-Schündeln, N. (2017). Quantifying the Disincentive Effects of Joint Taxation on Married Women's Labor Supply. In *American Economic Review, Papers & Proceedings*, volume 107, pages 100–104. American Economic Association. 1, 3
- Bick, A. and Fuchs-Schündeln, N. (2018). Taxation and Labour Supply of Married Couples Across Countries: A Macroeconomic analysis. *Review of Economic Studies*, 85(3):1543–1576. 1, 3
- Blau, F. D. and Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3):789–865. 10
- Blundell, R., Chiappori, P. A., Magnac, T., and Meghir, C. (2007). Collective Labour Supply: Heterogeneity and Non-Participation. *Review of Economic Studies*, 74(2):417–445. 3
- Bobonis, G. J. (2009). Is the Allocation of Resources Within the Household Efficient? New Evidence from a Randomized Experiment. *Journal of Political Economy*, 117(3):453–503. 3
- Borella, M., De Nardi, M., and Yang, F. (2023). Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labor Supply? *Review of Economic Studies*, 90(1):102–131. 1, 4
- Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77(4):531–553. 17
- Bursztyn, L., Cappelen, A. W., Tungodden, B., Voena, A., and Yanagizawa-Drot, D. (2024). How Are Gender Norms Perceived? *Working Paper*. 1, 3
- Bursztyn, L., Fujiwara, T., and Pallais, A. (2017). 'Acting Wife': Marriage Market Incentives and Labor Market Investments. *American Economic Review*, 107(11):3288–3319. 3, 26
- Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029. 3
- Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326. 50, 51, 52, 54
- Casarico, A. and Lattanzio, S. (2023). What Firms Do: Gender Inequality in Linked Employee-Employee Data. *Journal of Labor Economics*, (forthcoming). 5
- Chetty, R. (2009). Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy*, 1(2):31–52. 25
- Chetty, R. (2012). Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica*, 80(3):969–1018. 4
- Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence From Danish Tax Records. *Quarterly Journal of Economics*, 126(2):749–804. 4, 7, 12
- Chetty, R., Friedman, J. N., and Saez, E. (2013a). Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings. *American Economic Review*, 103(7):2683–2721. 4

- Chetty, R., Guren, A., Manoli, D., and Weber, A. (2013b). Does Indivisible Labor Explain the Difference between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities. *NBER Macroeconomics Annual*, 27(1):1–56. 4
- Chetty, R., Looney, A., and Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4):1145–1177. 4
- Chiappori, P.-A. (1988). Rational Household Labor Supply. Econometrica, 56(1):63-89. 3
- Chiappori, P.-A. (1992). Collective Labor Supply and Welfare. *Journal of Political Economy*, 100(3):437–467. 3
- Chiappori, P.-A. and Donni, O. (2011). Non-Unitary Models of Household Behavior: A Survey of the Literature. In Molina, A., editor, *Household Economic Behaviors*. Springer Verlag, Berlin. 1, 7
- Chiappori, P.-A. and Mazzocco, M. (2017). Static and Intertemporal Household Decisions. *Journal of Economic Literature*, 55(3):985–1045. 1, 7
- Di Marzio, I., Mocetti, S., and Rubolino, E. (2023). Market Externalities of Tax Evasion. *Bank* of Italy Temi di Discussione, forthcoming. 6
- Donni, O. (2008). Labor Supply, Home Production, and Welfare Comparisons. *Journal of Public Economics*, 92(7):1720–1737. 3
- Fernández, R. (2007). Women, Work, and Culture. *Journal of the European Economic Association*, 5(2-3):305–332. 1, 3, 16
- Fernández, R. and Fogli, A. (2009). Culture: An Empirical Investigation of Beliefs, Work, and Fertility. *American Economic Journal: Macroeconomics*, 1(1):146–177. 1, 3, 17
- Fernández, R., Fogli, A., and Olivetti, C. (2004). Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics*, 119(4):1249–1299. 3
- Field, E., Pande, R., Rigol, N., Schaner, S., and Moore, C. T. (2021). On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms. *American Economic Review*, 111(7):2342–2375. 1
- Fisman, R. and Miguel, E. (2007). Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets. *Journal of Political Economy*, 115(6):1020–1048. 17
- Folke, O. and Rickne, J. (2020). All the Single Ladies: Job Promotions and the Durability of Marriage. *American Economic Journal: Applied Economics*, 12(1):260–287. 26
- Gelber, A. M., Jones, D., and Sacks, D. W. (2020). Estimating Adjustment Frictions Using Nonlinear Budget Sets: Method and Evidence from the Earnings Test. *American Economic Journal: Applied Economics*, 12(1):1–31. 4
- Giuliano, P. (2020). Gender and Culture. Oxford Review of Economic Policy, 36(4):944–961. 3
- Goldin, C. (2014). A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, 104(4):1091–1191. 2, 10
- Goldin, C. and Katz, L. F. (2011). The Cost of Workplace Flexibility for High-Powered Professionals. *Annals of the American Academy of Political and Social Science*, 638(1):45–67. 2

- Guiso, L., Sapienza, P., and Zingales, L. (2006). Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives*, 20(2):23–48. 1, 3
- Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica*, 87(5):1439–1474. 26
- Ichino, A., Olsson, M., Petrongolo, B., and Skogman Thoursie, P. (2024). Economic Incentives, Home Production and Gender Identity Norms. *The Journal of Labor Economics*, (forthcoming). 1, 3
- Keane, M. P. (2011). Labor Supply and Taxes: A Survey. *Journal of Economic Literature*, 49(4):961–1075. 1
- Kleven, H. J., Kreiner, C., and Saez, E. (2009). The Optimal Income Taxation of Couples. *Econometrica*, 77(2):537–560. 1
- Kleven, H. J. and Waseem, M. (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *Quarterly Journal of Economics*, 128(2):669–723. 4, 7, 12
- Lundberg, S. and Pollak, R. A. (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy*, 101(6):988–1010. 14
- Lundberg, S. and Pollak, R. A. (1996). Bargaining and Distribution in Marriage. Journal of Economic Perspectives, 10(4):139–158. 27
- Miller, G. (2008). Women's Suffrage, Political Responsiveness, and Child Survival in American History. *Quarterly Journal of Economics*, 123(3):1287–1327. 19
- Olivetti, C. and Petrongolo, B. (2016). The Evolution of Gender Gaps in Industrialized Countries. *Annual Review of Economics*, 8(1):405–434. 2
- Oppenheimer, V. K. (1997). Women's Employment and the Gain to Marriage: The Specialization and Trading Model. *Annual Review of Sociology*, 23:431–453. 27
- Rubolino, E. (2022). Taxing the Gender Gap: Labor Market Effects of a Payroll Tax Cut for Women in Italy. *CESifo Working Paper No.* 9671. 4
- Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3):180–212. 6, 7, 12, 14
- Von Wachter, T. (2020). The Persistent Effects of Initial Labor Market Conditions for Young Adults and their Sources. *Journal of Economic Perspectives*, 34(4):168–194. 23
- Zinovyeva, N. and Tverdostup, M. (2021). Gender Identity, Coworking Spouses, and Relative Income within Households. *American Economic Journal: Applied Economics*, 13(4):258– 284. 25

Online Appendix (Not for Publication)

Α	Back	cground and Data Appendix	35
	A.1	Personal Income Tax Schedule	35
	A.2	The Spouse Tax Credit	37
	A.3	Take-Up Rate of Spouse Tax Credit	40
	A.4	Gender Norms By Birth Cohort	41
	A.5	Social Security Data	42
B	Rob	ustness Checks and Other Specifications	43
	B .1	Excluding Tax Credit Recipients	43
	B.2	Sensitivity to Polynomial Order and Extensive Margin Responses	44
	B.3	Responses from Wage Earners	48
	B.4	Bunching Before Marriage	49
	B.5	Regression Analysis of Figure 4 and 5	50
	B.6	Epidemiological Approach	55
	B.7	Cross-Cohort Approach	56
	B. 8	Cross-Municipality Approach	57
	B.9	Bunching Responses to Other Tax Discontinuities	58
	B .10	Heterogeneity in Labor Supply Choices	60
C	Mat	ch to Assign Spouse's Income	61
	C .1	Probabilistic Match Approach	61
	C.2	Machine Learning Approach	64

A Background and Data Appendix

A.1 Personal Income Tax Schedule

Taxable income (euros per-year)	Marginal tax rate (%)
If composed only of income from real estate (up to 500 euros)	0
If composed only of retirement income (up to 7,500 euros) + income from land (up to 185,92 euros) + income from a main residence	0
< 15,000	23
15,001-28,000	27
28,001-55,000	38
55,001-75,000	41
> 75,000	43

Table A1: Personal Income Tax Schedule

Note: This table displays information on the Italian personal income tax (IRPEF). Taxpayers are exempted from paying income taxes if their income is composed exclusively of real estate (up to 500 euros) or only from retirement income (up to 7,500 euros) plus income from land (up to 185,92 euros) plus income from a main residence and associated fixtures. The tax base is defined as net of deductible expenses, such as social security and welfare contributions or donations to nonprofit organizations.

Figure A1: Tax Return Sample

FAMILIARI A	C	ARICO							
				CODICE FISCALE	moni		detrazione	n. mesi detrazione figli	
BARRARE LA CASELLA				(il codice del coniuge va indicato anche se non fiscalmente a carico)	a carico	%	affidamento figli	21 anni o più	
C = Coniuge F1 = Primo figlio	1		4	XXXXXXXXXXXXXXX	⁵ 7				
 F = Figlio A = Altro D = Figlio con disabilità 	2	F1 FIGLIO ³ D				7	8	10	
	3	F ² AD							NUMERO FIGLI IN AFFIDO PREADOTTIVO
	4	FAD							CONTRIBUENTE
	5	FAD							

Notes: The figure shows a sample of the Italian tax return where the main earner reports his or her dependent spouse ("coniuge a carico") and dependent children ("figli a carico").

A.2 The Spouse Tax Credit

Main earner's gross income (euros) (1)	Tax credit (euros) (2)	Tax credit (% of tax burden) (3)	Tax credit (% of gross income) (4)
15 000	202 (110*	12 20/	0.00/
-15,000	800-(110*gross income/15,000)	43.2%	9.9%
15,001-29,000	690	12.9%	3.1%
29,001-29,200	700	9.5%	2.4%
29,201-34,700	710	8.4%	2.2%
34,701-35,000	720	7.5%	2.1%
35,001-35,100	710	7.4%	2.0%
35,101-35,200	700	7.2%	1.9%
35,201-40,000	690	6.5%	1.8%
40,001-80,000	690*[(80,000-gross income)/40,000]	1.8%	0.6%
80,001-	0	0	0

Table A2: The Spouse Tax Credit

Note: This table illustrates the main features of the spouse tax credit (law 917/1986). The main earner in a couple is eligible to claim a spouse tax credit if her spouse reports gross income below 2,840.51 euros. Column 2 shows the tax credit amount as a function of the main earners' annual gross income. Column 3 displays how much of the main earner's final tax burden is reduced thanks to the spouse tax credit. The tax burden is calculated by applying the personal income tax schedule (see Table A1) at an income level equal to the median income in each main earner's gross income group. Column 4 calculates the spouse tax credit as a share of the main earner's gross income, computed at the median income level of the corresponding main earner's gross income bracket. Authors' calculations are based on tax returns data.



Figure A2: Gender Norms Index and Female Employment By Region

(b) Female Employment



Notes: The top graph displays the region-level share of respondents agreeing that "when jobs are scarce, men should have priority" from the 2017 European Values Study. The bottom panel depicts female employment rate, using data from ISTAT. Each point corresponds to the region-level average: PIE = Piedmont; AOS = Aosta Valley; LOM = Lombardy; TRE = Trentino Alto-Adige; VEN = Veneto; FRI = Friuli-Venezia Giulia; LIG = Liguria; EMI = Emilia-Romagna; TUS = Tuscany; UMB = Umbria; MAR = Marche; LAZ = Lazio; ABR = Abruzzi; MOL = Molise; CAM = Campania; PUG = Puglia; BAS = Basilicata; CAL = Calabria; SIC = Sicily; SAR = Sardinia.



Figure A3: Summary Statistics of Potential Spouse Tax Credit "Bunchers"

Notes: This figure presents mean values of our sample of potential spouse tax credit "bunchers", defined as those reporting income between 2,840.51 and 4,840.51 euros before marriage (samples based on alternative definitions of the upper limit present similar results). We present mean values separately for men (light blue bars) and women (dark blue bars) of the following variables retrieved from tax returns data: age (logged), immigrant status (dummy), probability of leaving our sample (dummy), home ownership (dummy), claim a tax deduction (dummy), receive a tax credit (dummy), declare corporate income (dummy), declare holding income (dummy), declare pension income (dummy), declare employee income (dummy), declare farming income (dummy).

A.3 Take-Up Rate of Spouse Tax Credit



Figure A4: Take-Up Rate of Spouse Tax Credit

Notes: This binscatter shows the take-up rate of spouse tax credit between married men (black circles) and married women (red diamonds) as a function of the tax credit (as a share of the total gross tax burden). Since taxpayers reporting less than 5,000 euros do not pay taxes, the sample includes all married taxpayers with income above 5,000 euros.

A.4 Gender Norms By Birth Cohort

Figure A5: Gender Norms by Birth Cohort



Notes: The figure shows the share of female respondents that agree with the following statements: i. "men are the main responsible for the provision of family needs" (blue circles); ii. "men should have more right to a job than women" (red squares). We report estimates by age group (horizontal axis). Data from a nationwide survey conducted by the Italian Institute of Statistics (ISTAT) (see *Indagine sulle discriminazioni in base al genere, all'orientamento sessuale, all'appartenenza etnica*).

A.5 Social Security Data

We also use a confidential matched employer-employee dataset on a 7 percent random sample of "collaborator" workers, provided by the Italian Social Security Institute (INPS). Differently from the tax returns data, this dataset is not limited only to Veneto taxpayers, but to the whole Italy. For each unique employer-employee match, we observe information on gross earnings as well as standard demographics, such as gender, date of birth, place of residence, and education. In appendix Table A3, we report summary statistics. Our final sample focuses on workers above age 30 and is a panel composed of 862,199 "collaborator" workers, observed over the 2009-2021 period.

	N of observations (1)	Average (2)	SD (3)
Gross earnings	4,282,988	12,340.25	26,755.33
Age	4,282,988	44.250	9.634
Female	4,282,988	0.355	0.479
College degree	4,282,988	0.292	0.455
South	4,282,988	0.383	0.486

Table A3: Descriptive Statistics of the "Collaborator" Workers Dataset

Note: This table reports summary statistics of the "collaborator" workers dataset. The sample includes 862,199 workers, observed over the 1998-2021 period. The variable "College degree" is a dummy for workers with at least a bachelor's degree. The variable "South" is a dummy for workers living in Southern Italy.

B Robustness Checks and Other Specifications

B.1 Excluding Tax Credit Recipients



Figure B1: Bunching Responses to the Spouse Tax Credit - Excluding Tax Credit Recipients

Notes: These figures present density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). The left-hand side graph focuses on married male taxpayers; the right-hand side graph on married female taxpayers. In each graph, we report the number of taxpayers (by 150 euros bins) for gross income. The graphs also report counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2. Taxpayers who receive the spouse tax credit are excluded from the sample.

B.2 Sensitivity to Polynomial Order and Extensive Margin Responses

In this Appendix, we test the sensitivity of our bunching estimates to some assumptions that we make to estimate bunching responses. First, we test whether our estimates are sensitive to different polynomial orders used to estimate the counterfactual distribution. Because we explicitly estimate the upper bound of the dominated range, z_u , to ensure that excess bunching equals missing mass, one source of bias in z_u is functional form misspecification. We therefore carry out a sensitivity analysis with respect to the polynomial degree p. In Appendix Table B1, we show that our baseline bunching estimates are not substantially affected by different order choices of the polynomial.

Second, we test the sensitivity of our estimates to extensive margin responses. Our methodology is robust to extensive margin responses, including real participation responses as well as movements between the formal and informal sectors, as long as they do not take place locally around the tax notch. In this case, the determination of the excluded range should not be substantially affected by extensive margin responses as long as it is defined by a narrow range above the cutoff. In fact, since intensive (bunching) margin responses only occur locally, this approach allows us to identify only intensive margin responses. Yet, since the convergence method described above might rely on a larger range, it is potentially sensitive to extensive margin responses. We test the sensitivity of our estimates to different upper bounds on the income bandwidth where the convergence method is allowed to work. Appendix Figure B2 and Figure B3 show that our bunching estimates hold regardless of how we define the income window of interest. Therefore, extensive margin responses do not appear to be a meaningful source of bias.

Polynomial order:	Bunching estimate for women
5	1.775***
6	1.227***
7 (baseline)	(0.204) 1.740***
8	(0.267) 2.349***
9	(0.363) 1.640***
10	(0.383) 1.561***
	(0.389)

Table B1: Robustness to Polynomial Order

Notes: This table tests the sensitivity of our bunching estimate to the polynomial order's choice. We report bunching and standard error estimates with polynomial order ranging from 5 to 10 (including our baseline estimate that uses a seventh-degree polynomial).



Figure B2: Bunching Responses - Men - Different Income Bandwidths

Notes: These figures present density distributions around the tax notch determining eligibility for the spouse tax credit for married men, changing the income upper bound (10th percentile –5,850 Euros–, 25th percentile –12,000 Euros–, 50th percentile –21,000 Euros–, 75th percentile –36,000 Euros–). In each graph, we report the number of taxpayers (by 150 euros bins) for gross income as well counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2.

(a) Upper bound: 10th percentile (b) Upper bound: 25th percentile b: 4.229 (1.372) b: 1.586 (0.344) 5500° 5500-5000-5000 A500-A500. A000. A000. Taxpayers count Taxpayers count 3500-3500 3000 3000 2500 2500 2000 2000. 1500 1500 ,000 ,000. ,000 2000 3000 5000 6000 10000 ,2000 4000 1,000 5000 1000 ,000 Gross income Gross income (c) Upper bound: 50th percentile (d) Upper bound: 75th percentile b: 2.227 (0.322) b: 1.747 (0.196) 5500-5500-5000 5000. A500-4500 A000-A000 Taxpayers count Taxpayers count 3500 3500. 3000 3000 2500-2500. 2000. 2000. 1500-1500 1000-,000 ,005 ,000 Gross income Gross income

Figure B3: Bunching Responses - Women - Different Income Bandwidths

Notes: These figures present density distributions around the tax notch determining eligibility for the spouse tax credit for married women, changing the income upper bound (10th percentile –5,850 Euros–, 25th percentile –12,000 Euros–, 50th percentile –21,000 Euros–, 75th percentile –36,000 Euros–). In each graph, we report the number of taxpayers (by 150 euros bins) for gross income as well counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2.

B.3 Responses from Wage Earners

Figure B4: Bunching at the Spouse Tax Credit Notch of Wage Earners



Notes: These figures present density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line. The sample includes employees taxpayers: the left-hand side graph focuses on married males; the right-hand side graph on married females. In each graph, we report the number of taxpayers (by 150 euros bins) for gross income. The graphs also report counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2.

48

B.4 Bunching Before Marriage

Figure B5: Bunching Responses to the Spouse Tax Credit - Before Marriage

(a) Married Men (before marriage) (b) Married Women (before marriage)



Notes: These figures present density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). Left-hand side graphs focus on married male taxpayers, in the years preceding the marriage; right-hand side graphs on married female taxpayers in the years preceding the marriage. In each graph, we report the number of taxpayers (by 150 euros bins) for gross income. The bottom graphs also report counterfactual distributions (in red), bunching estimates and bootstrapped standard errors, computed as described in Section 3.2.

B.5 Regression Analysis of Figure 4 and 5

In this appendix, we detail the regression analysis we conducted to create the plots in Figure 4 and Figure 5.

"Unadjusted" model. To generate the plots in Panels A.i, B.i, and C.i, we estimate a regression discontinuity design model exploiting the 2,840 Euros threshold as the cutoff of the analysis. More specifically, the model provides non-parametric RD estimates following Calonico et al. (2014) under optimal bandwidth and polynomial order selection. The results of this analysis are reported in Table B2.

Adjusted model. To create the binscatter plots reported in Panels A.ii, B.ii, and C.ii, we estimate the following model:

$$Bunch_{i} = \beta \cdot Tax payer Feature_{i} + \gamma_{m(i)} + \delta_{t(i)} + \phi \cdot X_{i} + u_{i}, \tag{8}$$

where the explanatory variable of our interest is *Taxpayer Feature*_{*i*,*t*} which varies across individuals (or by municipalities in some specifications). First, in Panel A.ii, this variable captures the female labor force participation in the origin country, expressed in percentage, in 2000. Second, in Panel B.ii, it represents the age of the taxpayer, expressed in years. Finally, in Panel C.ii, it captures the percentage of female politicians elected in the city council of the town where the taxpayer resides.

The other variables of the regression are defined as follows. *Bunch_i* is equal to 1 if taxpayer *i* reports income in the income interval [2,500-2,840.51]. $\gamma_{m(i)}$ represent municipal fixed effect. To gain precision, we also include year fixed effects, $\delta_{t(i)}$, and we control for marital status and age of the individual in X_i (only in Panels A and C).

The sample in the analysis includes only foreign-born immigrants for the analysis of Panel A, while it comprises all taxpayers for the exercises of Panels B and C. The coefficient of interest, β , computes the effect of the main treatment *Taxpayer Feature*_i on the probability of bunching.

Numerical results for the graphical analysis shown in Figure 4 (Panels A.ii, B.ii, and C.ii) are reported in Table B3, column 1, and show statistically significant estimates.

Adjusted - gender difference. Finally, to create the binscatter plots reported in Panels A.iii, B.iii, and C.iii, we estimate the following model:

$$Bunch_{i} = \beta_{1} \cdot Tax payer Feature_{i} + \beta_{2} Female_{i} + \beta_{3} \cdot Tax payer Feature_{i} * Female_{i} + \gamma_{m(i)} + \delta_{t(i)} + \phi \cdot X_{i,t} + u_{i}, \quad (9)$$

where the variable *Female*_{*i*} indicates female taxpayers and all the other terms are defined as in equation (8). The coefficient of interest, β_3 , computes the effect of the main treatment *Taxpayer Feature*_{*i*} on the probability of bunching in difference across genders. The results of this analysis are reported in Table B3, column 2, and show that the impact of the explanatory variables is significantly different between male and female taxpayers, as already suggested by Figure 4 (Panels A.iii, B.iii and C.iii).

Other potential explanations. Table B4 reports the RD analysis shown in Figure 5. Also in this case the estimated model provides non-parametric RD estimates following Calonico et al. (2014), under optimal bandwidth and polynomial order selection, and it focuses on the threshold of 2,840 Euros. Except for panel C, which provides a statistically significant but very small coefficient, all variables are balanced around the threshold regardless of the gender of the taxpayer.

Dep. var.:	Female labor force participation (origin country)	Taxpayer age	Sh. female politician (town council)
RD estimate (Above 2,840 Euros cutoff)	1.360*	-0.377*	0.0005
Ν	32,075	373,665	366,741

Table B2: The Role of Gender Norms: Numerical results – Figure 4, column 1

Notes: This table provides numerical estimates to the test reported in Figure 4, column 1. The table shows non-parametric RD estimates following Calonico et al. (2014) under optimal bandwidth and polynomial order selection. The threshold of the analysis is the 2,840 Euro cutoff.

Table B3:	The Role of	Gender I	Norms:	Numerical	results -	Figure 4,	columns 2	2 and 3
						()		

	Dep. var.: Bunching around the threshold			
	Effect for female taxpayers	Effect for female taxpayers in diff. wrt male taxpayers		
Panel A: Female labor force participation - Origin country	-0.00029***	-0.00033***		
	(0.00009)	(0.00009)		
Ν	22,277	53,522		
Panel B: Age	0.00012^{***}	0.00009**		
	(0.00004)	(0.00004)		
Ν	224,142	499,810		
Panel C: Perc. female politicians - Town of residence	-0.01651***	-0.01036**		
	(0.00573)	(0.00444)		
Ν	222,818	496,509		

Notes: This table provides numerical estimates to the test reported in Figure 4, column 2, according to equation 8. The dependent variable is a dummy variable that indicates bunching behavior below the 2,840 Euro threshold. The main explanatory variables capture the female labor force participation in the origin country in panel A, the taxpayer's age in panel B, and the percentage of female politicians elected in the town council in the municipality of residence.

	Female taxpayers	Male taxpayers
Panel A: Spouse's income		
RD estimate (Above 2,840 Euros cutoff)	59.3355	-15.2071
	(251.3967)	(86.1718)
Ν	367,870	541,271
Panel B: Property ownership		
RD estimate (Above 2,840 Euros cutoff)	-0.0060	-0.0029
	(0.0114)	(0.0091)
Ν	373,665	548,920
Panel C: Other tax credits		
RD estimate (Above 2,840 Euros cutoff)	0.0265***	0.0059
	(0.0067)	(0.0070)
Ν	373,665	548,920
Panel D: Labor market attachment		
RD estimate (Above 2,840 Euros cutoff)	-0.0031	0.0050
	(0.0088)	(0.0100)
Ν	373,665	548,920

Table B4: Alternative explanations: Numerical results – Figure	• 5
--	------------

Notes: This table provides numerical estimates of the test reported in Figure 5. The table shows non-parametric RD estimates following Calonico et al. (2014) under optimal bandwidth and polynomial order selection. The threshold of the analysis is the 2,840 Euro cutoff.

B.6 Epidemiological Approach

Figure B6: Female Employment Rate in Origin Country by Income



Notes: This figure the average origin country's female labor force participation by men's income (left-hand side graphs) and women's income (right-hand side graph). The red vertical line indicates the spouse tax credit cutoff. The black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals for each bin average.

B.7 Cross-Cohort Approach



Figure B7: Average Age by Income

Notes: This figure depicts the average age by men's income (left-hand side graphs) and women's income (right-hand side graphs). The red vertical line indicates the spouse tax credit cutoff. Black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals.

B.8 Cross-Municipality Approach

Figure B8: Share of Female Politicians in Town Councils by Income



Notes: This figure depicts the average share of female politicians elected in the town council by men's income (left-hand side graphs) and women's income (right-hand side graphs). The red vertical line indicates the spouse tax credit cutoff. Black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals.

57

B.9 Bunching Responses to Other Tax Discontinuities



Figure B9: Bunching Responses to the Tax Exemption Cutoff

Notes: These figures present density distributions around the tax notch determining tax exemption for selfemployees (4,800 Euros, denoted by the dashed vertical line). The left-hand side graph focuses on male taxpayers; the right-hand side graph on female taxpayers. In each graph, we report the number of taxpayers (by 100 euros bins) for taxable income. The graphs also report counterfactual distributions (in red), bunching estimates, and bootstrapped standard errors, computed as described in Section 3.2. To avoid a distorted estimate of the counterfactual distribution, we exclude taxpayers with income below the spouse tax credit cutoff.



Figure B10: Bunching Responses to the Other Tax Discontinuities

Notes: These figures present density distributions around the tax bracket cutoffs, 15,000, 28,000, 55,000 and 75,000 Euros (denoted by the dashed vertical lines). Left-hand side graphs focus on male taxpayers; right-hand side graphs on female taxpayers. In each graph, we report the number of taxpayers (by 150 euros bins) for taxable income. To avoid a distorted estimate of the counterfactual distribution, we exclude taxpayers with income below the spouse tax credit cutoff.

B.10 Heterogeneity in Labor Supply Choices



Figure B11: Heterogeneity in Labor Supply Choices

Notes: The figure presents density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). The left-hand side graphs focus on male workers; the right-hand side graphs focus on female workers. In each graph, we report the number of workers by 50 euro gross income bin. Panel A reports separate distributions for workers living in Southern (red squares) and Northern Italy (blue circles). Panel B shows separate distributions for workers with a college degree (red squares) and without (blue circles). The sample is composed of a 7 percent random sample of "collaborator" workers, provided by the Italian Social Security Institute on self-employed taxpayers working as "collaborators" for Italian firms.

C Match to Assign Spouse's Income

The individual tax return data don't allow us to match married taxpayers and to measure the spouse's income. To overcome this data limitation, we apply several techniques to predict the spouse's income of each (married) taxpayer. First, we begin with a probabilistic matching approach, and, second, as a robustness check, we train a machine learning model.

C.1 Probabilistic Match Approach

The main method we rely on is a probabilistic matching approach. To conduct this prediction, we use all available demographic information to be as accurate as possible. With this goal, we define for each individual a "reference" group of potential spouses, given their age and municipality of residence. More specifically, we rely on three main criteria:

- 1. For each individual, we focus on taxpayers of the other gender who are married. Therefore, the total pool of taxpayers considered to conduct this simulation is a group of 1,329,135 individuals;
- 2. For each individual, aged from 18 to 80 years old, we define the potential spouses as those in an age window of -5/+5 years with respect to the taxpayer's age. The average size of an age group averaged across all municipalities, is 294 individuals;¹
- 3. For each individual, we consider taxpayers who reside in the same municipality.

Relying on criteria 1., 2. and 3. allows us to assign each married taxpayer to a group of potential spouses with the same civil status, with a "similar" age, and residing in the same municipality. The predicted spouse's income is simply the average income of the reference group for each taxpayer. Figure C1 shows the average predicted spouse income by age for male and female spouses. As expected, this predicted income increases with age and the gender pay gap emerges from the figure.

Lastly, we conduct a second simulation, which is a refinement of the main method, in which we constrain the pool of potential spouses to the taxpayers receiving the spouse tax credit. Given this new constraint, the considered taxpayers are by definition only married to individuals whose declared income is situated below 2,840.51 Euros. Therefore, this refined method only allows us to simulate the income of the spouses of the taxpayers who bunch at the tax credit threshold. This refined simulation is more precise than the previous method as the average age group of potential spouses averaged across municipalities, drops to 52 individuals. Moreover, this new method allows to conduct two additional analyses. First, we can test the precision of the previous simulation, less accurate but available for the universe of taxpayers. In particular, the correlation between the spouse's income simulated with this refined method and the main one is very high, 0.81 (significant at standard levels).

¹Importantly, we also conduct robustness checks using different time windows to define the relevant age. In particular, we also focus on windows -2/+2, -3/+3, and -10/+10 and the results are always similar. These outcomes are reported in Figure C3 and the variable is always continuous around the threshold, as indicated by the coefficients of the RD analysis displayed in the Figure's notes.

This result is reassuring and suggests that also the main method yields a reasonably accurate simulation. Second, we can use this refined simulated income to have a more precise idea of the distribution of the spouse's income for female and male taxpayers below the threshold. This is reported in Figure C2. As expected, male partners have considerably larger incomes than female partners.



Figure C1: Predicted spouse's income over spouse's age

Notes: This scatterplot shows the predicted spouse's income for married men (black circles) and married women (red diamonds) as a function of their age. The income is predicted according to the procedure described in the appendix C.





Notes: This scatterplot shows the predicted spouse's income for married men (black circles) and married women (red diamonds) as a function of taxpayers' declared income. Only taxpayers with income below the tax credit threshold are considered. The spouse's income is predicted according to the refined procedure described in the appendix C.

Figure C3: Test on spouse's income for female taxpayers - Alternative predictions

A. Time window of two years B. Time window of three years C. Time window of ten years



Notes: Each graph shows the average spouse's income (vertical axes) against the second earner's income (horizontal axes) for women. The red vertical line indicates the spouse tax credit cutoff. The black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals for each bin average. Panel A shows the spouse's income simulated relying on a time window of two years, Panel B on a time window of three years, and Panel C on a time window of ten years. The coefficient of an RD analysis with gross income as running variable is -25.5 (SE=275.3) for panel A, 19.6 (SE=266.8) for panel B, and 95.6 (SE=202.6) for panel C.

C.2 Machine Learning Approach

To corroborate the results obtained with the probabilistic match approach, we predict the spouse's income with a different methodology, relying on machine learning techniques.

Specifically, we train a machine learning algorithm with the goal of predicting a spouse's income and we rely on fine-grained ground truth data on household income which are provided by the Survey on Household Income and Wealth (SHIW), collected by the Bank of Italy. This survey collects granular information on family income and wealth and it is conducted regularly by the institute with the goal of providing a representative picture of the country. We focus on the waves 1989-2020 to create a dataset with information on taxpayers' income and their spouses. The final dataset contains information on 54,128 taxpayers, representing 27,064 Italian households.² In terms of predictors, we focus on a large number of individual characteristics: age and gender of the individual, type of main income, net declared income, real estate income, year of the survey, size of residence municipality, and geographic location. Finally, to improve the predicting capacity of the model, we do not focus on the continuous version of the variable to predict, the spouse's income, but we create a discrete version, which captures the income quintile of the spouse.

This large dataset allows us to train a machine learning algorithm that is able to predict a spouse's income quite accurately. In particular, we employ the XGBoost (Extreme Gradient Boosting) algorithm, a gradient boosting framework known for its efficiency, accuracy, and ability to handle both linear and non-linear relationships in the data. To enhance the model's predictive performance and prevent overfitting, we use a 5-fold cross-validation approach. This method divides the dataset into five subsets, training the model on four subsets while validating it on the remaining one, iteratively. Furthermore, we implement a grid search to fine-tune the hyperparameters, systematically exploring combinations of key

²The dataset has a panel component which we exclude from the final data in order to avoid an overrepresentation of certain households. When information on a household is available in multiple years, we focus on the most recent one.

parameters such as learning rate, maximum depth, and number of estimators to optimize model accuracy. Importantly, we find that our model performs well in predicting the income of the spouse with a mean AUC-ROC of 0.784, as shown in Figure C4, Panel A, which shows a strong correlation between true spouse's income (reported in the vertical axis) and predicted one (reported in the horizontal axis).

With this alternative measure of the spouse's income, we can now conduct the test to check the continuity of the main earner's income at the cutoff. This test is reported in Figure C4, Panel B. Also with this alternative measure, similarly as in Figure 5, Panel B, we do not see any discontinuous change in the income of the spouse for women taxpayers, providing additional support to the previous results.

Figure C4: Test on spouse's income for female taxpayers - Alternative predictions



Notes: The left plot shows the correlation between the true spouse's income as reported in the Survey on Household Income and Wealth (SHIW) and the predicted spouse's income, generated with the ML model. The right graph shows the average value of the predicted spouse's income (vertical axes) against the second earner's income (horizontal axes) for female taxpayers. The red vertical line indicates the spouse tax credit cutoff. The black circles denote bin averages (by 150 euros), and the gray dashed lines show a quadratic fit and 95 percent confidence intervals for each bin average. The coefficient of an RD analysis with gross income as running variable is -0.0052 (SE=0.0101).