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

Tommaso Giommoni\(^1\) and Enrico Rubolino\(^2\)

\(^1\)ETH Zurich
\(^2\)University of Lausanne

June 15, 2022

Abstract

The standard model of household behavior predicts that couples cooperate to maximize family income. This paper shows that gender identity norms represent an important friction preventing family income maximization. For identification, we focus on an Italian policy that grants a large tax credit to the main earner in a couple when the second earner reports income below a cutoff. Using novel tax returns data, we show large bunching responses at the tax credit cutoff from second earner women, but no response from second earner men. This result suggests that household decisions are not Pareto-efficient when men are the second earner within the couple. Gender differences in bunching emerge after marriage, exacerbate with childbirth, and do not reflect any gender-specific difference in scope for bunching. In support of the view that gender norms drive our results, we find that gender differences in bunching are relatively larger among immigrants coming from more conservative societies, and natives living in more gender-traditional municipalities. Additionally, these results have important implications on female labor market outcomes: we show that the spouse tax credit persistently limits women's careers and amplifies the gender income gap.

JEL codes: H24; H31; J16.
Keywords: gender norms; gender inequality; spouse tax credit; income taxation.

*We thank Sonia Bhalotra, Marius Brulhart, Marko Köthenbürger and seminar participants at the University of Essex for comments.

Tommaso Giommoni: ETH Zurich, Department of Management, Technology, and Economics, Leonardstrasse 21, 8092 Zurich, Switzerland; E-mail: giommoni@ethz.ch; +41 44 632 29 25; Enrico Rubolino: University of Lausanne, Faculty of Business and Economics, Quartier de Chambronne, Internef 261, 1015 Lausanne, Switzerland; enrico.rubolino@unil.ch; +41 21 692 34 84. This version: June 2022.
1 Introduction

The standard model of household behavior states that individuals maximize a single household utility under a budget constraint (see, e.g., Chiappori 1988; Chiappori 1992; Apps and Rees 1996; Blundell et al. 2007). In this model, individuals care about their family welfare, over and above their own well-being. Taken literally, the model predicts that the distribution of income within the couple would not matter: couples cooperate to maximize the family income and household resources are allocated in the most economically efficient manner. This paper shows that gender identity norms represent an important friction preventing family income maximization. We provide empirical evidence that household income maximization choices are not Pareto-efficient when men are the second earner within the couple.

For identification, we take advantage of the Italian spouse tax credit: the main earner in a couple receives a large tax credit if the second earner reports gross annual income below 2,840.51 euros. The size of the tax credit is a negative function of the main earner’s gross income. For main earners reporting less than 15,000 euros, the tax credit accounts, on average, for more than 40 percent of their tax burden (that is around one-tenth of their gross income). The tax credit then linearly decreases with income, and it is eventually phased out for main earners reporting more than 80,000 euros. The policy thus offers a large cash transfer to low-income families. Over the last decade, we observe that more than one-third of Italian families received the spouse tax credit. Because men tend to earn more than their wives in the vast majority of couples, a pattern observed virtually everywhere else in the world (Blau and Kahn 2017), there are large gender gaps in the spouse tax credit take-up rate.

The spouse tax credit offers a propitious testing ground for evaluating the standard model of household behavior. Since the Italian tax system is individually assessed, the policy creates a substantial “notch” in the budget constraint of family: a discontinuity in the choice set of second earner gross income versus family net income (Saez 2010; Chetty et al. 2011; Kleven and Waseem 2013). Under the standard model, the tax notch should induce second earners, who would otherwise report more income, to bunch right at the tax credit cutoff. By contrast, in a world where gender identity norms matter, the decision to bunch would be the result of a cost-benefit analysis, where the benefit of having larger family net income would be discounted by the cost of violating the adopted gender norm.1 For instance, in couples

---

1Gender identity norms prescribe how men and women should behave (see the seminal contributions by Akerlof and Kranton 2000 and Akerlof and Kranton 2010). Departures from the set of prescriptions defining the gender identity norm, such as who should be the main earner in the couple, would generate psychological costs and affect choices. Bertrand et al. (2015) provide a notable example of the non-monetary costs associated with violating gender identity norms: couples where the wife outearns the husband are more likely to divorce and to report lower marriage satisfaction in survey.
that embrace the male breadwinner model, men would adopt behaviors that allow them to “fill the gap” when the male breadwinner model fails. This kind of gender identity norm would dissuade second earner men from reporting income below the tax credit cutoff, but would not deter second earner women from bunching. Therefore, in the presence of traditional gender norms, the spouse tax credit can backfire, with important (unintended) negative impacts on gender income inequalities and aggregate economic inefficiencies.

Building on this idea, we combine novel tax returns data with a bunching approach. Our results provide striking evidence that gender identity norms shape individual behavior in reporting income.\(^2\) We find sizable bunching at the tax credit cutoff from second earner women, but no response from second earner men. There is excess bunching below the tax notch by around 1.2 times the height of the counterfactual female income distribution, suggesting that the density of second earner women located in an income range strictly below the tax credit cutoff is 1.2 times larger than the density that we would have observed in the absence of the policy.\(^3\)

Our bunching estimate suggests that the female marginal buncher reduces her reported income by about 186 euros to let her husband to enjoy a tax credit of nearly 700 euros, thus increasing family (reported) net income by around 514 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 514 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 percent of the annual net family income.

We then examine the distributions of spouses’ income before and after marriage. In family specialization models à la Becker (1991), married men mostly focus on working, while their wives are responsible for non-market production. If marriage “activates” gender identity norms, gender differences in bunching rate would thus emerge just after marriage. Using an event study approach, we show that gender differences in bunching rate emerge right after marriage. By contrast, we find no significant difference before marriage, suggesting that pre-existing factors determining tax credit eligibility do not affect the formation of couples. We also find that gender difference in bunching response exacerbates with motherhood, thus suggesting that the spouse tax credit strengthens child penalties (Kleven et al. 2019b; Casarico and Lattanzio 2021).

\(^{2}\)Following the modern public economics literature (Saez et al. 2012), we focus on taxable income responses, which incorporate both real effects, such as labor supply responses, and tax evasion or tax avoidance responses.

\(^{3}\)We 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 be located close to the tax credit cutoff.
Because women and men may sort into different occupations, one concern is that our finding reflects structural differences by gender in scope for responding to discontinuities in the budget constraint. For instance, women can sort into occupations that allow them more flexibility in adjusting their labor supply (Goldin 2014). To assuage these concerns, we investigate whether gender differences in bunching emerge at other points of the income distribution where the marginal tax rate discontinuously changes. We find no evidence of different responses by gender at other tax kinks. This rules out the possibility that our main finding would simply reflect gender differences in scope for bunching (either for evasion or labor supply reasons).

To analyze whether gender norms are responsible for our results, we present two additional results. First, we conduct an epidemiological study of gender norms based on foreign-born immigrants. We show that gender differences in bunching rate are relatively larger among immigrants coming from countries with more traditional gender norms. For instance, gender differences in bunching rate are much larger among immigrants born in places with lower female employment, such as Iraq or Saudi Arabia, than in places with higher female employment, such as China or Norway. This result holds even adjusting for their current municipality of residence. Second, we find larger bunching differences in gender-traditional municipalities, identified as those i) that were less likely to support the deregulation of abortion in a 1981 referendum; ii) where fewer female politicians are elected; iii) located in rural areas. At the individual level, we detect stronger responses among older women, that tend to report more conservative views on questions about gender norms in survey data.

In the final part of the paper, we study whether the spouse tax credit has any impact on secondary earners’ career and gender inequalities. Although the economic rationale of the spouse tax credit is to offer insurance against labor market shocks, the policy can persistently affect the work (or income reporting) incentives of second earners. We assess the likelihood of lingering in bunching responses, separately by gender. Focusing on the first individual-specific bunching episode observed in the data, we find that the probability of bunching in the successive year is around 17 percentage points larger for women with respect to men (41 versus 24 percent, respectively). This gender difference survives for many years: after 7 years since the first bunching period, there is still a significant 20 percent probability of a bunching response from women, while the corresponding probability for men is not significantly different from zero.

We then show that our results have important implications for gender income gaps. We find that gross incomes of married women are 4 log points lower than those of married men at the first income decile, that is where the spouse tax credit strikes. Significant gender gaps do not emerge at other income deciles, with the
notable exception of the top decile. In support of the argument that the tax credit contributes to exacerbate gender income inequalities, we show that the gender gap in the bottom income decile is relatively smaller among unmarried individuals, where the spouse tax credit does not matter by definition. Although we cannot rule out alternative explanations for the emergence of a gender gap at this point of the income distribution (e.g., self-selection in married status by low-income women), this result seems to suggest that the spouse tax credit significantly contributes to gender income inequalities.

1.1 Literature Contribution

This paper contributes to three literatures. First, our results show that cultural factors, such as gender identity norms, significantly shape individual behaviors. This is consistent with the seminal contribution from Akerlof and Kranton (2000) and Akerlof and Kranton (2010), and it is in line 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; Alesina et al. 2013; Giuliano 2022). Our analysis complements previous work on the link between gender norms and gender differences in labor market outcomes. Previous studies have focused on divorce rate and couples’ satisfaction (Bertrand et al. 2015), childcare allocation (Ichino et al. 2021), labor market choices following a spouse’s layoff (Halla et al. 2020) or misreporting income information in survey (Roth and Slotwinski 2021).

Our analysis focuses on a novel, perhaps more compelling, outcome: adopting behaviors that maximize family income. To our knowledge, we also offer the first evidence of monetary costs that couples face where men are secondary earners.

Second, our results emphasize how policy-making can backfire in presence of traditional gender norms. By incentivizing second earner women to report income below a (small) income threshold, we show that a spouse tax credit limits women’s income, exacerbating gender income gaps. This result relates with studies showing how public policies can hold back female employment. In most countries, taxes and benefits depend on one’s marital status and tend to reduce the labor supply of the secondary earner. Guner et al. (2012) show that switching to a tax system in which married individuals can file taxes separately would substantially increase female labor participation. More recently, Borella et al. (2022) show that eliminating marriage-related provisions in the US would significantly increase married women’s

4A few studies have also elicited gender attitudes directly from survey data. Levine (1993), for example, finds that attitudes, as reflected in responses to General Social Survey questions, are an important predictor of female labor force participation. Focusing on Austria, Vella (1994) shows that gender attitudes significantly predict female participation in market work. Using World Value Survey data, Fortin (2005) finds that gender roles explain work and fertility outcomes across developed countries.
labor market participation.

Finally, our findings have implications for optimal taxation of couples. Kleven et al. (2009) study optimal income taxation of couples in a unitary model of family decision making. In their model, the optimal tax schedule crucially depends on whether the second earner labor force participation depends on high labor market opportunities or low home production ability. In the former case, the optimal tax scheme displays positive tax rate on second earners and negative jointness: the tax rate on one earner declines with the spouse’s earnings. In the latter case, the optimal solution is a negative tax rate on the second earner with positive jointness, whereby the second earner subsidy is phased out with primary earnings. This solution is somehow close to the design of the Italian spouse tax credit: the second earner tax distortion would tend to zero asymptotically as income of the main earner becomes large.

We show that the optimal response of “bunching” below the tax credit cutoff is observed only for female second earners. Since women are still much more likely to be secondary earners in most countries, our result is important for virtually every tax system in the world. Our findings imply that gender identity norms directly enter in the family utility function. Differentiating income tax rates by gender would 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 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: gender differences in bunching responses at the tax credit cutoff. Section 5 shows that this result is particularly concentrated in contexts with more traditional gender norms. In section 6, we discuss the implications of our findings for second earners’ career and gender income inequality. Finally, section 7 concludes.

---

5 Assuming a collective model of the household with a marriage market, Gayle and Shephard (2019) find that the optimal tax system for couples is characterized by negative jointness. Although moderate, the welfare gains from negative jointness are increasing in the gender wage gap, providing an instrument to address within household inequality.

6 See also Boskin and Sheshinski (1983), Apps and Rees (1988), and Apps and Rees (1999) for papers in this tradition.
2 Institutional Background and Data

2.1 Gender Norms in Italy and their Socio-Economic Implications

Similarly to other Southern-European countries, Italy is a conservative society with traditional gender norms. The vast majority of families embraces the male breadwinner model, where wives are mostly out of the labor force or enter in the labor market as second earners. According to the OECD Family Database, Italy ranks in 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. In terms of gender pay differences, Italy looks relatively better: the gender gap in median earnings of full-time employees was around 5 percent in 2018 (OECD average is around 13 percent). Casarico and Lattanzio (2019) show that the gender pay gap declined steadily over the last two decades in Italy.

Panel A of Figure 1 shows trends in the share of couples where the wife is the main earner, divided by macro-region. We find that, on average, women are the head of the household in less than 10 percent of couples. The figure also shows that there are wide geographical differences: in 2020, women lead in around 15 percent of families living in the Northern Italy, while this happened in only 7 percent of couples from Southern Italy. The share of couples where the woman is the main earner has been steadily growing: it increased of around 5 percentage points (from 6 to 11 percent) over the 2013-2020 period. Panel b in the figure shows that such geographical differences are strongly correlated with measures of gender norms, such as the share of respondents agreeing that “men should have more right to a job than women” from the 2017 European Values Study.

The 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.” The survey also reveals other interesting facts about the Italian society. For example, around one-third of working women agrees or strongly agrees that “husbands are the main responsible for the provision of family needs”. This figure increased to around 56 percent of non-working female respon-

---

7 Since women’s labor force participation and education levels have increased in most countries, a similar pattern emerges in other developed countries. For example, among families in which both members received earnings, the share of families where the wife outearns the husband has increased from 15.9 to 29.3 percent in the U.S. between 1981 and 2015 (Blau and Kahn 2017). Furthermore, it seems that in the vast majority of couples where the wife is the main earner, this disparity tends to be persistent (Winkler et al. 2005).

8 Alesina and Ichino (2009) discuss the implications of unpaid family work for labor supply decision of Italian women.
Figure 1: Gender Norms and Wife Head of Household

(a) Trends in Woman Head of Household

(b) Gender Norms vs Women Leading

Notes: The left hand-side graph displays trends in the share of couples where a woman is the head of household, as reported in a survey called Aspects of Daily Life (Indagine Multiscopo sulle Famiglie: Aspetti della Vita Quotidiana), provided by the Italian Institute of Statistics (ISTAT). We report separate series from couples living in the North of Italy (red squares), Center Italy (blue triangles), and South of Italy (green diamonds). The black solid line depicts the (unweighted) average. The right hand-side graph compares the region-specific share of couples where the woman leads the couples with the share of respondents agreeing that “men should have more right to a job than women”, using data from the 2017 European Values Study. Each point corresponds to the region-level average. The figure also reports the slope and the corresponding standard error. On average, a 1 percentage point increase in the share of respondents agreeing with the above statement leads to a 0.152 percentage points decrease in the probability that women lead the household.

dents. 30 percent of female respondents also agrees or strongly agrees that ”overall, men’s conditions in the current Italian society are largely better than women’s conditions.” These figures portray Italy as a gender-conservative environment, with considerable gender inequalities in many social and economic aspects. It thus provides a useful setting to study whether gender identity norms affect economic outcomes.

What happens when such conservative gender norms are violated? In Appendix B, we offer survey evidence to shed light on this question. Using data on married couples from 2013 to 2020 survey on Aspects of Daily Life (Indagine Multiscopo sulle Famiglie: Aspetti della Vita Quotidiana), we collect information on who is the head of the household. The survey asks questions on a range of topics, including economic outcomes, health status and life satisfaction, allowing us to explore the effect of gender norms violation on several socio-economic outcomes. We also retrieve basic demographics (region of residence, age, marital status, marriage tenure, education level) and labor market information (occupation, sector).

We show that the violation of the male breadwinner model has important negative implications on couples’ life and economic satisfaction. Specifically, we present three main suggestive results. First, in couples where the wife’s income exceeds the husband’s, both the wife and the husband report to be less satisfied with their mar-
riage. Relative to comparable couples where the husband is the main earner, couples where the wife outearns the husband are between 1 and 1.6 percentage points less likely to respond to feel “happy” or “very happy” with their marriage. This evidence is similar to previous (more robust) results based on American families (Bertrand et al. 2015). We also show that this effect spreads through the family: when the male breadwinner model fails, daughters are 2.1 percentage points less likely to feel “satisfied” or “very satisfied” about their family, while sons appear to be less responsive.

Second, husbands are more likely to report anxiety disorders when the male breadwinner model fails. In couples where the wife is the main earner, husbands are 1.7 percentage points more likely to report to suffer of anxiety disorders. This result is consistent with social psychology research stating that infringing of internalized rules generates apprehension and nervousness.9

Finally, despite our empirical exercise is based on comparing couples with similar observable economic conditions, we find that wives report to be significantly less satisfied with the economic conditions of their family when they are the main earner in the couple. One potential explanation for this finding is that second earner husbands do not adopt behaviors that maximize family income. We study this potential explanation in the context of the Italian spouse tax credit, that we describe below.

2.2 Spouse Tax Credit and Income Taxation in Italy

All Italian residents are subject to personal income taxation (IRPEF, Imposta sul Reddito delle PErsona 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 over the period we study.

The final tax burden is calculated net of tax credits. The Italian personal income tax system provides a wide array 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). The main earner in a couple receives a tax credit if the second earner reports gross annual income below 2,840.51 euros.10 Entitlement is also allowed for second earners that are out of the labor force. Spouse tax credit eligibility is

9 In personality development, researchers agree on the importance of internalization of rules for behavior. For instance, Thomas (1996) points out that identity, or self, must be constantly “defended against anxiety in order to limit disruption and maintain a sense of unity” (p. 284).

10 During 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 post-war 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).
self-reported by the main earner when filling tax returns. In the case of third-party reported earnings, workers can apply for receiving the tax credit directly from the Italian Social Security Institute (INPS) website.

Table 1 illustrates the main features of the spouse tax credit. The table reports the size of the tax credit (column 2), how much of the main earner’s final tax burden is reduced thanks to spouse the tax credit (column 3), and the importance of the tax credit in proportion to main earner’s gross income (column 4). The tax credit is computed by a simple formula for main earners reporting income below 15,000 euros or between 40,000 and 80,000 euros, while it is a fixed amount for other income groups. 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. 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 income lower than 10,000 euros would get a tax credit of 726.7 euros, that 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, which corresponds to around 10 percent of their gross income. The relevance of the tax credit then monotonically decreased over the main earner’s income distribution, accounting for less than one-tenth of the tax burden for incomes above 29,000 euros.\(^\text{11}\)

### 2.3 Data and Descriptive Evidence

We use administrative data provided by 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 (ISTAT) and the fifth most populous (ISTAT). As income taxes in Italy are filled individually, the unit of observation is the individual. The dataset contains taxable income data (divided in 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 over the 2007-2014 period.

Following standard practice in the literature, our bunching analysis on second earners focuses on self-employees.\(^\text{12}\) Since income is self-reported, self-employees

\(^{11}\)The Italian spouse tax credit is similar in spirit to the U.S. Earnings Income Tax Credit (EITC): a large cash transfer for low-income earners that introduced substantial kinks in the budget constraints of families (see, e.g., Saez (2010) for evidence of behavioral responses to the EITC along the intensive margin; Kleven et al. (2019a) for responses over the extensive (employment) margin). Yet, while the spouse tax credit eligibility is based on second earner’s income, the EITC only grants working individuals based on family income and number of children.

\(^{12}\)For instance, Saez (2010) finds that bunching responses on kink points of the U.S. Earnings Income Tax Credit (EITC) are exclusively concentrated among EITC recipients with self-employment income. EITC recipients with only wage earnings display no evidence of bunching. Kleven et al.
Table 1: The Spouse Tax Credit

<table>
<thead>
<tr>
<th>Main earner’s gross income (euros)</th>
<th>Tax credit (euros)</th>
<th>Tax credit (% of tax burden)</th>
<th>Tax credit (% of gross income)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-15,000</td>
<td>800</td>
<td>43.2%</td>
<td>9.9%</td>
</tr>
<tr>
<td>15,001-29,000</td>
<td>690</td>
<td>12.9%</td>
<td>3.1%</td>
</tr>
<tr>
<td>29,001-29,200</td>
<td>700</td>
<td>9.5%</td>
<td>2.4%</td>
</tr>
<tr>
<td>29,201-34,700</td>
<td>710</td>
<td>8.4%</td>
<td>2.2%</td>
</tr>
<tr>
<td>34,701-35,000</td>
<td>720</td>
<td>7.5%</td>
<td>2.1%</td>
</tr>
<tr>
<td>35,001-35,100</td>
<td>710</td>
<td>7.4%</td>
<td>2.0%</td>
</tr>
<tr>
<td>35,101-35,200</td>
<td>700</td>
<td>7.2%</td>
<td>1.9%</td>
</tr>
<tr>
<td>35,201-40,000</td>
<td>690</td>
<td>6.5%</td>
<td>1.8%</td>
</tr>
<tr>
<td>40,001-80,000</td>
<td>690*[(80,000-gross income)/40,000]</td>
<td>1.8%</td>
<td>0.6%</td>
</tr>
<tr>
<td>80,001-</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

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 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 main earner’s gross income, computed at the median income level of the corresponding main earner’s gross income bracket.

can adjust their taxable income for tax-related reasons. By contrast, since employees’ wages are third-party reported, employees have limited room for adjusting their income and thus respond to the tax credit cutoff.

Table A2 reports the summary statistics on our sample of self-employees, which includes around 2.7 millions taxpayers. Panel A of Table A2 focuses on male taxpayers, who represent the 69.7% of the sample: the average gross declared income is 34,695 Euros and the 18.7% of them receive the spouse tax credit. Panel B of Table A2 shows the summary statistics for the sample of female taxpayers: the average gross declared income is 25,049 Euros and the fraction of spouse tax credit’s recipients is 2%.

The main advantage of using tax returns data is that they provide information about the exact location of taxpayers over the income distribution. Moreover, in contrast to survey data, tax returns data have almost no measurement error. These administrative data are thus suitable for our empirical analysis, which consists in estimating whether the spouse tax credit affects income reporting behavior.

To offer prima facie evidence on the impact of the spouse tax credit on second earners’ reported income, Figure 2 depicts the fraction of married men (left-hand side graph) and married women (right-hand side graph) located below the cutoff

(2011) find that there is virtually no tax-related manipulation in wage earnings of audited Danish taxpayers because of third-party reporting by firms.
determining spouse tax credit eligibility (i.e., taxpayers reporting income between 840.51 and 2,840.51 euros) with respect to the fraction of taxpayers reporting income above the cutoff (i.e., between 2,840.52 and 4,840.51 Euros). To ensure comparability, the two graphs share a common scale, with darker (lighter) areas denote municipalities where the portion of taxpayers located below the cutoff is relatively larger (smaller).

Figure 2: Map of Taxpayers’ Fraction Located Below the Spouse Tax Credit Cutoff

(a) Men – Married

(b) Women – Married

Notes: These figures show the distribution of taxpayers located below the spouse tax credit cutoff across municipalities in the Veneto region. Left panel focuses on married men; the right panel on married women. For each municipality, we report the difference between the fraction of taxpayers declaring an income below the spouse tax credit threshold (between 840.5 and 2,840.5 Euros) and the fraction of taxpayers declaring an income above it (between 2,841.5 and 4,840.5 Euros). To ensure comparability, the two graphs share a common scale, with darker (lighter) areas denote municipalities where the portion of taxpayers located below the cutoff is relatively larger (smaller).

Two main remarks emerge from this figure. First, there are gender differences in the probability of reporting income below the spouse tax credit cutoff: women are much more likely to report incomes below the spouse tax credit than men. Second, the figure shows substantial geographical dispersion in the share of taxpayers located below the spouse tax credit cutoff. For instance, gender differences in the fraction of taxpayers located just below the cutoff are negligible in municipalities located around the Adriatic sea (South-East in the map). This area, mostly composed of municipalities located in the Venice province, includes municipalities with higher female labor force participation, where gender attitudes are likely to be less pervasive. By contrast, gender differences are more intense in rural areas, such as the
mountainous province of Belluno (North-East in the map), where female employment is relatively lower.

2.4 Take-Up Rate
This section provides information on the spouse tax credit take-up rate. Figure 3 depicts the spouse tax credit’s take-up rate (as a share of total married taxpayers), separately for married women (red diamonds) and married men (black circles). We present the take up rate in 20 bins of tax credit amount as a share of total tax burden (see appendix Figure A1 for an illustration of the take-up rate by gross income). The pattern emerging from this figure leads to three main observations.

Figure 3: Take-Up Rate of Spouse Tax Credit

![Figure 3: Take-Up Rate of Spouse Tax Credit](image)

Notes: This binned scatter 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). The sample includes all married taxpayers with income above 5,000 euros.

First, there is 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.

Second, men’s take-up rate presents an inverted U-shaped pattern over the tax credit distribution. Intuitively, when the tax credit matters less (as a share of the
main earner’s tax burden), the take-up rate is low because the economic returns are relatively smaller. The take-up rate then linearly increases up to the point where the tax credit accounts for around one-fifth of the main earner’s tax bill. However, when the tax credit is high enough, the take-up rate starts to gradually decline. One explanation for this pattern is that the second earner’s market income becomes an importance source of total family income at main earner’s low income levels, when the tax credit matters relatively more. Therefore, second earners’ wives become more likely to contribute to family income by reporting income well above the tax credit cutoff.

Finally, the women’s take-up rate distribution is fairly flat. This result could suggest that second earners’ husbands decision to bunch at the cutoff determining spouse tax credit eligibility does not respond to economic incentives faced by their wife. This figure thus provides prima facie evidence that the gender of the second earner matters: couples where the male breadwinner model is satisfied are more likely to maximize family income by responding on the economic incentives created by the spouse tax credit.

3 Conceptual Framework and Empirical Strategy

3.1 Theoretical Foundations

The spouse tax credit creates a substantial “notch” in the budget constraint of families, that is a discontinuity in the choice set of second earner gross income versus family (net) income. Under the conventional (collective) model of household behavior (see, e.g., Chiappori 1988; Chiappori 1992; Apps and Rees 1996; Blundell et al. 2007), this notch should induce second earners, who would otherwise report more income, to instead bunch right at the tax credit cutoff.

Figure 4 offers a simple illustration of how individuals would respond to the tax notch created by the spouse tax credit. Before couple formation, individuals choose gross income, \( y \), that maximizes their own utility subject to their budget constraint. Gross income is distributed according to a smooth density distribution \( h(y) \) and any heterogeneity in income is due to preferences or idiosyncratic shocks. When individuals marry and an household is created, the second earner in the couple will face a tax notch at income level \( y^* \). Such a notch introduces an incentive for reporting income just at the point where the tax credit cutoff is met. Therefore, in a frictionless world, the tax credit should create a hole in the second earners’ gross income distribution falling into a segment \([y^*, y^* + \Delta y^*]\), and excess bunching just at the tax credit cutoff.

In this simple model, the gender of the second earner would not matter: couples
cooperate to maximize family income and household decisions are Pareto-efficient. In this paper, we argue that gender identity norms could be an important factor in driving economic decisions. 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. Hence, people’s economic actions can in part be explained by a desire to conform with one’s sense of self. 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.

A large 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). For instance, in family specialization models à la Becker (1991), married men mostly focus on working, while their wives are responsible for non-market production. Norms prescribing that “men work in the labor force and women work in the home” could explain why women have been slow at increasing their labor force participation. In couples that embrace the male breadwinner model, women would be willing to reduce their labor supply and underinvest in their career, while men would adopt behaviors that allow them to “fill the gap” when the male breadwinner model fails.

The spouse tax credit offers a propitious testing ground for evaluating the importance of gender identity norms. A natural test for evaluating the importance of gender norms would be to observe different responses to the tax credit depending on gender of the second earner. If gender identity norms affect behavior, the outcomes that we should observe are those reported in panel B. The density distribution of second earner wives will present excess bunching at the tax credit cutoff (distribution in blue), while the tax notch will not generate any change in the distribution of second earner husbands (distribution in red).

As a benchmark, we have considered a static model. Importantly, however, if career concerns are important and the spouse tax credit affects not only current reported income, but also income reported in the future, the corresponding bunching would be smaller. We will come back to this point when interpreting our empirical results.
Notes: This figure illustrates the incentives that would lead second earners within a couple to bunch at the spouse tax credit cutoff. The tax credit creates a notch in the budget constraint of families: a discontinuity in the choice set of second earner gross income versus family (net) income. When individuals marry and an household is created, the second earner in the couple will face a tax notch at the gross income level $y^*$. Such a notch introduces an incentive for reporting income just at the point where the tax credit cutoff is met. Therefore, in a frictionless world, the tax credit should create a hole in the second earners’ gross income distribution falling into a segment $[y^*, y^* + \Delta y^*]$, and excess bunching just at the tax credit cutoff. The segment $[y^*, y^* + \Delta y^*]$ thus corresponds to the dominated area. At the tax notch, second earners will enjoy larger family net income than in the dominated area. If, instead, gender norms prevent second earners’ men to adjust their gross income, then the tax notch will not generate any change in the distribution of second earners’ husbands gross income distribution. Panel B displays the two possible density distributions of second earners’ gross income.

3.2 Bunching Approach

We study whether the second earner’s gender affect family income maximization by examining bunching at the spouse tax credit cutoff. Following previous studies (Saez 2010; Chetty et al. 2011; Kleven and Waseem 2013), we first group 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:

$$n_j = \sum_{i=0}^{p} \beta_i \cdot (m_j)^i + \sum_{i=L}^{U} \gamma_i \cdot 1(m_j = i) + u_j,$$  \hspace{1cm} (1)

where the first term on the right-hand side is a $p$-th degree polynomial; 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 ($p = 7$).
7).

We can then calculate counterfactual bin counts as the predicted values from equation (1) omitting the contribution of dummies in the excluded range:

\[ \hat{n}_j = \sum_{i=0}^{p} \beta_i \cdot (m_j)^i. \] (2)

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

\[ \hat{B} = \sum_{j=L}^{0} (n_j - \hat{n}_j) = \sum_{j=L}^{0} \hat{\gamma}_j. \] (3)

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. A larger \( \hat{B} \) estimate would imply a greater distortion in reported income by second earners due to the tax credit.

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.\(^{14}\)

4 Gender Differences in Bunching Responses

4.1 Bunching Responses to the Spouse Tax Credit

We present our main empirical results in Figure 5, which plots the empirical distribution of gross income by gender. The top panel presents histograms of gross income distribution. To construct these histograms, 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’

\(^{13}\)To determine the excluded range, we follow the procedure proposed by Chetty et al. (2011). 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.

\(^{14}\)Since we observe the actual income distribution, the estimated standard error mostly reflects misspecification of the polynomial used to estimate the counterfactual distribution, rather than sampling errors.
distribution up to an income level of 25,000 euros.\textsuperscript{15}

The figure shows that there is a spike in the fraction of female taxpayers just below the tax credit cutoff (right panels). By contrast, the distribution of male income is smooth and do not present any visible spike at the tax credit cutoff (left panels). Although the shape of the income distribution differs by gender (due, e.g., to gender income gaps), we do not detect any visible spikes at other points of the gross income distribution.

Figure 5: Bunching Responses to the Spouse Tax Credit

(a) Men – Married

(b) Women – Married

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; right-hand side graphs on married female taxpayers. 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.

The bottom graphs offer a comparison of the observed distributions (blue dots) with the counterfactual distribution (red solid line). We also report excess bunching

\textsuperscript{15}In these graphs, we do not make any restriction on our sample of self-employees taxpayers. In Appendix Figure A2, 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 that is well above the tax credit cutoff, thus not affecting the counterfactual distribution.
estimates, obtained from equation (3), and bootstrapped standard errors. The figure provides clear evidence of gender differences in bunching response to the spouse tax credit. Relative to the counterfactual distribution, there is a clear excess mass of female earners reporting income just below the cutoff, while we do not observe any excess mass in the distribution of male taxpayers. We estimate excess bunching of 1.238 (-0.008) times the height of the counterfactual distribution of women (men). The standard error associated with our excess bunching estimate is 0.239 for women, 0.363 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 5.03), while it is not rejected for the male distribution (t-statistics of -0.02).

These estimates suggest that the density of second earner women located in an income range strictly below the tax credit cutoff is 1.238 times larger than the density that we would have observed in the absence of the policy. On average, our bunching estimate suggests that the female marginal buncher reduces taxable income by 1.238 bins, which corresponds to around 186 euros. This estimate suggests that the female marginal buncher reduces her reported income by around 186 euros to let her husband to enjoy a tax credit larger than 700 euros, thus increasing family net income by around 514 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 514 euros less than comparable couples where the wife is the second earner.

Second earners can respond along two main margins: changes in labor supply or evasion and avoidance responses. Our result holds regardless of what margins underlies changes in second earners’ reported gross income. Intuitively, whether the second earner 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. Therefore, distinguishing evasion responses, including within-couple income shifting, from labor supply responses is not critical for the conclusions we draw here, as there are no a-priori reasons to believe that the margin of response could systematically differ depending on the second earner’s gender.¹⁶

One interpretation of our result is that couples are more likely to cooperate to maximize family income when the gender identity norm that men should be the main earner is satisfied. Our results provide clear evidence that a significant portion of second earner women maximizes their family income by keeping their reported income just below the tax credit cutoff, while gender norms could have prevented men to adjust their reported income. We will now investigate whether such gender

¹⁶Since 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. However, our tax returns data cannot be linked with other dataset, such as matched employer-employee data, providing labor supply information. We are thus unable to distinguish whether reported income responses reflect evasion or labor supply effects in our data.
4.2 Bunching Responses Emerge After Marriage

If marriage “activates” the gender identity norm, gender differences in bunching at the tax credit cutoff should emerge right after marriage. Married women will start to spend more time in non-market activities, which could reduce the effort that they put into their market jobs, and thus be more likely to report incomes just below the tax credit cutoff.

To test this hypothesis, we implement a difference-in-differences (DiD) design comparing bunching behavior between female and male taxpayers, before and after the marriage year. We define a dichotomous variable, $Bunch_{i,t}$, that is equal to 1 in the first year $t$ when individual $i$ reports gross income below the spouse tax credit cutoff. This approach has two main advantages compared to our “static” bunching approach. First, we can account for individual-level unobserved heterogeneity when estimating how the spouse tax credit affects income reporting behavior. Second, we can investigate the dynamics of “bunching” behavior around marriage.\(^{17}\)

We run specifications like the following:

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

where $Bunch_{i,t}$ is a dummy variable taking value one in the first “bunching” year, that is when the income drops below the spouse tax credit for individual $i$ in year $t$. $F_i$ 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 DiD coefficient $\beta_k$ can be interpreted as the probability of reporting income below the spouse tax credit cutoff 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, $\delta_{i,m(i)}$, allows us to construct potentially more realistic counterfactuals by comparing gender differences in the outcome variable within a given municipality. Individual fixed effects, $\gamma_i$, accounts for any time-invariant individual-specific characteristics or unobserved factors. Finally, $u_{i,t}$ is an error term. We cluster the standard errors at the individual level.

Figure 6 plots the $\beta_k$ coefficient estimates and both 90 and 95 percent confidence intervals. The figure provides two main findings. First, there is no significant differ-

---

\(^{17}\)Since we do not observe marriage status in year 2007, this analysis is based on tax returns covering the 2008-2014 period (instead of the baseline 2007-2014 period).
ence by gender in the years leading to marriage: the fraction of women reporting less than 2,840.51 euros is relatively similar to the fraction of men. This result suggests that couples are not formed in a way that would predict eligibility for the spouse tax credit. It is consistent with some studies showing that impact of taxation on marital status is modest (Alm and Whittington 1995; Alm and Whittington 1997; Alm and Whittington 1999).

Figure 6: Bunching Responses Around Marriage Event

Notes: This figure reports the outcome of the difference-in-differences analysis (model 4) that focuses on the impact of marriage on bunching behavior. The figure reports 95% (delimited by horizontal bars) and 90% (bold line) confidence intervals. The specification always includes individual fixed effects and municipality-year fixed effects. The coefficient for the corresponding difference-in-differences analysis is 0.013 (SE=0.003).

Second, the figure shows that gender differences emerge at marriage. On average, we estimate that, once married, the fraction of women reporting income below the tax credit eligibility cutoff is around 1.3 percentage points larger than the fraction of men. This result implies that marriage activated the gender identity norm, leading wives to be significantly more likely than their husbands to report income below the spouse tax credit cutoff. This result is in line with specialization models à la Becker (1991), suggesting that married men mostly focus on working, while their wives are responsible for non-market production.
4.3 Bunching Responses Exacerbate With Childbirth

We then examine whether childbirth also contributed to generate gender differences in bunching behavior. Considerable empirical evidence indicates the existence of a child penalty for women (see, e.g., Kleven et al. (2019a) for cross-country evidence; Casarico and Lattanzio (2021) for the Italian context). While there are a number of reasons that might explain the negative relationship between childbirth and women’s labor market outcomes, some studies have shown that gender norms become more intense around childbirth (Kuziemko et al. 2018).

We examine the impact of childbirth by running specifications as in equation (4), but where the event time is defined by the number of years since childbirth. There are two main empirical challenges that we need to tackle. Our first challenge is the definition of childbirth event, that we do not directly observe in our data. We impute the childbirth year as the first year where a taxpayer receives a child tax credit. Since a child tax credit is received by both parents starting from childbirth, it turns out to be a reasonable proxy for defining the year where the first child was born.

Second, there could be other contemporaneous events taking place at childbirth year. Since, as previously shown, gender differences in bunching began to materialize at marriage, we are concerned that childbirth event times may somehow overlap with marriage event times, biasing upward our childbirth estimates. Although potentially important, we find that marriage timing perfectly overlaps with childbirth timing just for a negligible portion of our sample. This suggests that this issue should be not particularly meaningful. Yet, we augment our baseline model with marriage tenure-year fixed effects to account for the marriage-driven effects that we documented above.

Figure 7 presents our main results on whether childbirth affects bunching behavior. The figure reports the $\beta_k$ coefficient estimates and both 90 and 95 percent confidence intervals obtained from regression as in (4), where the event time is the number of years elapsed from childbirth. We find that gender differences in bunching rate emerged one year after childbirth. On average, we find that women are 0.6 percentage points more likely to bunch than men after childbirth.

Taken together, these results suggest that gender differences in bunching behavior emerged around two of the two most, if not the most, important events in couples’ life: marriage and childbirth. One explanation for this result is that gender norms turn on around these events. We will now first test the robustness of these results. Next, we will provide further evidence on the importance of gender norms for explaining our findings.
Figure 7: Bunching Responses Around Childbirth Event

Notes: This figure reports the outcome of the difference-in-differences analysis (model 4) that focuses on the impact of a childbirth event on bunching behavior. The figure reports 95% (delimited by horizontal bars) and 90% (bold line) confidence intervals. The specification always includes individual fixed effects and municipality-year fixed effects. The coefficient for the corresponding difference-in-differences analysis is 0.006 (SE=0.002).

4.4 Robustness Checks

In principle, our results could be masking gender differences in scope for bunching at the spouse tax credit cutoff. For instance, women might sort into occupations that allow more flexibility in adjusting their reported taxable income (either for labor supply or tax evasion reasons). If there are gender differences in scope for responding to economic incentives, then women would be more responsive to any budget set discontinuity.

To investigate whether this hypothesis may be likely, we study whether gender differences in bunching behavior emerge at other tax notches or kinks where a discontinuity is present in own gross versus net income. We focus on two cases: i. a tax notch created by the tax exemption cutoff: self-employed income below 5,000 euros is tax exempted; ii. the marginal tax rate increase - from 15 to 23 percent - at the 15,000 income bracket. If unobservable gender differences in scope for adjusting reported incomes are driving our results, these unobservable characteristics would lead us to
observe larger bunching by women at these two points of the income distribution as well. Figure A3 and Figure A4 provide evidence of a similar bunching response by gender at these two points of the income distribution. This result assuage concerns regarding the possibility of gender differences in scope for bunching.18

The identification assumption underlying causal inference on gender differences in responses to the spouse tax credit is that the gender-specific income distribution would be smooth in the absence of the tax credit cutoff. We can relax this assumption by examining the distribution of wage earners. Since their earnings are reported by employers, wage earners face substantial frictions in adjusting their reported income. Appendix Figure A5 displays the income distribution for all wage earners, while Figure A6 provides distributions by gender. We do not detect any significant excess mass in any of these distributions. This yields credence to our identifying assumption.

We then test the sensitivity of our bunching estimates to some assumptions that we make to estimate the counterfactual income distribution. First, we test whether our estimates are sensitive to different polynomial orders used to estimate the counterfactual distribution. Appendix Table A3, we show that our baseline bunching estimates are not substantially affected by different polynomial orders. Second, we scrutinize whether our bunching estimate are sensitive to the income bandwidth being analyzed. Appendix Figure A7 and Figure A8 show that our male-specific and female-specific, respectively, bunching estimates hold regardless of how we define the income window of interest.

These analyses do not completely rule out the possibility that an unobservable difference between women and men is driving our results. For instance, it may be that self-employees low-income men face substantial frictions when adjusting their taxable income, while women do not. Under this explanation, our finding would result from characteristics of second earner women at that specific point of the income distribution. Our next step is to compare second earner men and women at the same income level, but in contexts with different gender identity norms.

5 Are Gender Norms Responsible for Bunching?

In this section, we bring additional evidence to bear on the hypothesis that gender identity norms are responsible for our results. We present two additional exercises. First, we focus on immigrants to test whether gender differences in bunching rate are

---

18 As long as bunching mostly reflects tax evasion, an alternative explanation is that women are more risk averse than men, and thus more willing to incur in elusive behaviors. Yet, most of the existing evidence points to the opposite direction. In a review of the existing experimental evidence, Croson and Gneezy (2009) report that women are more risk averse than men, on average.
relatively larger for individuals coming from more traditional societies. Second, we correlate bunching rates across municipalities and individuals with different proxies for traditional gender norms.

5.1 Immigrants

We start by proposing an epidemiological study of gender norms using foreign-born immigrants. Following Fernández and Fogli (2009), we exploit the “portability” of cultural factors: when individuals emigrate, they may take some aspects of their culture, including gender norms, with them. This suggests that studying immigrants may be a useful strategy for isolating the role of gender norms from other economic and institutional factors. These individuals, living and working in Italy, face the same markets and institutions, but they potentially differ in their cultural heritage, as reflected in their country of origin. Following Fernández and Fogli (2009), we proxy gender norms with past female labor force participation from the woman’s country of origin. We retrieve information on female labor force participation from the World Bank database.¹⁹

Since these women come from different societies, but live and work in the same economic and formal institutional environment, our standard bunching approach would not allow us to investigate how gender norms operate in isolation from other competing factors, such economic factors and institutions of the destination municipality. For instance, it is plausible that immigrants self-select into municipalities that present characteristics (either cultural or institutional) that are closer to those of their origin country. To account for this issue, we estimate bunching responses adjusting for selection in their municipality of residence. Therefore, our empirical approach compares bunching rates across individuals living in the same municipality, but with different gender norms based on their origin country. Our strategy will thus allow us to study how gender norms operate in isolation from other factors that vary across municipalities. To gain precision, we also include year fixed effects and a set of individual controls (age and marital status).

Figure 8 presents our results. The figure relates the proportion of male (left-hand side graph) and female taxpayers (right-hand side graph) reporting income below the spouse tax credit cutoff (vertical axis) with origin country’s female labor force participation (horizontal axis). It provides striking graphical evidence that gender differences in bunching rates are strongly related to female labor force participation. Conditional on their municipality of residence, women born in high-female employment countries, such as China or Norway, are much less likely to report incomes

---

¹⁹Following the classification proposed by the International Labour Organization, female labor force participation is computed relative to the share of female population older than 15. We use the female employment share observed in 2000.
just below the spouse tax credit cutoff, compared to women born in low-female employment countries, such as Iraq or Saudi Arabia. This relationship is statistically significant and economically meaningful: a 10 percentage point increase in origin country’s female employment raises the share of women reporting income just below the spouse tax credit cutoff by 0.32 percentage points (the slope coefficient is 0.032, with standard errors of 0.009). By contrast, there is no significant relationship for male taxpayers: the figure presents a flat relationship and the estimated slope is not significantly different from zero.

Figure 8: Gender Differences in Bunching Rate Among Immigrants

(a) Men

(b) Women

Notes: The figure relates the proportion of male (left-hand side graph) and female taxpayers (right-hand side graph) reporting income below the spouse tax credit cutoff (vertical axis) with origin country’s female labor force participation (horizontal axis). The proportion of taxpayers below the cutoff is defined by the share of (either male or female) taxpayers declaring income between 2,500 and 2,840.50 Euros. The scatter-plot controls for municipality fixed effects, year fixed effects, and individual specific controls (age and marital status). The sample includes taxpayers who declared less than 10,000 Euros in the first year of the dataset. The coefficient of the regression is -0.0029 (SE=0.005) for male and -0.032 (SE=0.009) for female.

We believe this result is important for two main reasons. First, this result again suggests that gender norms are an important determinant of the bunching response to the spouse tax credit. This evidence appears robust to differential selection of immigrants in their municipality of residence. Second, we show that differences in the intensity of traditional gender norms are able to explain not only gender differences in bunching rates, but also across women that were grew in contexts that are significantly different in terms of gender norms (as proxied by female employment). The latter result motivates our next empirical exercise: we relate variation in proxies for gender norms across individuals and municipalities with the intensity of bunching responses.
5.2 Correlate Bunching Rates with Proxies for Gender Norms

This section relates bunching rates with several proxies for gender norms, leveraging both cross-individual and geographical variation in the intensity of traditional gender norms.

We start by exploiting cross-cohort variation in progressivity of gender views and gender differences in bunching rates. We use data from 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, to construct an index of gender roles that varies across cohorts. Specifically, we measure the share of women that strongly agrees with two statements that would likely capture views on traditional gender: i. “men are the main responsible for the provision of family needs”; ii. “men should have more right to a job than women”. The left-hand side panel of Figure 9 shows that views about gender roles have become more progressive among younger cohorts. The share of women with very conservative gender norms has decreased by at least 20 percentage points over the last 50 years. Namely, while around one-third of women born in the early post-war period was likely to agree that “men should have more right to a job than women”, just one-tenth of women born in the late 1990s and early 2000s shares this view.

We then investigate whether this pattern mirrors gender differences in bunching rate. We present gender differences by age in the right-hand side graph, where we depict the gender difference in the proportion of taxpayers located barely below the spouse tax credit cutoff. To account for cross-municipality heterogeneities, we compute bunching rates controlling for municipality fixed effects. The figure shows that our time series of gender progressivity views mirrors the size of bunching differences by gender, suggesting that gender norms are an important determinant of the size of bunching rate even across cohorts.20

Next, we leverage cross-municipality variation in the intensity of gender norms. We propose three main proxies, and we then split municipalities according to whether they score below or above the median value in each of these proxies. Although these proxies are likely to be confounded with other potential cross-municipality heterogeneity, we believe that this result might offer some suggestive evidence on the link between gender norms and gender differences in bunching behavior.

First, we construct a municipality-level index of support for the deregulation of abortion in a 1981 referendum.21 We assume that gender norms are less progressive

---

20 In Table 2, we report unconditional bunching estimates and standard errors estimated for taxpayers above versus below the median age of taxpayers in our sample. Consistent with the evidence presented in the figure, we find that the bunching estimate is statistically significant and large among older women, while it is not statistically significant among younger women and for men (of any age).

21 This referendum, that took place on the 17th of May 1981, asked Italians their opinion on the so-
Figure 9: Gender Differences in Bunching Are Larger Among the Elderly

(a) Intensity of Traditional Gender Norms
(b) Gender Bunching Differences By Age

Notes: The left-hand side graph 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).

The right-hand scatter-plot shows the regressions of taxpayers age on taxpayers bunching behaviour (defined as declaring between 2,500 and 2,840.50 Euros), in difference between female and male taxpayers. The specification includes municipality and year fixed effects and individual specific controls (civil status). The sample includes taxpayers who declared less than 10,000 Euros in the first year of the dataset.

in municipalities that were less willing to opt for a “yes” in the referendum. Second, we look at the share of female politicians that are elected in the town council 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. Finally, consistent with patterns emerging from survey data, we focus on urban density: we expect that individuals living in rural places share more traditional gender norms.

Table 2 shows bunching estimates and standard errors for married women and married men in each of the sample described above. The table provides evidence that bunching estimates among female taxpayers are systematically larger in municipalities with more traditional gender views. In line with our results based on immigrants, we find that differences in the intensity of traditional gender norms seem to have a limited impact on the probability of bunching among men. These results are thus consistent with the notion that gender norms are a powerful factor for explaining gender differences in bunching rate also across places and individuals.

called law 194 (see “Norme per la tutela sociale della maternità e sull’interruzione volontaria della gravidanza”), passed three years before.
Table 2: Bunching Responses Are Larger in Municipalities with More Traditional Gender Norms

<table>
<thead>
<tr>
<th></th>
<th>Bunching estimate for women</th>
<th></th>
<th>Bunching estimate for men</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Below median</td>
<td>Above median</td>
<td>Below median</td>
<td>Above median</td>
</tr>
<tr>
<td>Support 1981 referendum on free abortion</td>
<td>1.287***</td>
<td>0.101</td>
<td>0.014</td>
<td>-0.314</td>
</tr>
<tr>
<td></td>
<td>(0.296)</td>
<td>(0.254)</td>
<td>(0.401)</td>
<td>(0.309)</td>
</tr>
<tr>
<td>% of female politicians in town council</td>
<td>1.103***</td>
<td>0.613***</td>
<td>0.149</td>
<td>0.366</td>
</tr>
<tr>
<td></td>
<td>(0.265)</td>
<td>(0.268)</td>
<td>(0.241)</td>
<td>(0.255)</td>
</tr>
<tr>
<td>Urban density</td>
<td>1.382***</td>
<td>0.316</td>
<td>0.273</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(0.250)</td>
<td>(0.232)</td>
<td>(0.273)</td>
<td>(0.260)</td>
</tr>
<tr>
<td>Age</td>
<td>0.145</td>
<td>1.129***</td>
<td>-0.128</td>
<td>-0.184</td>
</tr>
<tr>
<td></td>
<td>(0.261)</td>
<td>(0.307)</td>
<td>(0.364)</td>
<td>(0.366)</td>
</tr>
</tbody>
</table>

Note: This table reports bunching estimates and bootstrapped standard errors, computed as described in Section 3.2, for married female and married male taxpayers. We split our original sample according to whether a municipality ranks below or above the median value of the following variables: i. share of votes that support the abrogation of free abortion in a 1981 referendum; ii. share of female politicians elected in town council; iii. urban density index. Finally, in the last column, we split taxpayers according to their age.

6 Second Earners’ Career and Gender Gaps

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. In this section, we first study whether the spouse tax credit generates long-lasting effects on secondary earners. Then, we discuss the implications of our findings on gender income inequalities.

6.1 Women Are More Likely to Lingering in Bunching

We assess the likelihood of lingering in bunching responses, separately by gender. Focusing on the first individual-specific bunching episode observed in the data, we then estimate the probability that the same individual will bunch in the n-th year following the first bunching episode.

Formally, we estimate the following model:

\[
Bunch_{i,t} = \sum_{k=1}^{7} \beta_k \cdot D_{i,t}^{k} + \gamma_i + \delta_{i,m(i)} + u_{i,t},
\]

where \(Bunch_{i,t}\) is a dummy variable indicating whether taxpayer \(i\) reports an income level below the spouse tax credit cutoff at year \(t\). \(D_{i,t}^{k}\) is a dummy variable for \(k\) years after the first bunching episode. The model also includes individual fixed
effects, $\gamma_i$, and year fixed effects interacted with municipality fixed effects, $\delta_{t,m(i)}$. Standard errors are clustered at individual-level.

Figure 10 presents the $\beta_k$ coefficient estimates, specifically for married male and married female taxpayers. The figure shows the probability of bunching from year +1 ($\beta_k = 1$) up to year +7 ($\beta_k = 7$) compared to the first bunching year. We also report 90 percent and 95 percent confidence intervals.

**Figure 10: Persistence Below the Threshold**

Notes: This figure shows the probability of lingering in bunching, separately by gender. The figure reports coefficient estimates from 1 year up to 7 years after the first bunching episode observed in our sample, along with 95% (delimited by horizontal bars) and 90% (bold line) confidence intervals. Period 0 represents the period when the taxpayer crosses the threshold for the first time. The specification always includes individual fixed effects and year-municipality fixed effects.

The figure provides striking evidence of gender differences in bunching persistence. We find that the probability of bunching in the year following the first bunching episode is around 17 percentage points larger for women with respect to men (41 versus 24 percent, respectively). This gender difference survives for many years: after 7 years since the first bunching period, there is still a significant 20 percent probability of a bunching response from women, while the corresponding probability for men is not significantly different from zero.
6.2 The Impact of the Spouse Tax Credit on the Gender Income Gap

Our analysis has provided two sources of gender differences in behavioral responses to the spouse tax credit. First, married women are much more likely than married men to reduce their income to let their spouse to receive the tax credit. Second, conditional on bunching on the spouse tax credit cutoff, women are more likely to lingering in bunching. Both these effects would positively contribute to generate gender income inequalities. This section provides descriptive evidence on how much the policy, and the the behavioral responses that it triggers, contributes to generate gender income gaps.

To evaluate the impact of the spouse tax credit on the gender income gap, we compute the log difference of male and female declared income in each income decile. We compute this measure for two samples: i. married taxpayers, that are directly affected by the spouse tax credit; ii. unmarried taxpayers, that should be less affected by the policy.

We present the gender gap in Figure 11. The figure shows that the gross incomes of married women are 4 log points lower than those of married men at the first income decile, that is where the spouse tax credit strikes. Significant gender gaps do not emerge at other income deciles, with the notable exception of the top decile. In support of the argument that the tax credit contributes to exacerbate gender income inequalities, we show that the gender gap in the bottom income decile is relatively smaller among unmarried individuals, where the spouse tax credit does not matter by definition.

Although we cannot rule out alternative explanations for the emergence of a gender gap at this point of the income distribution (e.g., self-selection in married status by low-income women), this result seems to suggest that the spouse tax credit significantly contributes to gender income inequalities.
7 Concluding Remarks

This paper leverages variation from an Italian policy that grants a large tax credit to the main earner in a couple when the second earner reports income below a cutoff. Using novel tax returns data, we document several important responses to the spouse tax credit. First, we show large bunching responses at the tax credit cutoff from second earner women, but no response from second earner men. This result suggests that household decisions are not Pareto-efficient when men are the second earner within the couple. Second, we show that gender differences in bunching are relatively larger among immigrants coming from more conservative societies, and natives living in more gender-traditional municipalities. Third, our results suggest that gender differences in bunching emerge after two salient events in couples’ life: marriage and childbirth.

Finally, we discuss the implications of the policy for female labor market outcomes and gender inequality. Although the economic rationale of the spouse tax credit is to offer insurance against labor market shocks, we find that the spouse tax credit persistently limits women’s careers and contributes to create a significant gender
References


Appendices

A Additional Tables and Figures

Table A1: Personal income tax schedule

<table>
<thead>
<tr>
<th>Taxable income (euros per-year)</th>
<th>Marginal tax rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>If composed only of income from real estates (up to 500 euros)</td>
<td>0</td>
</tr>
<tr>
<td>If composed only of retirement income (up to 7,500 euros) + income from land (up to 185,92 euros) + income from a main residence</td>
<td>0</td>
</tr>
<tr>
<td>&lt; 15,000</td>
<td>23</td>
</tr>
<tr>
<td>15,001-28,000</td>
<td>27</td>
</tr>
<tr>
<td>28,001-55,000</td>
<td>38</td>
</tr>
<tr>
<td>55,001-75,000</td>
<td>41</td>
</tr>
<tr>
<td>&gt; 75,000</td>
<td>43</td>
</tr>
</tbody>
</table>

Note: This table displays information the Italian personal income tax (IRPEF). Taxpayers are exempted from paying income taxes if their income is composed exclusively of real estates (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 net of deductible expenses, such as social security and welfare contributions or donations to non profit organizations.
Table A2: Descriptive statistics – Self-employed workers

<table>
<thead>
<tr>
<th>Panel A: Male</th>
<th>Average value</th>
<th>Standard deviation</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>.635</td>
<td>.481</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Age</td>
<td>45.73</td>
<td>10.77</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Foreign</td>
<td>.068</td>
<td>.253</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Gross income</td>
<td>34,695.34</td>
<td>54,866.94</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Taxable income</td>
<td>29,045.2</td>
<td>52,209.76</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Spouse tax credit (recipient)</td>
<td>.187</td>
<td>.390</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Spouse tax credit (amount)</td>
<td>123.84</td>
<td>263.76</td>
<td>1,892,849</td>
</tr>
<tr>
<td>Income tax</td>
<td>7,682.23</td>
<td>21,475.43</td>
<td>1,892,849</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Female</th>
<th>Average value</th>
<th>Standard deviation</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>.648</td>
<td>.477</td>
<td>822,493</td>
</tr>
<tr>
<td>Age</td>
<td>45.87</td>
<td>11.16</td>
<td>822,493</td>
</tr>
<tr>
<td>Foreign</td>
<td>.071</td>
<td>.257</td>
<td>822,493</td>
</tr>
<tr>
<td>Gross income</td>
<td>25,049</td>
<td>34,918.3</td>
<td>822,493</td>
</tr>
<tr>
<td>Taxable income</td>
<td>21,185.89</td>
<td>33,079.84</td>
<td>822,493</td>
</tr>
<tr>
<td>Spouse tax credit (recipient)</td>
<td>.020</td>
<td>.143</td>
<td>822,493</td>
</tr>
<tr>
<td>Spouse tax credit (amount)</td>
<td>14.40</td>
<td>99.67</td>
<td>822,493</td>
</tr>
<tr>
<td>Income tax</td>
<td>5,184.49</td>
<td>13,022.9</td>
<td>822,493</td>
</tr>
</tbody>
</table>

Note: The table displays the descriptive statistics. The variables Gross income, Taxable income, Spouse tax credit, Income tax and Spouse tax credit (amount) are expressed in Euros.
Table A3: Robustness to Polynomial Order

<table>
<thead>
<tr>
<th>Polynomial order</th>
<th>Women</th>
<th></th>
<th>Men</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td></td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.965***</td>
<td></td>
<td>-0.257</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.234)</td>
<td></td>
<td>(0.252)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.138***</td>
<td></td>
<td>-0.642***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td></td>
<td>(0.272)</td>
<td></td>
</tr>
<tr>
<td>7 (baseline)</td>
<td>1.238***</td>
<td></td>
<td>-0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.247)</td>
<td></td>
<td>(0.341)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1.197***</td>
<td></td>
<td>-0.111</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.251)</td>
<td></td>
<td>(0.363)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.433***</td>
<td></td>
<td>-0.224</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.244)</td>
<td></td>
<td>(0.318)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.497***</td>
<td></td>
<td>-0.116</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.287)</td>
<td></td>
<td>(0.347)</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* 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 A1: Take-Up Rate of Spouse Tax Credit By Income

*Notes:* This binscatter shows the take-up rate of spouse tax credit between married male and female as a function of the main earner’s gross income. The sample includes all married taxpayers with an income above 5,000 Euros.
Figure A2: Bunching Responses to the Spouse Tax Credit - Excluding the credit recipients

(a) Employees – Male married

(b) Employees – Female married

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 graph focuses on married male taxpayers; 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.
Figure A3: Bunching Responses to Tax Exemption Cutoff for Self-employees

(a) Men

(b) Women

Notes: These figures present density distributions around the tax notch determining tax exemption for self-employees, 4,800 Euros (denoted by the dashed vertical line). Left-hand side graph focuses on male taxpayers; right-hand side graph on female taxpayers. In each graph, we report the number of taxpayers (by 100 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.
Figure A4: Bunching Responses to the First Tax Bracket Cutoff (15,000 Euros)

(a) Men  
(b) Women

Notes: These figures present density distributions around the first tax bracket cutoff, 15,000 Euros (denoted by the dashed vertical line). Left-hand side graph focuses on male taxpayers; right-hand side graph on 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.
Notes: The figure presents density distributions around the tax notch determining eligibility for the spouse tax credit (denoted by the dashed vertical line). The sample includes all marries employees taxpayers. In the graph, we report the number of taxpayers (by 150 euros bins) for gross income. The graph also reports counterfactual distributions (in red), bunching estimates and bootstrapped standard errors, computed as described in Section 3.2.
Figure A6: Bunching (placebo) – Employees by gender

(a) Employees – Male married

(b) Employees – Female married

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: left-hand side graph focuses on married males; 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.
Figure A7: Bunching Responses - Male - Different Income Bandwidths

(a) Upper bound: 10th percentile

(b) Upper bound: 25th percentile

(b) Upper bound: 50th percentile

(b) Upper bound: 75th percentile

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.
Figure A8: Bunching Responses - Female - Different Income Bandwidths

(a) Upper bound: 10th percentile

(b) Upper bound: 25th percentile

(b) Upper bound: 50th percentile

(b) Upper bound: 75th percentile

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 Survey Evidence on Implications of Violating Gender Norms

This Appendix provides evidence on the implications of violating gender identity norms. We use data on married couples from 2013 to 2020 survey on Aspects of Daily Life (Indagine Multiscopo sulle Famiglie: Aspetti della Vita Quotidiana), provided by the Italian Institute of Statistics (ISTAT). For each couple, we collect information on who is the head of the household. In addition to basic demographics (region of residence, age, marital status, marriage tenure, education level) and labor market information (occupation, sector), the survey asks questions on a range of topics. We focus on two main groups of questions.

First, we retrieve information on life satisfaction, with questions focusing on both family, economic, friendship, and free-time satisfaction. For each question, respondents can choose answers on a scale of 1 (very happy) to 4 (very unhappy). We define a binary variable that simply takes value 1 if the respondent reports to feel “happy” or “very happy”. Second, we focus on their reported health status. We define a binary variable that takes value 1 if the respondents report to feel in very good or good health status.

To study whether the fact that the wife is the head of the household affects survey responses, we estimate the following linear probability model:

\[ y_i = \beta \cdot 1(WifeHeadHousehold_i) + \gamma \cdot X_i + \delta \cdot C_{s(i)} + u_i \]  

where \( y_i \) is the answer to a survey question by individual \( i \). In the analysis, we examine wives’ and husbands’ responses separately. In this way, we can study whether it is the wife or the husband who dislikes the reversal of traditional gender identity norms. \( 1(WifeHeadHousehold) \) is a dummy taking value 1 in couples when wives are the head of the household. \( X_i \) contain both demographic and labor market controls (region fixed effects, year fixed effects, education groups, age groups, marriage tenure, civic status, occupation fixed effects, and sector fixed effects). \( C_{s(i)} \) collect demographic and labor market information on the spouse of the respondent. We cluster the standard errors at the region level.

Table B1 presents our main results. We find that the violation of the male breadwinner model has important negative implications on couples’ life and economic satisfaction. Although we miss a plausibly exogenous source of variation in the probability of being the head of the household, we believe that this exercise provides three key suggestive results.

First, in couples where the wife’s income exceeds the husband’s, both the wife and the husband report to be less satisfied with their marriage (see panel A). Relative to
comparable couples where the husband is the main earner, couples where the wife outearns the husband are between 1 and 1.6 percentage points less likely to respond to feel “happy” or “very happy” with their marriage. In Table B2, we also show that this effect spreads through the family: when the male breadwinner model fails, daughters are 2.1 percentage points less likely to feel “satisfied” or “very satisfied” about their family, while sons appear to be less responsive.

Second, husbands are more likely to report anxiety disorders when the male breadwinner model fails (panel D). In couples where the wife is the main earner, husbands are 1.7 percentage points more likely to report to suffer of anxiety disorders. This result is consistent with social psychology research stating that infringing of internalized rules generates apprehension and nervousness.

Finally, despite our empirical exercise is based on comparing couples with similar observable economic conditions, we find that wives report to be significantly less satisfied with the economic conditions of their family when they are the main earner in the couple (panel B).

As a robustness check, we test whether this pattern is also spuriously reflected in other outcomes, for which the relationship with violating the male breadwinner should be less obvious, if any. In panel C, we replicate the analysis on friends’ satisfaction. Reassuringly, we find negative but imprecisely estimated coefficient.
Table B1: Failure of Male Breadwinner Model and Couples’ Outcomes

<table>
<thead>
<tr>
<th>A. Outcome: Marriage satisfaction</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.010***</td>
<td>-0.009**</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Husband</td>
<td>-0.016***</td>
<td>-0.015***</td>
<td>-0.015***</td>
<td>-0.015***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Observations</td>
<td>71,238</td>
<td>71,238</td>
<td>71,238</td>
<td>68,979</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Outcome: Economic satisfaction</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife</td>
<td>-0.026***</td>
<td>-0.030***</td>
<td>-0.031***</td>
<td>-0.024***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Husband</td>
<td>-0.005</td>
<td>0.000</td>
<td>-0.006</td>
<td>-0.014*</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Observations</td>
<td>71,450</td>
<td>71,449</td>
<td>71,448</td>
<td>54,640</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. Outcome: Friends’ satisfaction</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife</td>
<td>-0.004</td>
<td>-0.006</td>
<td>-0.006</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Husband</td>
<td>-0.012*</td>
<td>-0.010*</td>
<td>-0.010</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Observations</td>
<td>71,209</td>
<td>71,209</td>
<td>71,209</td>
<td>68,946</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D. Outcome: Anxiety</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wife</td>
<td>0.005</td>
<td>0.004</td>
<td>0.004</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Husband</td>
<td>0.017***</td>
<td>0.015**</td>
<td>0.016**</td>
<td>0.012*</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Observations</td>
<td>38,501</td>
<td>38,501</td>
<td>38,500</td>
<td>37,294</td>
</tr>
</tbody>
</table>

Region FE Yes Yes Yes Yes
Year FE Yes Yes Yes Yes
Demographics Yes Yes Yes Yes
Occupation FE No Yes Yes Yes
Sector FE No Yes Yes Yes
Partner demographics No No Yes Yes
Partner occ and sec FE No No No Yes

Note: This table shows how failure of the male breadwinner model affect survey responses from wives (first row in each panel) and husbands (second row). In panel A, B, and C, the outcome variable is equal to 1 if the respondent reports to feel “happy” or “very happy”. In panel D, the outcome variable is equal to 1 if the respondent reports to suffer of anxiety “often” or “very often”. In each panel, we report the β estimate obtained from equation (1) and region-level clustered standard errors. Each coefficient reports the difference in the outcome variable of interest between couples where the wife is the main earner (the male breadwinner model is violated) and couples where the husband is the main earner (male breadwinner model is satisfied). Therefore, each coefficient allows to test whether violating the male breadwinner model has significant implications.
### Table B2: Failure of Male Breadwinner Model and Children’s Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Outcome variable:</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Daughter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friend</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daughter</td>
<td></td>
<td>-0.021*</td>
<td>0.002</td>
<td>-0.023*</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>(0.012)</td>
<td>(0.016)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Son</td>
<td></td>
<td>-0.007</td>
<td>0.006</td>
<td>0.005</td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td>(0.009)</td>
<td>(0.018)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Region FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Demographics</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Note:** This table shows how failure of the male breadwinner model affect survey responses from daughters (first row in each panel) and sons (second row). In columns (1)-(3), the outcome variable is equal to 1 if the respondent reports to feel “happy” or “very happy”. In column (4), the outcome variable is equal to 1 if the respondent reports to suffer of anxiety “often” or “very often”. In each panel, we report the $\beta$ estimate obtained from equation (1) and region-level clustered standard errors. Each coefficient reports the difference in the outcome variable of interest between children where their mother is the main earner (the male breadwinner model is violated) and children where their father is the main earner (male breadwinner model is satisfied). Therefore, each coefficient allows to test whether violating the male breadwinner model has significant implications.