International Tax and Public Finance
https://doi.org/10.1007/s10797-026-09975-3

Willing but unable to pay? Gender, preferences, resources, and tax compliance

Andrea Lopez-Luzuriaga[1] · **Carlos Scartascini**[1]

Received: 19 December 2024 / Accepted: 5 May 2026

Abstract

The existing literature shows that women are generally more likely than men to pay taxes. However, there is less consensus on gendered responses to interventions designed to boost compliance among non-payers, and the underlying mechanisms remain unclear. In this study, we exploit a field experiment designed to increase property tax compliance to investigate this gender disparity. Because female- and male-led households differ along many dimensions, including preferences, household structure, and economic resources, we interpret the female–male contrast as the joint product of these channels rather than the effect of gender alone. Within this framing, the cross-sectional baseline gap in compliance is descriptive, while the within-gender treatment effects of the experimental messages are causally identified by random assignment. Our findings confirm that taxpayers in female-led households are more diligent in paying taxes than those in male-led households at baseline. Notably, while receiving a deterrence letter encourages women to pay earlier, it does not increase their overall compliance; men, in contrast, show a marked improvement in overall compliance upon receiving the same letter. We also find that the size of the tax bill influences women's compliance behavior (the likelihood of payment after treatment rises significantly for smaller tax bills) but has no similar effect on men. Survey data targeted at the same population suggest that, although women may be more motivated to pay (they perceive higher enforcement risk and exhibit higher tax morale) they also face tighter liquidity constraints, lower and less stable income, and a stronger perception that the tax is excessive. These patterns are consistent with a simple analytical model that links compliance to tax morale, risk aversion, and budget constraints, and in which liquidity-constrained taxpayers remain at a corner solution unresponsive to enforcement messages. The findings highlight how tax policies and enforcement procedures can exacerbate post-tax income inequality between genders, particularly in low-enforcement contexts where the tax base is weakly correlated with current income.

Extended author information available on the last page of the article

Published online: 29 May 2026

Keywords Taxes · Tax compliance · Liquidity constraint · Field experiment

JEL Classification H26 · H24 · D31 · J16

1 Introduction

Do women and men behave differently when faced with tax obligations? Abundant evidence from field interventions (Advani et al., 2023; Alstadsæter & Jacob, 2017; Cabral et al., 2015; Kleven et al., 2011; Wenzel, 2006) and laboratory experiments (Alm & Malézieux, 2021; Bazart & Pickhardt, 2011; D'Attoma et al., 2020, 2017; Eisenhauer et al., 2011; Finocchiaro Castro & Rizzo, 2014; Fortin et al., 2007; Kogler et al., 2016) shows that women are more likely to comply with their tax obligations than men. The main hypotheses for explaining this difference focus on two channels: women tend to be more risk-averse than men (Charness et al., 2018; Engstrom et al., 2015; Hibbert et al., 2013; Skatun, 2017), and women tend to have higher tax morale (Alm & Torgler, 2006; Cyan et al., 2016; Shafiq, 2015; Torgler, 2005; Torgler & Valev, 2010). A growing literature on gender and taxation, however, makes clear that the observed differences between female and male taxpayers are rarely the product of a single channel. Female- and male-led households also differ systematically along household structure, income level, income stability, caregiving responsibilities, and exposure to social norms regarding which household member is registered as the responsible taxpayer (Coelho et al., 2024; Lin & Slemrod, 2024; Lin et al., 2025; Maag et al., 2023). The "gender effect" on tax outcomes is conventionally understood, both in this literature and in the present paper, as the joint product of these intrinsic and structural channels.

If women are more likely to pay their taxes than men, does this imply they would be more responsive to a letter from the tax agency? The existing literature provides mixed and limited evidence. If women exhibit higher tax morale or are more risk-averse, an intervention that increases the salience of enforcement could be more effective for them.[1] On the other hand, if women face tighter budget constraints, because female-led households earn lower and less stable income on average, an intervention that does not relax the cash constraint may have a smaller effect on overall compliance, even when it successfully shifts perceptions. The net direction of the effect is therefore an empirical question, and one that field-experimental evidence has rarely addressed directly.

In this article, we investigate how taxpayers in female- and male-led households respond differently to messages aimed at enhancing property tax compliance, by re-analyzing the field experiment of Castro and Scartascini (2015) along gender lines. The original experiment, conducted in the municipality of Junín, Argentina, in 2011, included three treatment arms: one emphasizing penalties and detection probability (deterrence message) and two conveying tax morale messages (reciprocity and peer-effects messages). We complement the experimental analysis with two surveys cov-

[1] This would correspond to an interior solution in the standard taxpayer's decision problem.

ering the same population, which allow us to characterize differences in motivations and resources between male- and female-led households.

Our analysis distinguishes between two separate empirical questions. The cross-sectional comparison of women and men at baseline is, like in much of the gender-and-taxation literature, descriptive: it reflects the joint contribution of preferences, household structure, income, and norms, and we cannot fully decompose it with the data at hand. The treatment effects of the deterrence, reciprocity, and peer-effects messages, however, are causally identified by the within-gender randomization in the original experiment. The differential treatment effect across gender (the interaction term) likewise rests on within-gender comparisons that do not require us to take a stand on what generates the baseline gap. Throughout the paper, we are explicit about which findings rest on which identification argument.

The Junín property tax has features that make it especially well-suited to our analysis. The city government calculates the tax based on simple administrative inputs (the linear frontage of the lot, the public services available at the property's location) rather than on a market-based valuation, and the bill is issued bi-monthly. Informational asymmetries are absent, leaving taxpayers with a binary decision: to pay the full bill or not (no partial payments are accepted). The tax design, monitoring, availability of payment plans, and other associated features do not vary by gender. Importantly, mortgage financing in Argentina accounts for less than 1% of GDP, one of the lowest rates globally, so housing is illiquid and the property tax is poorly correlated with the taxpayer's current income flow. This disconnect between the tax base and current income is common in the developing world due to thin credit markets, limited options for leveraging assets as collateral, and heavy reliance on indirect taxes; in Latin America and the Caribbean, personal income taxes account for approximately 10% of total revenue, compared with around 25% in the OECD (Acosta-Ormaechea et al., 2022; Corbacho et al., 2013).[2] This setting is therefore one in which liquidity constraints are a plausible margin of variation in compliance behavior, unlike most of the existing literature on tax compliance, which is set in high-income contexts with proportional income taxation.

Our empirical findings are the following. First, taxpayers in female-led households are more likely to pay than those in male-led households at baseline, both in terms of the unconditional probability of payment and in terms of timely payment. Second, the deterrence message produces sharply different responses across gender: it shifts the *timing* of payment for women (women in the deterrence group are 3.6 percentage points more likely to pay by the first due date and 2.8 percentage points more likely to pay by the second due date), but does not increase their overall probability of payment. For men, the pattern is the opposite: the deterrence message has no effect on the probability of paying by the first due date but increases the overall probability of payment by 2.0 percentage points. We characterize the female pattern as a response on the *timing margin* (paying earlier, conditional on paying) and the

[2] Moreover, in Argentina and other developing countries, a significant proportion of taxpayers owing income tax are part of a simplified tax regime. In these regimes, the tax owed remains constant within broad income brackets. For instance, in Argentina in 2021, individuals at the lower bound of the first bracket paid about 2% of their sales in income taxes, while those at the upper bound of the same bracket paid less than 1%.

male pattern as a response on the *extensive margin* (paying versus not paying). Third, the heterogeneous effects of the deterrence message by tax size are consistent with a liquidity-constraint mechanism for women but not for men: the female deterrence effect is positive and large (up to 10 p.p.) for low tax bills and disappears as the tax liability grows. The male deterrence effect, by contrast, does not vary with tax size.

To interpret these findings, we develop a simple analytical model in which a taxpayer chooses whether to pay a government-assessed tax bill or to face a stochastic penalty. The model incorporates three parameters that vary across taxpayer types: tax morale (an intrinsic motivation to comply), risk aversion, and a minimum-consumption constraint that binds when the tax bill is large relative to current income. The predictions are stated in terms of taxpayer types rather than gender per se: taxpayers with higher tax morale or higher risk aversion respond more strongly to a deterrence intervention, while taxpayers facing a binding liquidity constraint remain at a corner solution and do not respond at all. The empirical predictions for the female–male contrast then follow as a special case of the general model, given the documented differences in preferences and resources between the two groups. Survey evidence corroborates the mechanism: female-led households perceive higher enforcement risk, exhibit higher tax morale, and are simultaneously poorer, less likely to receive a steady wage, and more likely to view the property tax as excessive.

Our paper contributes to several literatures. We add to the growing literature on gender and taxation (Brehm & Malkova, 2023; Coelho et al., 2024; Lin & Slemrod, 2024; Lin et al., 2025; Maag et al., 2023) by providing field-experimental evidence on how the bundle of attributes associated with a registered taxpayer's gender shapes the response to enforcement. We add to the literature on the effectiveness of targeted messages in promoting tax compliance (Antinyan & Asatryan, 2024; Mascagni, 2018; Slemrod, 2019) by showing that the same message can produce qualitatively different responses across subpopulations, with implications for how tax authorities design interventions. We add to the literature on tax policy and gender inequality in developing economies (Bird & Zolt, 2008; Grown & Valodia, 2010) by documenting a mechanism, a differential response to enforcement under liquidity constraints, through which blunt enforcement instruments can widen post-tax inequality. Finally, we add to the literature on property taxation (Brockmeyer et al., 2021; Castro & Scartascini, 2015; Chirico et al., 2019, 2016) by highlighting the role of liquidity constraints in the payment decision, which is distinctive to property taxes in illiquid settings.

The policy implications follow directly. In environments with weak enforcement and a tax base that is poorly correlated with current income, blunt enforcement instruments may exacerbate post-tax inequality between genders, since the population most willing to comply is also the most cash-constrained. We argue that complementary tools, such as income-contingent payment plans or targeted property tax discounts for low-income households, can preserve the equity-enhancing role of stronger enforcement while addressing the differential burden it places on liquidity-constrained taxpayers.

2 Background and data

The data for this analysis come from a large-scale field experiment conducted by Castro and Scartascini (2015) to investigate the determinants of property tax compliance in Junín, Argentina, in 2011. We describe the institutional setting of the property tax (Section 2.1), the experimental design (Section 2.2), the construction of the gender variable and the dependent variables (Section 2.3), and the auxiliary survey data we use to interpret the results (Section 2.4).

2.1 The property tax in Junín

The property tax (*Contribución por Servicios a la Propiedad*, hereafter CVP) is levied on homes, farms, business premises, and most other real estate in the city of Junín. Approximately 96% of the properties in our sample are residential; commercial properties account for approximately 4%, and the remaining categories (rural lots, garages, etc.) account for less than 1% combined. The composition is similar across male- and female-led households.

The city government calculates the tax based on administrative inputs that the municipality observes directly: the lot's linear frontage in meters, the public services provided to the block (street lighting, trash collection, street cleaning), and a small number of other neighborhood-level variables. Notably, the tax is *not* computed from a market-based valuation. Although the municipal registry records a cadastral value (*valor fiscal*) for each property, this value is updated infrequently and is not used in computing the bill. Two related institutional features make the cadastral value a poor proxy for market value: mortgage credit in Argentina accounts for less than 1% of GDP, so there is no transaction-based pressure to keep cadastral valuations current; and the registry is updated unevenly, so properties of similar market value can have very different cadastral values depending on when they last entered the registry. For these reasons, we use the inputs to the tax computation (frontage and public services) as our primary controls in the empirical analysis, and report robustness specifications that include the cadastral value as an additional covariate.

The tax bill is issued bi-monthly. Each bill has two due dates: the initial due date typically falls in the second week of the month, and the secondary due date in the following week. Payment is expected by the initial due date, but no late fees are levied if payment is made by the secondary due date. Outstanding liabilities accrue a monthly compound interest rate of 2%. We exploit this two-stage payment scheme to analyze compliance at different points within the billing cycle. Importantly, the city does not accept partial payments: for any given bill, the taxpayer either pays the full amount or pays nothing. Compliance is therefore binary at the property-bill level. The tax design, monitoring, availability of payment plans, and other associated features do not vary by gender.

2.2 Experimental design

The experiment was conducted in 2011 and included approximately 26,000 individual taxpayers registered to pay the property tax as of the August 2011 baseline.

The 3,000 individuals who had elected to pay the tax annually rather than bi-monthly were excluded from the randomization. The remaining taxpayers were randomly assigned to one of three treatment arms or to a control group, with stratification at the section (block) level. The intervention added a printed message to the tax bill, with three distinct treatment messages: a deterrence message detailing penalties for late payment, a reciprocity message describing the uses of collected funds, and a peer-effect message providing information about the overall compliance rate. The text of each message is shown in Table 1, and an example of a treated tax bill is reproduced in Figure A1 in the Online Appendix. Further details on the experimental implementation are provided in Castro and Scartascini (2015).

The original experiment was conducted in 2011, predating the now-standard practice of pre-registering field experiments in economics. The experiment was therefore not pre-registered, and the present paper, as a re-analysis of the original data with a focus on gender heterogeneity, is likewise not pre-registered against the original protocol. We report estimates for all three treatment arms across all three timing outcomes, for both the full sample and the male and female subsamples, with no selective reporting; the full set of results, including robustness specifications, is reported in the main text or the Online Appendix.

2.3 Variables

The taxpayer database includes the names of each property owner and the individual responsible for paying the tax. From this information, we infer the gender of the registered taxpayer. In Argentina, parents select their children's names from a pre-approved list of approximately 10,000 female and male names, with negligible overlap between the two lists.[3] Using this list, we construct a gender variable for 92%

Table 1 Message Included in the Property Tax Bill

Message/Group	Text	Image
Deterrence	Did you know that if you do not pay the CVP on time for a debt of AR$ 1,000 you will have to disburse AR$ 268 in arrears at the end of the year and the Municipality can take administrative and legal action?	
Reciprocity	In the first 6 months of this year, CVP's collection contributed to placing 28 new streetlights, water connections in 29 streets and sewerage networks in 21 blocks	
Peer-effects	Did you know that only 30% of taxpayers do not pay the CVP? What about you?	
Control	No message	No image

[3] See https://data.buenosaires.gob.ar/dataset/nombres.

of the sample, or about 21,500 taxpayers, 34% of whom are women.[4] Throughout the paper, we refer to the gender of the registered taxpayer, and we describe households in which the registered taxpayer is female (respectively, male) as "female-led" (respectively, "male-led") households. We are explicit in Section 5 about the limits of this terminology: female- and male-led households differ along several dimensions besides the gender of the registered taxpayer, and our data do not permit a full decomposition of these differences.

Dependent variables. We construct three binary indicators of compliance:

- *Paid by the 1st due date*: equal to one if the taxpayer paid the bill on at least one of their properties on or before the initial due date, and zero otherwise.
- *Paid by the 2nd due date*: equal to one if the taxpayer paid the bill on at least one of their properties on or before the secondary due date.
- *Paid (overall)*: equal to one if the taxpayer paid the bill on at least one of their properties on or before the end of the bimonthly billing cycle.

The "at least one property" coding is the natural choice for this sample, in which the median taxpayer owns a single property (mean of 2.2 properties, with a small right tail of multi-property owners). For the median taxpayer, the three coding rules—at least one bill paid, at least 60% paid, all paid—are equivalent. We report robustness specifications under the stricter codings in Tables A6 and A7 in the Online Appendix; the qualitative pattern of results is robust across all three definitions.

2.4 Baseline differences and auxiliary surveys

Female- and male-led households differ systematically along several dimensions, both at baseline in the experimental sample and in the broader population of the region. We summarize the relevant differences here, with full statistics reported in Tables A1 and A3 in the Online Appendix.

Compliance and property characteristics in the experimental sample. At baseline (the bimonthly period preceding the experiment), 44% of women paid the property tax, compared with 39% of men; 24% of women paid on time versus 21% of men; and 54% of women owe at least one outstanding bill, versus 60% of men. These descriptive differences are consistent with the established meta-analytic finding that tax compliance is, on average, several percentage points higher for women than for men (Hofmann et al., 2017). Properties owned by female-led households also differ from those owned by male-led households: women's properties are on blocks with more streetlights (2.8 versus 2.7 on average) and a higher public-services index, indicating more central locations; men's properties have, on average, one more meter of linear frontage; and men own more properties on average (2.4 versus 1.9). The cadastral value is somewhat higher for women's properties on average, reflecting the

[4] Our analysis is limited to gender as assigned at birth. Since 2012, it has been possible in Argentina to change one's legal name and gender based on self-identification. In the first ten years of the law, only 0.027% of Argentines changed their gender, so we are confident that gender at birth serves as a good proxy for gender identification in our sample.

higher value associated with central, well-serviced locations even though men's lots are physically larger. We control for these property characteristics throughout the empirical analysis.

Household resources from the EAHU. To characterize differences in household resources between male- and female-led households, we use the Urban Household Survey of 2011 (*Encuesta Anual de Hogares Urbanos*, EAHU). The EAHU does not allow us to match individual taxpayers in our experimental sample to survey respondents, but it provides representative information for the broader region in which Junín is located.[5] The patterns are stark. Male-led households have approximately 37% higher monthly household income than female-led households and a 13 percentage point higher probability of receiving a steady wage. Female-led households have fewer members on average (2.4 versus 3.1) and are much less likely to include a spouse or partner of the household head (15% versus 77%). These differences imply that female- and male-led households differ in income level, income stability, and household composition—each of which can plausibly affect both the level of tax compliance and the response to enforcement messages.

Survey on attitudes toward the property tax. We also draw on a survey conducted by the city government of Junín after the experiment, which targets the household member responsible for property tax payment and gathers their attitudes toward the tax. The survey is not linked at the individual level to the experimental sample, but it provides information on perceptions among the relevant population. We use this survey in Section 5 to characterize gendered differences in perceptions of enforcement risk and tax burden.

The "gender bundle." Taken together, female- and male-led households in our setting differ along the gender of the registered taxpayer, household structure, income, income stability, property characteristics, location, and—as the literature documents—likely also along risk aversion, tax morale, and exposure to social norms regarding which household member is registered as the responsible taxpayer (Alm & Malézieux, 2021; Lin et al., 2025; Torgler & Valev, 2010). We refer to this constellation of correlated attributes as the "gender bundle." Throughout the empirical analysis, we are careful to distinguish what can be identified from the experimental data: the cross-sectional comparison between female- and male-led households at baseline is descriptive, while the within-gender treatment effects of the experimental messages are causally identified by the random assignment of treatment.

Balance. Reassuringly, the gender variable is balanced across treatment and control groups, both in the full sample and within each gender subsample (Table A2 in the Online Appendix). Treatment assignment is therefore orthogonal to all observed and unobserved characteristics, both overall and conditional on the gender of the registered taxpayer.

[5] The Buenos Aires province is divided into six sub-regions for the EAHU: Buenos Aires (city), Gran La Plata, Bahía Blanca, Partidos del GBA, Mar del Plata, and a residual category for smaller cities. Junín belongs to the residual category.

3 Empirical results

We estimate the following linear probability model:

$$y_{ib} = \alpha + \sum_{j=1}^{3} \beta_j T_i^j \; + \; \gamma \, \mathrm{Male}_i \; + \; \sum_{j=1}^{3} \delta_j \left(T_i^j \times \mathrm{Male}_i \right) \; + \; \mathrm{X}_i' \boldsymbol{\lambda} \; + \; \mu_b \; + \; \varepsilon_{ib}, \quad (1)$$

where y_{ib} is one of three payment outcomes for taxpayer i in randomization block b (paid by the 1st due date, paid by the 2nd due date, or paid overall); T_i^j for $j \in \{1, 2, 3\}$ is an indicator for assignment to the deterrence, reciprocity, or peer-effects treatment, respectively; Male_i is an indicator for a male-led household; and X_i is a vector of controls including the lagged dependent variable, the log of the number of properties owned, the log of the average linear frontage, indicators for trash collection and street lighting services, and a dummy for taxpayers who elected to pay monthly. μ_b is a randomization-block fixed effect. Standard errors are clustered at the randomization-block level. The coefficients β_j capture the treatment effect for female-led households (the omitted gender category), and δ_j captures the differential treatment effect for male-led households. We choose the controls in X_i to match the inputs the city government uses to compute the tax bill (frontage and public services), which provide more reliable variation in property characteristics than the cadastral value; we report a robustness specification that adds the log of the cadastral value in Table A5 of the Online Appendix.[6]

The specification in equation (1) extends Castro and Scartascini (2015) by adding the Male_i indicator and the treatment-by-gender interactions, which are the focal coefficients of the present analysis. The set of controls and the clustering structure are otherwise identical to that paper. As a presentational alternative, we also estimate equation (1) on the male and female subsamples separately, omitting γ and the interaction terms. The split-sample and pooled-with-interaction specifications are mathematically equivalent: the coefficient on T^j in the female-only regression equals β_j in the pooled regression, and the coefficient on T^j in the male-only regression equals $\beta_j + \delta_j$.

3.1 Average treatment effects

Treatment was randomized at the taxpayer level, and gender is balanced across the four treatment arms (Table A2 in the Online Appendix). The within-gender treatment effects (β_j for women, $\beta_j + \delta_j$ for men) and the differential treatment effect (δ_j) are therefore causally identified by random assignment within each gender stratum. The cross-sectional comparison of women and men at baseline is not so identified: it captures the joint contribution of gender, household structure, income, norms, and other correlated characteristics, and we cannot decompose it.

[6] We do not include the size of the tax bill itself as a control because the tax is a deterministic function of the inputs we already condition on (frontage and public services) and of block-level fixed effects. We use the size of the tax bill instead as the conditioning variable in the heterogeneity analysis of Section 3.2.

Table 2 reports the estimated treatment effects for the bimonthly period of September–October 2011, when the experimental messages were delivered. The three columns of each panel report, in order, the male-only regression, the female-only regression, and the pooled regression with treatment-by-gender interactions. Figures 1 and 2 display the results graphically.

The deterrence message produces qualitatively different responses across gender. The most striking pattern in Table 2 concerns the deterrence message, which had been identified by Castro and Scartascini (2015) as the most effective treatment in the original full-sample analysis. We find that the deterrence message shifts *when* women pay but not *whether* they pay, while it shifts *whether* men pay but barely affects when. Specifically, women in the deterrence group are 3.6 percentage points more likely to pay by the first due date and 2.8 percentage points more likely to pay by the second due date than women in the control group, both significant at the 5% level. The effect on overall payment, however, is a statistically insignificant 1.1 percentage points. Men in the deterrence group are 2.0 percentage points more likely to pay overall than men in the control group, significant at the 5% level, with a smaller and weakly significant effect on payment by the second due date (1.3 p.p., $p < 0.10$) and no effect on payment by the first due date.

We characterize the female pattern as a response on the *timing margin* (women who would have paid in any case pay earlier when prompted by a deterrence message) and the male pattern as a response on the *extensive margin* (men who would not otherwise have paid are induced to pay by the same message).[7]

The differential treatment effect for the deterrence message, captured by the coefficient on T1 $\times$ Male in the pooled specification, is statistically significant for the first-due-date outcome at the 10% level in the no-controls specification (Table A4 in the Online Appendix) and is consistent in sign with the within-gender results. The implied treatment effect for men in the pooled specification (the linear combination $\beta_1 + \delta_1$) equals 0.020 with a standard error of 0.008 for the overall payment outcome, identical by construction to the male-only coefficient and significant at the 5% level.

The reciprocity and peer-effects messages have weaker and less differentiated effects. The reciprocity message reduces women's probability of paying by the first due date by 2.4 percentage points (significant at the 10% level) and has a small positive but insignificant effect on men. The differential effect for men (δ_2) is positive and marginally significant for the first-due-date outcome. The peer-effects message has small and statistically insignificant effects across both genders and all three outcomes. The reciprocity result is consistent with the broader finding in Castro and Scartascini (2015) that taxpayers who already receive substantial public goods from the city government respond negatively when reminded of how their taxes are spent; women in our sample tend to live in better-serviced central areas, which may explain the reciprocity effect rather than gender per se.

[7] Strictly speaking, in property tax compliance, there is no intensive margin in the conventional sense, since the city does not accept partial payments; the only choice is whether to pay the full bill or not. We use the term *timing margin* to capture the analogous within-payer dimension of variation: among taxpayers who eventually pay, by which due date do they pay?

Table 2 Average treatment effect (Sep/Oct)

	Paid 1st due date			Paid 2nd due date			Paid		
	Men	Women	Int	Men	Women	Int	Men	Women	Int
Men			−0.009			−0.009			−0.008
			(0.009)			(0.007)			(0.007)
T1: Deterrence	0.006	0.036**	0.037**	0.013*	0.028**	0.029**	0.020**	0.011	0.012
	(0.008)	(0.016)	(0.016)	(0.007)	(0.012)	(0.012)	(0.008)	(0.010)	(0.010)
T1: Deterrence × Men			−0.031			−0.016			0.008
			(0.019)			(0.015)			(0.012)
T2: Reciprocity	0.006	−0.024*	−0.023*	0.005	−0.016	−0.014	0.006	−0.014	−0.013
	(0.011)	(0.012)	(0.012)	(0.009)	(0.011)	(0.011)	(0.009)	(0.011)	(0.011)
T2: Reciprocity × Men			0.030*			0.020			0.020*
			(0.017)			(0.012)			(0.011)
T3: Peer-Effect	0.005	0.000	0.000	0.006	0.001	0.002	0.001	−0.001	−0.001
	(0.006)	(0.014)	(0.014)	(0.007)	(0.011)	(0.011)	(0.009)	(0.011)	(0.011)
T3: Peer-Effect × Men			0.005			0.004			0.002
			(0.016)			(0.011)			(0.014)
N	14,003	7,301	21,304	14,003	7,301	21,304	13,995	7,292	21,287
Gender	Men	Women	Both	Men	Women	Both	Men	Women	Both
$F(\beta_{T1}= 0, \beta_{gender} = 0, \beta_{T1\times gender} = 0)$			5.780			5.502			2.495
$p(\beta_{T1}= 0, \beta_{gender} = 0, \beta_{T1\times gender} = 0)$			0.004			0.005			0.084
$F(\beta_{T2}= 0, \beta_{gender} = 0, \beta_{T2\times gender} = 0)$			1.347			0.992			1.070
$p(\beta_{T2}= 0, \beta_{gender} = 0, \beta_{T2\times gender} = 0)$			0.283			0.413			0.380
$F(\beta_{T3}= 0, \beta_{gender} = 0, \beta_{T3\times gender} = 0)$			0.497			0.809			0.538
$p(\beta_{T3}= 0, \beta_{gender} = 0, \beta_{T3\times gender} = 0)$			0.688			0.501			0.661
Pr. of paying for each category									
T1: Men	0.260		0.260	0.348		0.348	0.417		0.417
T1: Women		0.329	0.330		0.412	0.413		0.467	0.468
T2: Men	0.260		0.260	0.340		0.340	0.403		0.404
T2: Women		0.270	0.270		0.368	0.369		0.442	0.442
T3: Men	0.259		0.259	0.341		0.340	0.398		0.398
T3: Women		0.294	0.294		0.385	0.386		0.455	0.455

Table 2 (continued)

	Paid 1st due date			Paid 2nd due date			Paid		
	Men	Women	Int	Men	Women	Int	Men	Women	Int
C : Men	0.254		0.254	0.335		0.335	0.397		0.397
C : Women		0.294	0.293		0.384	0.384		0.456	0.456

Notes: All regressions include as controls the lagged variable, fixed effects for blocks, variables for public service provision (trash collection and street lightning services during the period), the (log of the) number of properties that each taxpayer has, the (log of the) average linear front size of the properties, and a dummy that controls for those taxpayers who elected to pay monthly

Standard errors in parentheses are clustered by randomization blocks

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

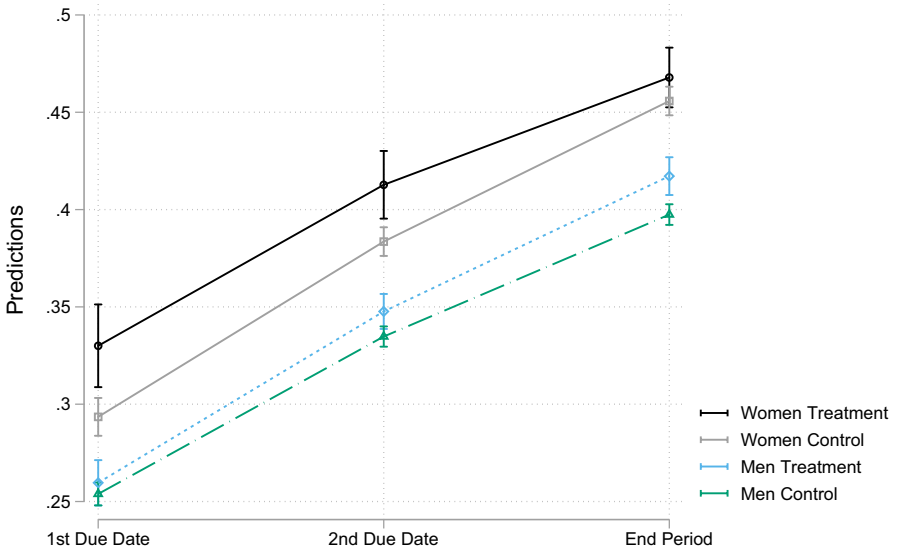

Fig. 1 Effects of the deterrence treatment

The main results are robust to several alternative specifications. Including the log of the cadastral value as an additional control leaves the point estimates and standard errors essentially unchanged (Table A5 in the Online Appendix). The pattern is also robust to alternative codings of the dependent variable for multi-property taxpayers: under the stricter coding requiring at least 60% of bills to be paid (Table A6 in the Online Appendix), and under the strictest coding requiring all bills to be paid (Table A7 in the Online Appendix), the qualitative pattern is unchanged. Notably, under the "all paid" coding, the female deterrence effect on overall payment becomes statistically significant at the 5% level (2.8 p.p., s.e. 0.013), consistent with the liquidity-constraint mechanism we develop in Section 4: among multi-property female taxpayers, the deterrence message appears to shift behavior on a margin that the lenient "at least one paid" coding does not capture. Restricting the sample to residential properties only (Table A8 in the Online Appendix) yields point estimates that are slightly larger than the baseline, consistent with the residential subsample being a

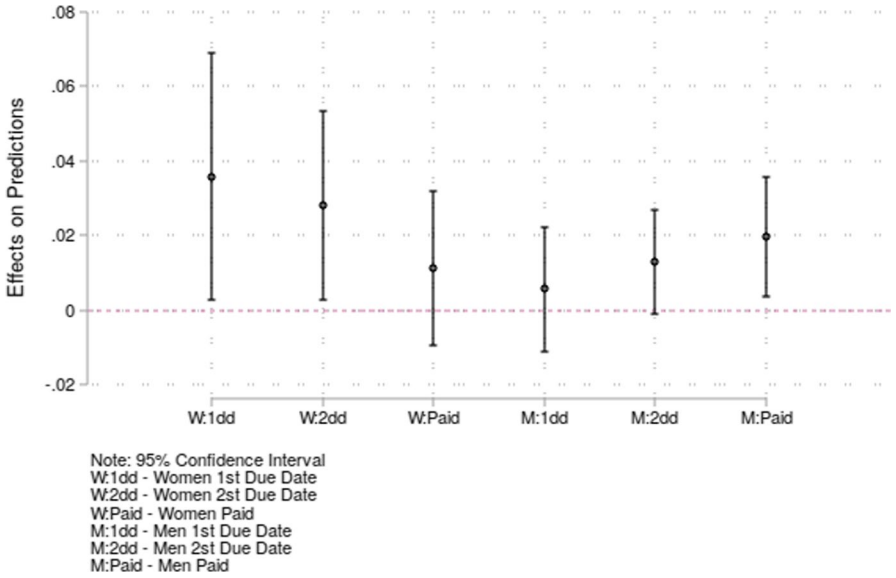

Fig. 2 Marginal effects of the deterrence treatment

more homogeneous group. Estimates on the small commercial-property subsample (Table A9 in the Online Appendix) are noisy, as expected given the smaller sample size, and we do not place weight on them.

3.2 Heterogeneous treatment effects by tax size

If the female deterrence effect on overall payment is muted by liquidity constraints, the effect should be concentrated among taxpayers with small bills—those for whom the cash constraint is least likely to bind—and should attenuate as the bill grows. The male deterrence effect, by contrast, should not display the same pattern if liquidity is not the primary margin of variation for men. This is exactly what the data show.

Figure 3 displays the marginal effect of the deterrence message on overall payment as a function of the log of the tax bill, separately for women and men. For women, the effect is positive and large at low levels of the tax bill (up to approximately 10 percentage points for the smallest bills) and declines monotonically as the bill grows, becoming statistically indistinguishable from zero in the upper portion of the distribution. For men, the marginal effect is roughly flat across the entire support of the tax distribution, hovering near the average treatment effect of 2 percentage points. The contrast is consistent with the liquidity-constraint mechanism we formalize in the next section: women appear to respond to the deterrence message on the extensive margin only when the implied cash outlay is small, while men's response does not appear to be conditioned on the size of the bill.

These findings motivate the analytical framework we develop next, which embeds tax morale, risk aversion, and a minimum-consumption constraint in a simple compliance model and derives predictions for how taxpayers of different types respond to a deterrence intervention.

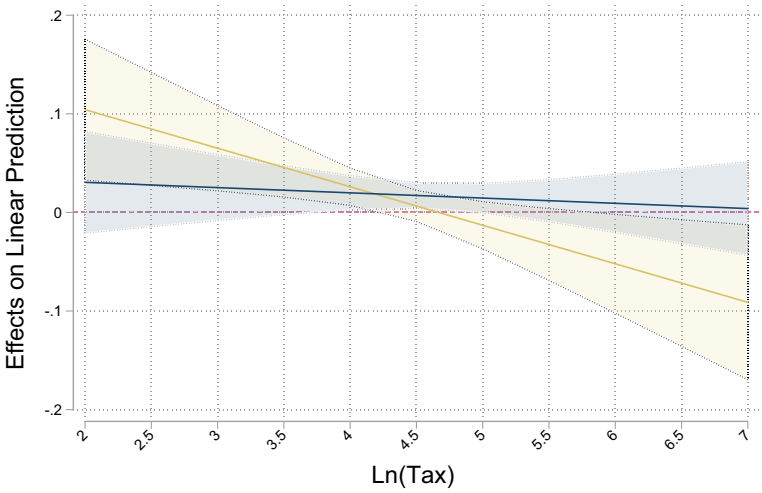

Fig. 3 Heterogeneous effects of the deterrence treatment by tax liability and gender. *Notes:* Blue line corresponds to men and the yellow to women

4 An analytical framework: tax morale, risk aversion, and liquidity

The empirical patterns documented in Section 3 call for a unified explanation. Any plausible mechanism must rationalize four observations simultaneously: (i) taxpayers in female-led households are more likely to pay at baseline; (ii) they respond to a deterrence message on the timing margin but not on the extensive margin; (iii) taxpayers in male-led households respond on the extensive margin; and (iv) the female extensive-margin response is strongly conditioned on the size of the tax bill, while the male response is not.

To organize these patterns, we sketch a simple compliance model whose full derivation is in Online Appendix C. The model is intentionally minimal and not novel: it embeds three standard ingredients, tax morale, risk aversion, and a minimum-consumption constraint, in the canonical taxpayer's decision problem, following Dwenger et al. (2016) on tax morale and adopting a binary pay/no-pay choice that mirrors the institutional setting of the property tax in Junín. We deliberately formulate the predictions in terms of taxpayer *types* characterized by these three parameters, rather than in terms of gender per se. The female–male contrast then arises as a special case of the general model, given the documented differences between female- and male-led households along each of the three dimensions. This formulation makes clear that the mechanism is not specific to gender: any subgroup with systematic differences in the relevant parameters would generate analogous patterns.

4.1 Predictions for taxpayer types

A taxpayer with current income Y chooses whether to pay a government-assessed property tax T in full or to pay nothing (partial payments are not permitted). If she pays, she enjoys after-tax income augmented by an intrinsic-motivation premium

$S \geq 0$. If she does not pay, she is detected with perceived probability p and faces a penalty proportional to T. As shown in Online Appendix C, the unconstrained problem yields an indifference probability p^* above which the taxpayer chooses to pay. Three parameters determine where any given taxpayer sits relative to p^* and how she responds to an intervention that shifts perceived enforcement upward.

Tax morale (S). Higher tax morale lowers p^*. Taxpayers with higher tax morale are more likely to comply at baseline and more likely to be brought into compliance by an intervention that shifts perceptions upward.

Risk aversion ($R_A(Y)$). Higher risk aversion likewise lowers p^*. Taxpayers who dislike the gamble of evasion are more likely to comply, and more likely to be moved by an intervention that increases the salience of detection.

Liquidity constraint ($\bar{C}$). The first two parameters determine behavior at an interior solution. We add a binding minimum-consumption constraint to capture the institutional feature, central to property taxation in illiquid settings, that the tax is calculated from property characteristics rather than from current income. Because mortgages are essentially unavailable in Argentina and the home cannot easily be converted into liquid resources, the tax bill can exceed what the taxpayer can pay out of current income while still meeting basic consumption needs (the same mechanism documented for Mexico City by Brockmeyer et al. 2021). When the constraint binds, the taxpayer is at a corner solution and does not pay the tax *regardless of perceptions*: shifting the perceived probability of enforcement does not relax the cash constraint.

Three taxpayer types. Combining these parameters, we can classify taxpayers into three types:

1. *High-morale or high-risk-aversion, unconstrained.* Already pay at baseline. An enforcement message reinforces the decision and may shift the timing of payment, but does not change the extensive-margin outcome.
2. *Low-morale and low-risk-aversion, unconstrained.* Do not pay at baseline. An enforcement message can move them across p^* and into compliance.
3. *Liquidity-constrained.* Do not pay at baseline. An enforcement message does not change behavior because the corner solution is not driven by perceptions.

4.2 Mapping to the female–male contrast

The empirical literature documents systematic differences across the three parameters between women and men. On tax morale, women score higher across countries and methods (Alm & Torgler, 2006; Cyan et al., 2016; Shafiq, 2015; Torgler, 2005; Torgler & Valev, 2010). On risk aversion, women are more risk-averse on average (Charness et al., 2018; Engstrom et al., 2015; Hibbert et al., 2013; Skatun, 2017). On liquidity, female-led households earn lower and less stable income on average, both globally and in our specific setting (Section 2). Whether these differences reflect biology, evolved preferences, socialization, structural labor-market constraints, or social norms about which household member is registered as the responsible taxpayer is the subject of an active literature (Lin et al., 2025) and is not something our data can adjudicate. What matters for the model's predictions is only that the differences exist.

If female- and male-led households differ along all three parameters in the directions documented above, the model predicts: (i) higher baseline compliance for female-led households, because higher tax morale and higher risk aversion both reduce p^*; (ii) a stronger deterrence response on the timing margin for women, because women already inclined to pay (Type 1) are reminded by the message to do so promptly, whereas men have a smaller share of Type 1; (iii) a stronger deterrence response on the extensive margin for men, because men are more likely to be in Type 2 and the message moves them across p^*; (iv) heterogeneity in the female extensive-margin response by tax size, since a larger share of female-led households are in Type 3 (liquidity-constrained), and the share rises with the size of the bill, so the average female treatment effect attenuates toward zero as the bill grows; and (v) a flat response by tax size for men, because the corner-solution share is smaller across the tax distribution. These predictions match the empirical patterns documented in Section 3.

4.3 What the model does not identify

The model is consistent with our empirical findings, but consistency is not identification of the underlying parameters. Our data do not allow us to separately measure tax morale, risk aversion, or the location of the liquidity threshold for individual taxpayers, so we cannot recover structural estimates of S, $R_A(Y)$, or $\bar{C}$. What we can do is test predictions that are diagnostic of the joint mechanism—the timing-versus-extensive split across genders, and the heterogeneity of the female extensive-margin response by tax size—both of which are difficult to rationalize without invoking liquidity constraints. Section 5 draws on auxiliary survey data to provide additional, more direct evidence on each of the three model parameters.

5 Discussion

The analytical framework of Section 4 predicts that female-led households should comply more at baseline (because of higher tax morale and risk aversion), should respond to enforcement on the timing margin (because the unconstrained Type-1 share is larger among women), should respond on the extensive margin only when liquidity does not bind (which explains the heterogeneity by tax size), and should be largely unresponsive when bills are large (because the corner-solution share grows with the bill). Section 3 shows that the empirical patterns line up with these predictions. In this section, we use auxiliary survey data to provide additional, more direct evidence on each of the three model parameters, and we close with an explicit acknowledgment of what the analysis cannot identify.

5.1 Survey evidence on the three channels

Tax morale and perceived enforcement. The post-experiment survey conducted by the city government of Junín asked respondents to evaluate, on a 7-point Likert scale, the proposition that "the risk that the city government becomes aware of property tax

evasion is low" and the proposition that "the consequences of being discovered not paying the property tax are unlikely." Figure A2 in the Online Appendix shows the response distributions by gender. On both items, women place greater weight than men on the high-risk end of the distribution: women are more likely to disagree that detection is unlikely and more likely to disagree that the consequences are unlikely. This is direct survey evidence that women internalize a higher perceived probability of enforcement, consistent with the model's tax-morale and risk-aversion channels lowering p^*.

Perceived tax burden. The same survey asked respondents whether they would be willing to pay a higher property tax in exchange for better municipal services, and whether they agreed that the property tax is too high. Figure A3 in the Online Appendix shows that women are substantially more likely to declare themselves unwilling to pay a higher tax and substantially more likely to view the existing tax as excessive. These responses are not what one would expect from women whose only difference from men were higher tax morale; if anything, higher tax morale should generate more positive views of taxation. The pattern is, however, exactly what we would expect if women face tighter budget constraints: a tax that absorbs a larger share of disposable income is correctly perceived as more burdensome.

Income and stability. The 2011 EAHU shows that male-led households earn approximately 37% more in monthly income than female-led households and are 13 percentage points more likely to receive a steady wage. Figure A4 in the Online Appendix displays the income decile distributions: female-led households are concentrated in the bottom three deciles, while male-led households spread more evenly across the upper half of the distribution. These differences are not unique to Junín or to Argentina; they reflect well-documented gender gaps in earnings and labor-market attachment that have been linked, in turn, to caregiving responsibilities, occupational segregation, and discrimination (Coelho et al., 2024). For our purposes, what matters is that the income gap is large enough, and the income-stability gap pronounced enough, that a property tax bill calculated from property characteristics rather than current income flows will more often bind on female-led households' liquidity.

5.2 How much of the baseline gap could be driven by income?

The cross-sectional baseline gap in compliance between female- and male-led households is approximately 5 percentage points (44% versus 39%). How much of this gap could plausibly be attributed to differences in household income alone, versus to other channels in the gender bundle? A precise decomposition is impossible with our data, since we do not observe individual taxpayer income. We can, however, use the EAHU and the existing literature on the income-elasticity of property tax compliance to provide a qualitative sense of the relative magnitudes.

The EAHU shows a 37% income gap and a meaningful gap in income stability. Brockmeyer et al. (2021) document, for the property tax in Mexico City, that liquidity-constrained taxpayers reduce consumption following a tax increase rather than fail to pay—suggesting that, at least at moderate tax-to-income ratios, income exerts pressure on payment behavior but is not the sole determinant. In the Junín setting, where mortgage credit is essentially unavailable and the tax is more strongly

disconnected from current income, we would expect liquidity to play a larger role. Even so, our reading of the survey evidence is that income differences contribute meaningfully to the baseline gap but do not exhaust it: women's higher perceived enforcement and higher willingness to view the tax as binding both point to channels beyond pure income. The bundle of attributes associated with female-led households (preferences, household structure, income, social norms) contributes jointly to the gap, and our data do not permit a clean decomposition.

This bounding-style reasoning concerns the cross-sectional baseline gap. The treatment effects of the experimental messages are causally identified by random assignment, regardless of which channel dominates the baseline gap, and the treatment-effect patterns themselves (timing margin for women, extensive margin for men, female heterogeneity by tax size) are the diagnostic predictions of the model.

5.3 Limitations

Several limitations of the analysis are worth stating plainly.

First, the municipal taxpayer registry does not record marital status, household composition, dependents, or income at the individual level. We cannot, therefore, separately identify how much of the female–male contrast in baseline compliance or in treatment response is attributable to gender narrowly construed (preferences, risk aversion) versus to correlated structural factors (household structure, income level, income stability, caregiving). This limitation is shared by essentially the entire field-experimental literature on gender and tax compliance (Lin & Slemrod, 2024; Maag et al., 2023; Torgler & Valev, 2010), and we view our contribution as documenting how the bundle of correlated attributes responds differentially to enforcement, rather than as decomposing the bundle into its constituent channels.

Second, the gender variable we construct identifies the registered taxpayer, not the de facto decision-maker in the household. Recent work shows that the choice of which household member is registered as the responsible taxpayer is itself shaped by social norms (Lin et al., 2025), which means that the population we identify as "female taxpayers" may differ systematically from the population of women who own property, in ways that the experimental setting cannot probe.

Third, the survey evidence we use to interpret the experimental findings comes from a different sample than the experimental sample, drawn from the same population but not linked at the individual level. We can use it to characterize the typical female- and male-led household but not to test the model's predictions at the individual level.

Finally, our analysis is based on a single experiment in a single city. The institutional features of the Junín property tax (thin mortgage market, tax based on property characteristics rather than income, no partial payment option) are common in the developing world but not universal. Generalization to other settings, particularly to higher-income economies with different property tax designs, is not warranted without further evidence.

We hope future field experiments on tax compliance will incorporate gender at the design stage and collect richer demographic and household-level covariates, both of which would allow the kind of decomposition we are unable to pursue here.

6 Conclusion

This paper has examined how taxpayers in female- and male-led households respond to enforcement messages in property tax payment in a developing country, where the tax base is poorly correlated with current income flows. Three findings stand out. First, taxpayers in female-led households are more likely to comply at baseline. Second, the deterrence message produces qualitatively different responses across gender: it shifts the timing of payment for women without changing whether they pay overall, while it shifts whether men pay without much affecting when. Third, the female extensive-margin response is concentrated among taxpayers with small tax bills and attenuates as the bill grows; the male response shows no such pattern. Survey evidence indicates that female-led households simultaneously perceive higher enforcement risk, exhibit higher tax morale, and face tighter liquidity constraints, lower and less stable income, and a stronger sense that the tax is excessive. The combination of patterns is consistent with a simple model in which tax morale and risk aversion lower the indifference probability p^* but a binding minimum-consumption constraint produces corner solutions for which enforcement messages are inert.

Throughout the paper we have been careful about what these findings can and cannot identify. The cross-sectional baseline gap between female- and male-led households is descriptive: it captures the joint contribution of preferences, household structure, income, social norms, and other characteristics that differ systematically between the two groups, and our data do not permit a clean decomposition. The treatment effects of the experimental messages, by contrast, are causally identified by random assignment within each gender stratum. The differential response we document (timing margin for women, extensive margin for men, with female heterogeneity by tax size) is therefore a causal statement about how the bundle of attributes associated with the gender of the registered taxpayer shapes the response to enforcement, even if it is not a structural statement about gender as such. We view this framing as consistent with the broader literature on gender and taxation, in which the "gender effect" is conventionally understood as the joint product of intrinsic and structural channels rather than as a parameter purged of all correlated covariates.

The findings have implications for how tax authorities design enforcement in settings where the tax base is weakly correlated with current income. In such settings, blunt enforcement instruments (deterrence letters, fines, escalating penalties) may exacerbate post-tax inequality between groups whose underlying willingness to comply is high but whose liquidity is low. Stronger enforcement is desirable for many reasons, but if it disproportionately reaches taxpayers who are already willing to comply and not the ones who are unwilling, its equity-enhancing role is muted. Complementary tools that recognize the disconnect between the tax base and current income, such as income-contingent payment plans, targeted discounts for low-income households, or installment options that smooth the cash outlay over the year, could preserve the revenue gains of stronger enforcement while addressing the differential burden it places on liquidity-constrained taxpayers. The case for such tools is sharpest in property taxes in developing countries, where mortgage markets are thin and property cannot easily be converted to liquid resources, but the underlying

logic generalizes to any setting where tax liability and current income are imperfectly correlated.

Our analysis points to several directions for future work. First, field experiments that incorporate gender at the design stage, and that collect richer demographic and household-level covariates, would allow the kind of decomposition our data do not support: separating the contribution of preferences, household structure, income, and norms to the gender bundle. Second, evidence from other developing countries with similar property tax designs would help establish the generality of the timing-versus-extensive split and the heterogeneity by tax size. Third, the differential effectiveness of complementary policy tools, income-contingent payment plans, targeted discounts, and automated payment options, is an open empirical question that lends itself naturally to randomized evaluation. Tax authorities should pursue enforcement methods that are, at a minimum, neutral with respect to characteristics correlated with the gender of the registered taxpayer; how to achieve such neutrality without sacrificing horizontal equity in tax design is, in our view, a critical agenda for both researchers and practitioners.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10797-026-09975-3.

Acknowledgements The research was financed by the Inter-American Development Bank (IDB) and the Institutional Capacity Strengthening Fund (ICSF) of the Inter-American Development Bank, funded by the Government of the People's Republic of China. The IDB, ICSF, and the Government of the City of Junín had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The opinions presented herein are those of the authors and thus do not necessarily represent the official position of the institutions to which they belong.

Author contributions Both authors jointly conceptualized and designed the study, as well as conducted the primary analysis and interpretation of the data. The drafting of the manuscript was a collaborative effort, with iterative back-and-forth writing. Both authors reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Acosta-Ormaechea, S., Pienknagura, S., & Pizzinelli, C. (2022). "Tax policy for inclusive growth in Latin America and the Caribbean". *IMF Working Paper*, *228*.

Advani, A., Elming, W., & Shaw, J. (2023). The dynamic effects of tax audits. *The Review of Economics and Statistics, 105*(3), 545–561.

Alm, J. & Malézieux, A. (2021). 40 years of tax evasion games: a meta-analysis. *Experimental Economics24*(3). Publisher: Springer & Economic Science Association, (pp. 699–750).

Alm, J., & Torgler, B. (2006). Culture differences and tax morale in the United States and in Europe. *Journal of Economic Psychology, 27*(2), 224–246.

Alstadsæter, A. & Jacob, M. (2017). Who participates in tax avoidance? Evidence from Swedish micro-data. *Applied Economics*, *49*(28). Publisher: Routledge, (pp. 2779-2796), https://doi.org/10.1080/00036846.2016.1248285,

Antinyan, A. & Asatryan, Z. (2024). Nudging for Tax Compliance: a Meta-Analysis*. *The Economic Journal*, ueae088. https://doi.org/10.1093/ej/ueae088/59573810/ueae088.pdf.

Bazart, C., & Pickhardt, M. (2011). Fighting income tax evasion with positive rewards. *Public Finance Review, 39*(1), 124–149.

Bird, R. M., & Zolt, E. M. (2008). Tax policy in emerging countries. *Environment and Planning C: Government and Policy, 26*(1), 73–86. https://doi.org/10.1068/cav3

Brehm, M. E., & Malkova, O. (2023). The child tax credit over time by family type: Benefit eligibility and poverty. *National Tax Journal, 76*(3), 707–741. https://doi.org/10.1086/725889

Brockmeyer, A., Estefan, A., Ramírez Arras, K., & Suárez Serrato, J. C. (2021). *Taxing Property in Developing Countries: Theory and Evidence from Mexico*. Working Paper 28637. Series: Working Paper Series. National Bureau of Economic Research.

Cabral, A.C.G., Myles, G., & Kotsogiannis, C. (2015). Self-Employment Underreporting in Great Britain: Who and How Much?. *14th journ é es Louis-Andr é G é rard-Varet*.

Castro, L., & Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization, 116*, 65–82.

Charness, G., Eckel, C., Gneezy, U., & Kajackaite, A. (2018). Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. *Journal of Risk and Uncertainty, 56*(1), 1–17.

Chirico, M., Inman, R., Loeffler, C., MacDonald, J., & Sieg, H. (2019). Deterring property tax delinquency in Philadelphia: An experimental evaluation of nudge strategies. *National Tax Journal, 72*(3), 479–506. https://doi.org/10.17310/ntj.2019.3.01

Chirico, M., Inman, R. P., Loeffler, C., MacDonald, J., & Sieg, H. (2016). An experimental evaluation of notification strategies to increase property tax compliance: Free-riding in the city of Brotherly Love. *Tax Policy and the Economy, 30*(1), 129–161. https://doi.org/10.1086/685595

Coelho, M., Davis, A., Klemm, A., & Osorio-Buitron, C. (2024). Gendered taxes: the interaction of tax policy with gender equality. *International Tax and Public Finance, 31*, 1413–1460.

Corbacho, A., Fretes Cibil, V., & Lora, E. (2013). *More than Revenue*. Inter-American Development Bank and Palgrave McMillan

Cyan, M. R., Koumpias, A. M., & Martinez-Vazquez, J. (2016). The determinants of tax morale in Pakistan. *Journal of Asian Economics, 47*, 23–34.

Dwenger, N., Kleven, H., Rasul, I., & Rincke, J. (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *American Economic Journal: Economic Policy, 8*(3), 203–232.

D'Attoma, J., Volintiru, C., & Malézieux, A. (2020). Gender, social value orientation, and tax compliance. *CESifo Economic Studies, 66*(3), 265–284.

D'Attoma, J., Volintiru, C., & Steinmo, S. (2017). Willing to share? Tax compliance and gender in Europe and America. *Research Politics, 4*, 2.

Eisenhauer, J. G., Geide-Stevenson, D., & Ferro, D. L. (2011). Experimental estimates of taxpayer ethics. *Review of Social Economy, 69*(1), 29–53.

Engstrom, P., Nordblom, K., Ohlsson, H., & Persson, A. (2015). Tax compliance and loss aversion. *American Economic Journal: Economic Policy, 7*(4), 132–164.

Finocchiaro Castro, M., & Rizzo, I. (2014). Tax compliance under horizontal and vertical equity conditions: An experimental approach. *International Tax and Public Finance, 21*(4), 560–577.

Fortin, B., Lacroix, G., & Villeval, M.-C. (2007). Tax evasion and social interactions. *Journal of Public Economics, 91*(11), 2089–2112.

Grown, C., & Valodia, I. (2010). *Taxation and Gender Equity*. Routledge.

Hibbert, A. M., Lawrence, E. R., & Prakash, A. J. (2013). Does knowledge of finance mitigate the gender difference in financial risk-aversion? *Global Finance Journal, 24*(2), 140–152.

Hofmann, E., Voracek, M., Bock, C., & Kirchler, E. (2017). Tax compliance across sociodemographic categories: Meta-analyses of survey studies in 111 countries. *Journal of Economic Psychology, 62*, 63–71.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica, 79*(3), 651–692.

Kogler, C., Mittone, L., & Kirchler, E. (2016). Delayed feedback on tax audits affects compliance and fairness perceptions. *Journal of Economic Behavior and Organization, 124*, 81–87.

Lin, E. Y., & Slemrod, J. (2024). Gender tax difference in the U.S. income tax. *International Tax and Public Finance, 31*, 808–840.
Lin, E. Y., Slemrod, J., Smith, E., & Yuskavage, A. (2025). Who's on (the 1040) first? Determinants and consequences of spouses' name order on joint returns. *International Tax and Public Finance, 32*, 957–994.
Maag, E., Matsui, A., & Menefee, K. (2023). How refundable tax credits can advance gender and racial equity. *National Tax Journal, 76*(3), 743–763.
Mascagni, G. (2018). From the lab to the field: A review of tax experiments. *Journal of Economic Surveys, 32*(2), 273–301. https://doi.org/10.1111/joes.12201
Shafiq, M. N. (2015). Aspects of moral change in India, 1990–2006: Evidence from public attitudes toward tax evasion and bribery. *World Development, 68*, 136–148.
Skatun, J. D. (2017). Taxation, risk aversion, and the wage gaps in tournaments. *Oxford Economic Papers, 69*(3), 834–845.
Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature, 57*(4), 904–54.
Torgler, B. (2005). Tax morale in Latin America. *Public Choice, 122*(1/2), 133–157.
Torgler, B., & Valev, N. T. (2010). Gender and public attitudes toward corruption and tax evasion. *Contemporary Economic Policy, 28*(4), 554–568.
Wenzel, M. (2006). A letter from the tax office: Compliance effects of informational and interpersonal justice. *Social Justice Research, 19*(3), 345–364.

Authors and Affiliations

Andrea Lopez-Luzuriaga[1] · Carlos Scartascini[1]

✉ Andrea Lopez-Luzuriaga
andrealo@iadb.org

Carlos Scartascini
carlossc@iadb.org

[1] Inter-American Development Bank, 1300 New York Ave NW, Washington, DC 20577, USA