Fiscal Capacity, Distributional Consequences, and Preferences over Taxation*

Jessica Gottlieb †
& Florian M. Hollenbach‡
& Thiago M.Q. Moreira§

First Draft: August, 2021
This Draft: June 27, 2025

Abstract

Most theories of taxation and redistribution assume that states are efficient in collecting taxes. Yet, we know they vary tremendously in their ability to enforce tax policies with distributional consequences. In this paper, we evaluate whether and how taxpayer perceptions of fiscal capacity affect beliefs and attitudes about taxation. Through three survey experiments conducted in the United States, we demonstrate that citizens update beliefs in response to information about asymmetric tax enforcement with effects on willingness to pay tax. Countervailing effects manifest through two mechanisms – fairness and material concerns. Our results imply that informing taxpayers about revenue losses from asymmetric compliance is a potential strategy to address, through tax morale, fiscal capacity weaknesses often at the root of asymmetric enforcement.

*Authors listed in alphabetical order. Equal authorship is implied. We thank seminar participants at Copenhagen Business School and University of Gothenburg for helpful comments. All remaining errors are our own. Pre-analysis plan available at https://osf.io/f3qrz/. The survey experiments received IRB approval at Texas A&M University (IRB2019-1549M) and the University of Houston (STUDY00003428).

†Associate Professor, Hobby School of Public Affairs, University of Houston. Email: jagottli@central.uh.edu. URL: https://sites.google.com/site/jessicaagottlieb/

‡Associate Professor, Department of International Economics, Government and Business, Copenhagen Business School, Porcelænshaven 24A, 2000 Frederiksberg, Denmark. Email: fho.egb@cbs.dk. URL: fhollenbach.org

§Assistant Professor, Department of Political Science, Louisiana State University. Email: tmoreira@lsu.edu. URL: thiagomqmoreira.com

The majority of existing work on citizen preferences over taxation and redistribution assumes the state to be a capable and equitable enforcer of tax policies. Yet, states vary tremendously in their ability to enforce tax policies and collect taxes (Bird and Zolt, 2004; Fjeldstadt and Moore, 2008; Gordon and Li, 2005; Besley and Persson, 2014). Similarly, taxpayers differ significantly in their ability to evade taxes. High-income earners often use sophisticated financial arrangements and expensive lawyers to hide income, such that tax evasion is not detectable with standard audits (Guyton et al., 2021), making tax enforcement substantially more complex and expensive. As a result, many resource-constrained fiscal administrations focus their limited enforcement capacity on lower and middle income taxpayers, leading to inequality in enforcement.

The ability of agencies to enforce tax laws thus varies both across countries as well as within countries across taxpayers. In the United States, the empirical context of this study, capacity of the Internal Revenue Service (IRS) and overall audit rates have long been in decline, with audit probabilities for those with income above $200,000 decreasing the most (U.S. Government Accountability Office, 2022; Frank, 2022).[1] We argue that such differential enforcement has important downstream consequences through multiple channels.

First, strength of enforcement (or probability of detection) affects taxpayer compliance (Slemrod, 2019). More unequal enforcement should therefore reduce revenue collected from high-income earners relative to the rest of the taxpayer distribution, undermining the de facto progressivity of collected revenues. In the US, for example, researchers estimate that between 2006 and 2013, the top 1% failed to report around 21% of their income to the IRS (Guyton et al., 2021). And while all governments struggle to enforce compliance among rich taxpayers, such discrepancies in enforcement across the income spectrum are even larger in states with weak capacity (Bachas, Jensen, and Gadenne, 2024).[2] Moreover, low enforcement capacity constrains both the state's revenue-maximizing tax rate (Bergeron, Tourek, and Weigel, 2024) as well as the progressivity of the tax schedule (Rubolino, 2023).

Second, asymmetric enforcement not only influences the distribution of voluntary compliance across income groups, but it also affects *beliefs* about compliance of these groups. As we argue

[1] Under Biden's Inflation Reduction Act (IRA), signed into law in August 2022, the IRS budget for enforcement of high income and wealthy tax payers was meant to increase significantly over the next ten years, a development that President Trump and Republicans in Congress have promised to reverse.

[2] Cobham and Janskỳ (2018) find that revenue losses from international corporate avoidance and evasion comprise, on average, more than 2% of low-income countries' GDP, compared to less than 0.5% in OECD countries. As a result of weak enforcement capacity, low-income countries predominantly raise revenues via indirect taxes on consumption and personal income taxes that target formal workers (Benedek, Benítez, and Vellutini, 2022; Choudhary, Ruch, and Skrok, 2024).

and show, beliefs about enforcement and compliance undermine taxpayers' willingness to pay. Asymmetric non-compliance by the rich not only makes it difficult to collect revenue that could be reinvested into the state's enforcement arm to address the problem of non-compliance, but it also has the potential to decrease fiscal revenue from other taxpayers due to fairness concerns.

Despite its importance, asymmetric enforcement and compliance has received little attention in the literature, in particular on *how* it may change voters' views about taxation. In this manuscript, we present the results from a series of three pre-registered survey experiments. We seek to understand how asymmetric enforcement affects beliefs about compliance and, in turn, tax morale. To gain traction on our question, we test how informing voters about the extent of asymmetric fiscal enforcement and its consequences for public revenue affects fairness concerns and willingness to pay tax. Our innovation relies on the idea that most taxpayers have beliefs about the state's tax capacity, but are relatively ill-informed about the distribution of non-compliance over the income spectrum. We first show that providing such information to taxpayers causes them to update their beliefs about fiscal capacity and perceived tax compliance. We then show that this updating has complicated effects on tax morale: it causes taxpayers to be less willing to pay tax for public goods but more willing to pay tax to boost enforcement.

In line with expectations from the existing literature, our first study shows that informing taxpayers about true asymmetric compliance rates decreases their perceived fairness of the tax system and willingness to pay for public goods. In Study 2, we exploit a countervailing sentiment – that asymmetric compliance is materially bad for most taxpayers – and prime the idea that increasing individual compliance can improve the state's ability to enforce upon the rich. When we prime material loss (rather than fairness) *and* link increased tax revenues to increasing enforcement capacity, we find an increase in willingness to pay. This finding suggests that informing taxpayers about revenue losses from asymmetric compliance is a potential strategy to improve, through tax morale, weaknesses in fiscal capacity that are at the heart of asymmetric enforcement.

Finally, in a third study, we exploit the recent passage of the Inflation Reduction Act (IRA) that established an unprecedented increase in fiscal capacity, with a focus on higher-income earners. We test whether priming the salience of this credible commitment to expand the public budget through increased enforcement on the rich can undo some of the depressive effect on tax morale of informing citizens about the unfairness of asymmetric non-compliance by the rich. We do not find support for this conjecture. Rather, we find that respondents in our third study had,

on average, higher levels of tax morale than in Study 2 and less negative responses to the same fairness prime than in Study 2. We speculate about why in the discussion section.

Our paper makes three important contributions to the literature on tax morale and taxation. First, we show that citizens' perceptions of the state's tax capacity vary tremendously and are malleable. Much of the existing literature applies a unitary conception of fiscal and state capacity – measured at the level of the state and applied uniformly within the polity.[3] One consequence of this conceptualization of state capacity is that it appears stable and hard to change (Besley and Persson, 2011; de Mesquita, Downs, and Smith, 2017). In contrast, we show that citizens differ in their assessment of the state and that perceptions about the state's tax capacity can change in response to new information, with potentially consequential impacts on tax attitudes and behaviors. Our second contribution is thus showing how perceptions of the state's (asymmetric) capacity to enforce the rules of taxation directly influence beliefs about tax compliance, and do so differently for different income groups, impacting views about the fairness of tax systems. And third, we provide evidence of a novel determinant of citizen willingness to pay tax – perceptions about the state's ability to collect.

1 Theoretical Framework

When taxpayers perceive low compliance among peers, it can reduce their own willingness to pay tax (Frey and Torgler, 2007; Castro and Scartascini, 2015; Alm, Bloomquist, and McKee, 2017).[4] This behavior is characteristic of collective action problems where one person's deviation from a cooperative outcome can trigger defection among others (Olson, 1965). However, not all taxpayers are equal: some have a much higher ability to contribute than others. Moreover, the likelihood that non-compliance is detected differs significantly across different groups of taxpayers (Guyton et al., 2021). We may, therefore, expect the average taxpayer to react differently to learning about non-compliance by the rich and poor: non-compliance by these groups ought to have different effects on the willingness to pay tax. In particular, if the average taxpayer discovers that the rich are paying at lower rates than they thought, they may have a relatively more negative reaction than finding out the same about the poor.

We propose two mechanisms through which these distributional concerns can affect willingness to pay: one normative and one instrumental. Let's assume the compliance asymmetry is in

[3] See Luna and Soifer (2017) for an important exception.

[4] Many have found the reverse to be true as well: increased perceptions of compliance increases tax morale (Frey and Torgler, 2007; Del Carpio, 2013).

the direction of the rich complying at lower rates than the poor. Then, there are two principal reasons a middle-class taxpayer might be especially concerned. First, are *fairness concerns*. The rich have a higher capacity to pay, i.e., a greater ability to bear the costs of payment, so their under-compliance can seem like an affront to the fairness of the social contract. Second, are *material concerns*. Contributions of the rich have a greater impact on the public budget, such that the marginal negative effect on potential public goods provision of non-compliance of a single rich taxpayer is much greater than that of a single poor taxpayer. In a nutshell, non-compliance by the rich is materially costlier to society (and to any individual taxpayer) than non-compliance by the poor.

We argue that the first mechanism – *fairness concerns* – will negatively affect willingness to pay tax. The fact that people who are wealthier than you are also less likely to pay tax only serves to exacerbate the incentive to defect. In addition to not wanting to be a sucker for paying when others are not, asymmetric non-compliance increases the tax system's perceived unfairness, further demotivating the average taxpayer. Indeed, perceptions of an unfair tax structure have been shown to reduce tax morale (Castañeda, 2024). Moreover, increasing compliance among the rich has a more positive effect on tax morale than increasing compliance among the poor (Campos-Vazquez et al., 2023).

By contrast, we argue that the second mechanism – *material concerns* – may have a countervailing effect on willingness to pay. If non-compliance by the rich has large negative material consequences for the average taxpayer, then one solution is to invest in enforcement capacity. In other words, taxpayers might be more willing to pay taxes if they think the additional revenue could be invested in the fiscal state, addressing the asymmetric enforcement problem. In that case, the marginal return to paying higher taxes would be significantly higher, as it raises revenues through better enforcement of high-income earners. For example, recent estimates suggest that "an additional Dollar spent auditing taxpayers above the 90th income percentile yields more than $12 in revenue," significantly more than audits of average taxpayers (Boning et al., 2025, 63). In line with this argument are existing findings that better enforcement improves compliance both directly by reducing evasion and indirectly by increasing tax morale (Filippin, Fiorio, and Viviano, 2013). Similarly, highlighting the material benefits of tax revenue to taxpayers increases tax morale, support for enforcement, and perceptions of peer compliance (Wilson and Rosenzweig, 2024).

Taxpayers are likely to form beliefs about compliance and enforcement based on their own

experience – and that of their peers – of paying tax or being audited. As a result, people may have stronger beliefs about compliance and enforcement in their own income group, with whom they have greater contact and interaction. In contrast, taxpayers are likely to have more uncertain beliefs about the *distribution* of compliance and enforcement across the income spectrum. Thus, taxpayers are likely to be relatively uninformed or uncertain about the asymmetry of compliance and enforcement in their tax system.

In a context of limited information about true taxpayer compliance and high uncertainty for citizens, it is likely that beliefs are malleable and citizens update their priors in response to new information (Bullock et al., 2015).[5] We, therefore, expect that providing information about the true enforcement capacity of the state, and the extent to which it is applied asymmetrically across the income spectrum, should cause people to substantially update their beliefs about compliance.

1.1 Empirical predictions

We test these theoretical expectations in a suite of three studies among similar samples of potential taxpayers in the United States. Because many of our inquiries are understudied, we sequenced our studies incrementally, designing each new study based on the insights of previous ones. We pre-registered our priors and design in a single design document that we updated with each new study, e.g., we pre-registered Study 2 only after completing Study 1.

In Study 1, we first interrogate two assumptions upon which our theory rests: that new information about true levels of tax enforcement will cause the average citizen to update their beliefs about 1) the state's ability to collect taxes; and 2) fellow citizen's compliance.

Hypothesis 1 *Informing citizens about true levels of tax enforcement will reduce perceptions of fiscal capacity.*

Hypothesis 2 *Informing citizens about the asymmetry of tax enforcement across rich and poor taxpayers will increase the asymmetry of perceived compliance rates across the rich and poor relative to simply informing them about low average enforcement rates.*

We then examine downstream effects of new information, testing Hypothesis 3 in Study 1, Hypotheses 4 - 5 in Study 2, and Hypotheses 6 - 7 in Study 3.

[5] In a similar context, Fernández-Albertos and Kuo (2015) find that respondents are often incorrect when asked about their placement in the income distribution and correcting these beliefs has direct effects on preferences over tax progressivity.

HYPOTHESIS 3 *Informing citizens about true levels of tax enforcement will reduce willingness to pay tax, mediated in part through the effect on perceptions of tax compliance.*

HYPOTHESIS 4 *Priming enforcement asymmetry and its effect on fairness will decrease willingness to pay for public goods and enforcement.*[6]

HYPOTHESIS 5 *Priming enforcement asymmetry and its effect on public revenue will increase willingness to pay for public goods and enforcement.*

HYPOTHESIS 6 *The negative effect on tax morale of priming the unfairness of asymmetric enforcement can be offset (or even reversed) by informing taxpayers about policies adopted to enforce compliance.*[7]

HYPOTHESIS 7 *The effect of information about these policies is stronger when accompanied by specific details about their mandate to address asymmetric enforcement by targeting compliance of the rich.*

2 Data and Research Design: Studies 1 and 2

We first report the results from Studies 1 & 2, fielded in May 2021 and January 2022. Both studies were fielded on the same platform among the same type of respondents but employed different treatment variants.[8] The treatment conditions we test in Study 2 were designed to answer questions raised by the results in Study 1. The questions used to measure outcomes are largely consistent across the studies.

2.1 Treatments

Studies 1 and 2 both aimed to manipulate perceptions of fiscal capacity by randomly assigning informational treatments (and a placebo control) to survey respondents.[9] Study 1 examined the effects of priming different features of low capacity: 1) general weakness of tax collection (*Low Enforcement*); 2) uneven enforcement across income groups (*Asymmetric Enforcement*), with the wealthy being enforced upon less. Motivated by the reality of the US context, we focused on the asymmetric enforcement benefiting the rich in Study 2, but primed two distinct consequences of

[6] In our PAP, we expected the fairness and revenue primes to decrease willingness to pay for both public goods and enforcement, but were ambivalent about which would have a larger effect on each of these outcomes.

[7] In Study 3, we inform US citizens about the Inflation Reduction Act of 2022.

[8] The provider ensured that in Studies 2 & 3 none of the respondents had participated in one of the previous studies.

[9] We stratified random assignment income quartiles using data from DQYDJ estimated United States household income brackets with the following cutoffs from our survey responses: 1) under $30,000; 2) $30,000 to $59,000; 3) $60,000 to $119,999; and 4) above $120,000.

such uneven enforcement: 1) unfairness, and 2) lost revenue. Table 1 summarizes the treatment variants under each study; both use the same placebo control condition. We provide the full text of the treatments in Appendix 7.[10]

Table 1: Description of Treatments

	Study 1: Dimensions of Low Capacity	Study 2: Effect of Asymmetric Enforcement
Treatment Variants	Low enforcement	Fairness concerns (rich are less likely to pay)
	Asymmetric enforcement (relatively weaker among rich)	Losses concerns (more revenue lost when rich don't comply)

The treatments in Study 1 were created by compiling true statements from national news sources that prime the different aspects of the state's effort to collect taxes. The *Low Enforcement* treatment provides information about the declining rate of tax auditing and the amount of money that is lost as a result. The *Asymmetric Enforcement* treatment provides information about the fact that the declining audit rate is especially attributable to a decline in audits among the highest earners.

In contrast to the first study, both informational treatments in Study 2 prime respondents to think about the differential effects of asymmetric enforcement capacity by the IRS, where high-income earners are enforced upon relatively less than low-income earners. But they each highlight a different aspect of asymmetric enforcement to better understand not only whether, but why respondents may modulate tax morale in response. The *Fairness* treatment primes the fact that asymmetric enforcement is unfair by illustrating how the estimated share of unpaid taxes increases with taxpayer income. The *Losses* treatment informs respondents about the lost revenue caused by asymmetric enforcement, documenting the incredibly large estimated amount of lost revenue due to owed taxes by the rich. The treatments are created using excerpts of statements from the U.S. Treasury about the difference in owed and paid taxes across income groups.

To simulate the time respondents spend reading the treatment text but without the attendant effect on tax beliefs, we include a placebo text for the control group. The text describes the growth in internet usage in recent years among Americans – information that is meant to be relevant for the internet survey population but is not expected to affect mood or attitudes related to our outcomes of interest.

[10] Table G.1 and G.2 show the balance across relevant covariates for the different treatment groups in Studies 1 and 2, respectively.

2.2 Sample

We recruited and surveyed a representative sample of the US population through Lucid Marketplace, restricting eligibility to US citizens 18 and older who can vote and pay taxes. We employ two attention checks (Aronow et al., 2020): a pre-screening question and another administered just prior to treatment. As pre-specified, we drop all respondents who failed the in-survey attention check and end up with 1422 and 1556 respondents in Studies 1 and 2, respectively.

2.3 Descriptive Statistics & Manipulation Checks

Our main outcome of interest in these studies is tax morale. But our main driver of changes in tax morale is citizens' beliefs about capacity and the resulting differential enforcement. Any effect on a respondent's willingness to pay tax presupposes that our treatments actually affected respondent perceptions of fiscal capacity. Citizens differ significantly in their perceptions of the state's capacity to enforce taxes. To measure perceived capacity, we ask respondents to rate the ability of the IRS to collect taxes and enforce tax laws (on a scale of 0 to 100). Across the first two studies, the average responses are quite similar, with the average perceived capacity being 60.5 and 59.8, respectively. Nevertheless, responses vary quite significantly, with inter-quartile ranges of 40 and 43 and standard deviations of 26.6 and 26.8, respectively.[11]

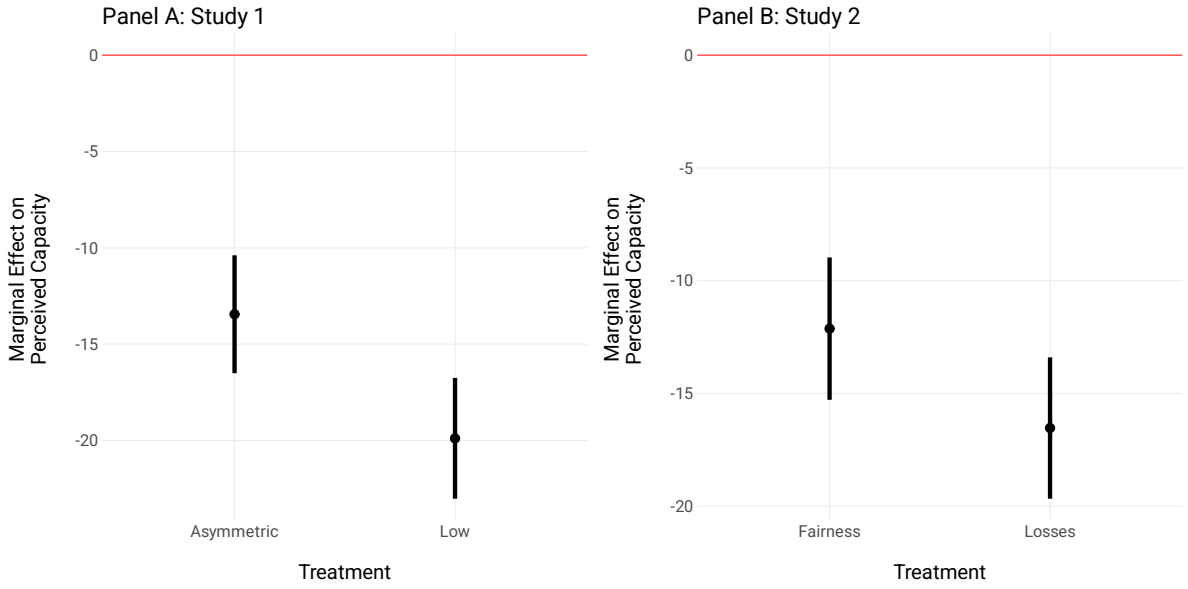
Turning to our treatments of interest, as a manipulation check and test of Hypothesis 1, we first examine whether the treatments in each study lowered perceptions of the ability of the IRS to collect taxes and enforce tax laws. As Figure 1 shows, in both studies, all treatments priming aspects of low capacity of the IRS to collect taxes significantly decreased perceived capacity relative to the placebo control condition. The marginal effects are based on simple regression models with the treatment indicators as main independent variables. We control for income quartiles (blocking variables) and calculate robust standard errors (HC2). The full results are presented in Table G.5 in the Appendix. In Study 1, shown in Panel A, the *Low* and *Asymmetric Enforcement* treatments lower respondents' perceived capacity by 19.9 and 13.5 points, respectively. These effects are quite large, corresponding to 0.74 and 0.5 standard deviations of the dependent variable.[12] In Study 2, the *Fairness* and *Losses* treatments lead respondents to lower their capacity beliefs by 12.1 (0.45 SD) and 16.5 (0.61 SD) points, respectively (Panel B in Figure

[11] In Table G.4 in the Appendix, we show the results from simple linear regression models with perceived capacity as the dependent variable and respondent characteristics included as independent variables. The most robust results is that respondents who had friends or family who experienced audits are substantially more positive about IRS capacity.

[12] This tests Hypothesis 1 from our PAP.

1). The effects of the treatments in both studies are substantively large, bigger than any of the estimated coefficients in the descriptive regressions.

Figure 1: Effect of Treatments on Perceived Capacity

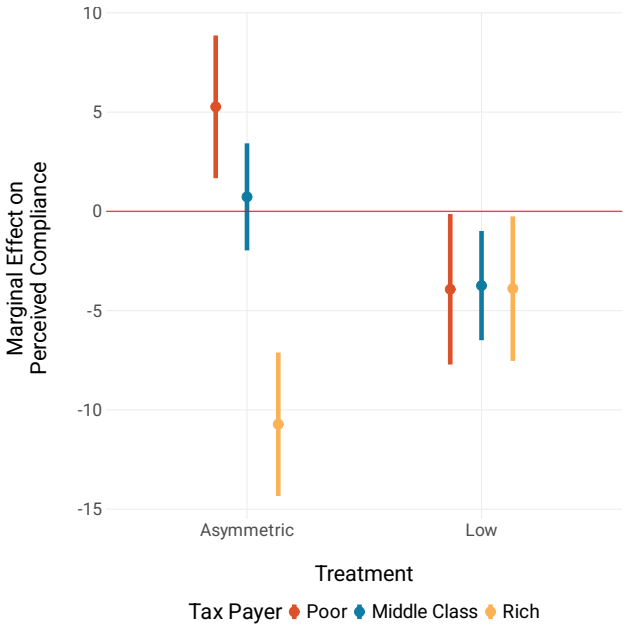


Notes: Coefficient plots are produced by jointly estimating the effect of the two unique treatment arms in each study, relative to the placebo control. The outcome measures responses to the following question: "If the federal government wants to collect income tax revenue, on a scale from 0 to 100 how do you rate the ability of the IRS to collect taxes and enforce tax laws in the United States?" Estimates are from simple linear regression models with treatment group indicators and controlling for income quartiles (blocking variable) with robust standard errors (HC2).

Interestingly, in Study 1, the *Low Enforcement* treatment reduced capacity perceptions by about 50 percent more than the Asymmetric treatment which primed respondents to think about the fact that some taxpayers are more affected by low IRS capacity than others. To help explain this differential effect on perceptions across the two treatment arms, we examine effects on respondent perceptions of tax compliance.[13] Consistent with our expectation in Hypothesis 2, the *Asymmetric* treatment leads to more asymmetric perceptions of compliance than the *Low Enforcement* treatment.[14] Figure 2 illustrates that the *Low Enforcement* treatment reduced respondent perceptions of tax compliance for all target income groups equally. The *Asymmetric* treatment increased perceptions of compliance among the poor and decreased perceptions of compliance among the rich, which was expected given that the *Asymmetric Enforcement* treatment particularly primed the differential audit rates for poor and rich. Full results for the underlying regression models are shown in Table G.6 in the Appendix.

[13] This outcome was only measured post-treatment in Study 1.
[14] H2c from our PAP.

Figure 2: Effect of Treatments on Perceived Compliance



Notes: Coefficient plots are produced by jointly estimating the effect of the two unique treatment arms in each study, relative to the placebo control for three separate outcome measures. Each outcome measures responses to the following question (separately for each income group): "Below we show three example households with different monthly incomes (in the 10th, 50th, and 90th percentiles). In your view, what percentage of US residents with those household incomes fully comply with paying taxes?" Regression models control for income quartiles, and HC2 standard errors are calculated.

Together, these results demonstrate that taxpayers are uncertain about fiscal capacity and tax compliance and that they differ in their beliefs. Importantly, their perceptions of these phenomena are malleable. Our treatments in Surveys 1 and 2 adequately inform respondents about the existing enforcement conditions at the IRS to update these perceptions among taxpayers.

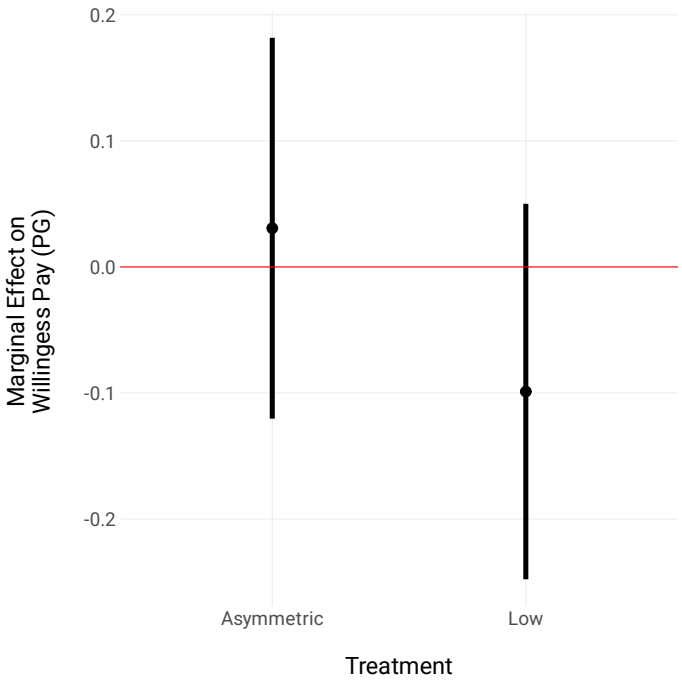
3 Results from Studies 1 and 2

As stated in Hypothesis 3, we expect that both the *Low* and *Asymmetric Enforcement* treatments would reduce willingness to pay tax, mediated through the effect on compliance beliefs.[15] Figure 3 shows the marginal effects for the continuous willingness to pay measure. The full model results for the binary and continuous measures are presented in Table G.7 in the Appendix. As shown in the plot, we do not find a direct effect of either enforcement treatment on willingness to pay tax. In fact, somewhat surprisingly, the *Low Enforcement* treatment has a slightly negative but insignificant effect on willingness to pay tax for public goods, while the *Asymmetric* treatment has a slightly positive but insignificant effect. The difference between the two treatment effects is

[15] H3 in our PAP.

statistically significant at the 0.1 level.

Figure 3: Marginal Treatment Effects on Willingness to Pay Tax - Study 1



Notes: The plot shows the marginal effect of the two treatment arms on the binary measure of respondents' willingness to pay tax for public goods provision. We jointly estimate the effect of the two unique treatment arms, relative to the placebo control, and report marginal effects. We asked respondents: "Let's assume the federal government develops a new program to provide important public goods (for example, infrastructure, health care, or education) to US residents and citizens. They are issuing a referendum on whether to undertake a 1% across-the-board income tax rate increase to fund the provision of the public good. How likely are you to vote for this tax increase?" The *numeric* measure uses the raw 4-point Likert scale; the *binary* measure codes both "unlikely" responses as 0 and both "likely" responses as 1. Estimates are based on OLS regressions with controls for income quartiles and robust standard errors (HC2).

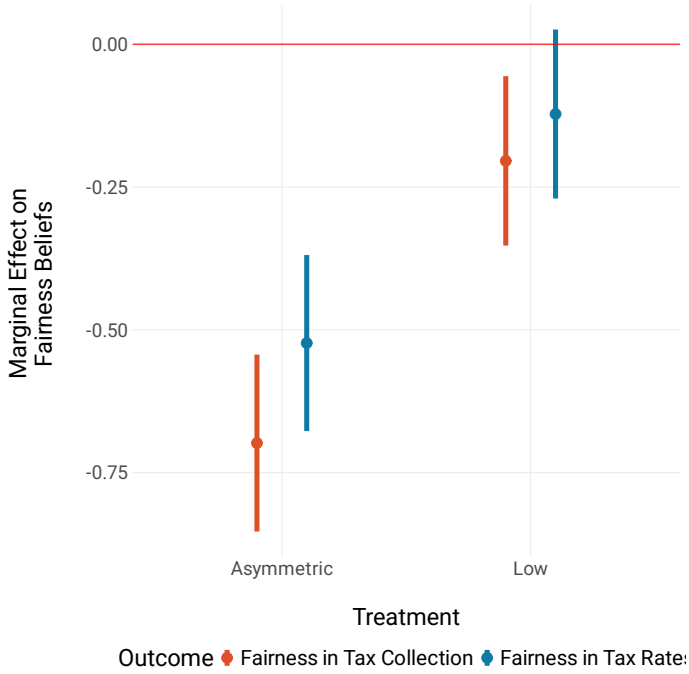
While there is no clear direct effect of either treatment, we do find some evidence of a mediated effect. As we show in Table G.8 in the Appendix, consistent with the uniform decrease in compliance beliefs reported in Figure 2, the *Low Enforcement* treatment has a negative effect on willingness to pay mediated by compliance beliefs of the rich and poor. And consistent with the reported asymmetric effect on compliance beliefs, the *Asymmetric Enforcement* treatment has a negative effect on willingness to pay only when mediated by the compliance beliefs about the rich. This evidence should be interpreted with caution, as we cannot assume sequential ignorability, e.g., among respondents in the treatment group, those with high compliance beliefs are not proper counterfactuals for those with low compliance beliefs.

Revisiting the effects on compliance perceptions from Study 1 in Figure 2, we can see that the *Asymmetric Enforcement* treatment induces respondents to believe that tax compliance across

income groups is more unequal. This could theoretically have countervailing effects on their tax morale. On the one hand, it could increase their perception of the unfairness of tax collection (i.e., *fairness concerns*), making them less likely to voluntarily comply with a system they see as unjust. On the other hand, it could highlight the potential lost revenue from the wealthy who do not pay (i.e., *material concerns*), making them more likely to pay into the system *if* doing so could correct this inequality.

We see some evidence for fairness concerns in Study 1. Consistent with expectations, both treatment arms decrease perceptions of the fairness of the tax system (see Figure 4).[16] The effect is significantly more pronounced in the *Asymmetric* treatment arm relative to the *Low Enforcement* arm. Indeed, as suggested by the mediation analysis in Table 2, just about all of the negative effect of the *Asymmetric Enforcement* treatment on tax morale is working through the negative effect on beliefs about the fairness of tax collection and tax rates.

Figure 4: Marginal Treatment Effects on Fairness Beliefs - Study 1



Notes: We jointly estimate the effect of the two unique treatment arms, relative to the placebo control, for the two outcome measures and report marginal effects. On a 5-point Likert scale, we separately asked respondents to rate the fairness of *tax collection* and federal income *tax rates* in the US. Both treatments lower respondents' beliefs about fairness in collection as well as tax rates, though the marginal effect of the *low* treatment on tax rate fairness is not statistically significant. The marginal effects of the *Asymmetric* treatment are substantially larger.

[16] H6 from our PAP. These estimates are based on OLS regressions with controls for income quartiles and robust standard errors (HC2) and the full results are presented in Table G.9 in the Appendix.

Table 2: Willingness to Pay for PG – Mediated by Fairness Beliefs – Study 1

Outcome:	Willingness To Pay Tax for PG			
	Treatment:			
	Low Enforcement		Asymmetric Enforcement	
	Mediator: Fairness of			
	Tax Collection	Tax Rates	Tax Collection	Tax Rates
ACME	-0.044*	-0.032	-0.122***	-0.095***
	[-0.087, -0.010]	[-0.068, 0.007]	[-0.172, -0.071]	[-0.147, -0.057]
ADE	-0.071	-0.078	0.154*	0.125+
	[-0.208, 0.050]	[-0.216, 0.074]	[0.008, 0.307]	[-0.014, 0.273]
Total Effect	-0.115+	-0.110	0.032	0.030
	[-0.257, 0.018]	[-0.255, 0.050]	[-0.119, 0.181]	[-0.119, 0.177]
Prop. Mediated	0.374+	0.262	-1.013	-0.915
	[-0.750, 3.227]	[-0.783, 1.839]	[-18.979, 19.341]	[-19.121, 12.391]

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the mediation package in R. Robust standard errors are calculated.

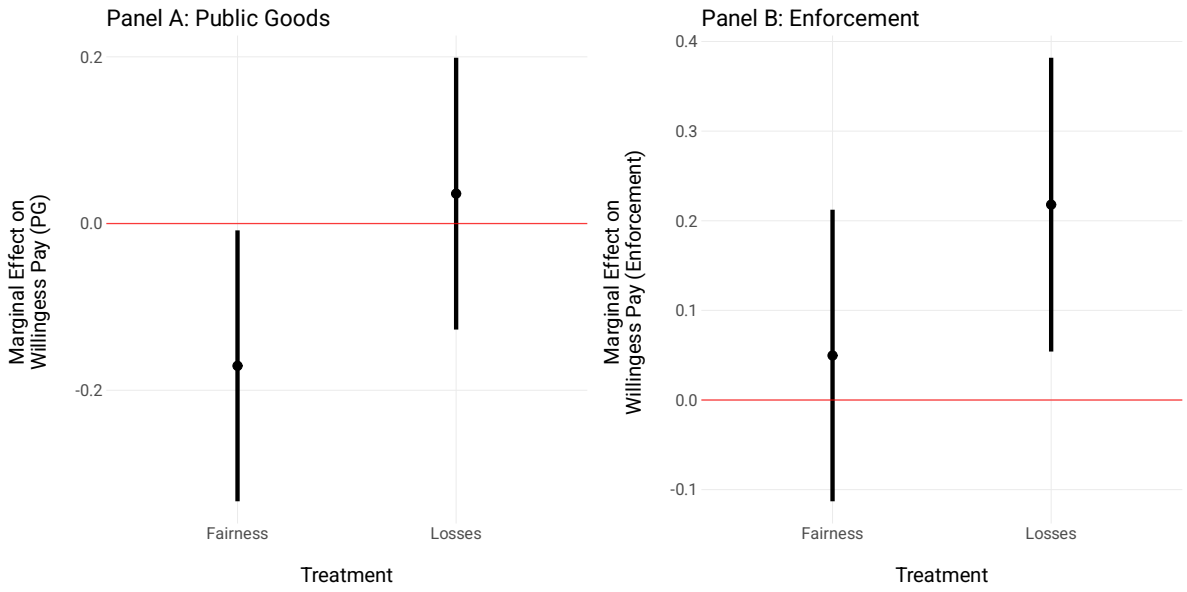
If the *Asymmetric Enforcement* treatment simultaneously primed these countervailing mechanisms – unfairness of the tax system and a missed opportunity to collect revenue – it could mar our ability to detect effects on tax morale in a single direction. Study 2 was thus designed to disentangle these mechanisms. Not only do we prime these different effects of asymmetric enforcement through our two treatment conditions – *Fairness* and *Losses* – but we also devised two different outcome measures to assess whether respondents would be more willing to pay tax if the revenue was spent differently. Specifically, the first measure asks whether respondents would be in favor of a small increase in the federal income tax rate if additional revenue would go directly to pay for better education of American children. The second question earmarks additional revenue to pay for better IRS enforcement capabilities and improved auditing rates, projected to significantly increase compliance of high-income taxpayers.

The main results of Study 2, summarized in Figure 5, provide evidence of these countervailing forces. As shown in Panel A, the *Fairness* treatment has a negative effect on respondent willingness to pay a new tax to finance public goods (relative to *Control* and relative to the *Losses* treatment arm), supporting Hypothesis 4.[17] By contrast, and in line with Hypothesis 5, Panel

[17]The continuous measures of willingness to pay are a five-point Likert scale, the binary measure codes both statements in agreement as 1 and statements of indifference or disagreement as 0. Unfortunately, the results from Study 1 on willingness to pay for public goods are not entirely comparable since respondents were not given an option to voice indifference.

B shows that the *Losses* treatment has a positive effect on respondent willingness to pay a new tax to improve IRS enforcement (relative to *Control* and relative to the *Fairness* treatment arm).[18] The latter is consistent with another recent study finding that providing U.S. respondents with information about the underfunding of the IRS and its consequences for public revenue caused an increase in support for funding the IRS (Anson and Kane, 2023).

Figure 5: Marginal Treatment Effects on Willingness to Pay - Study 2



Notes: This figure shows the marginal effect of the *Fairness* and *Losses* treatments relative to the placebo control for willingness to pay tax for public goods (Panel A) and increased tax enforcement (Panel B). The *Losses* treatment increases respondents' willingness to pay for enforcement, whereas the fairness treatment decreases respondents' willingness to pay for public goods. Estimates are from simple linear regression models with treatment group indicators and controlling for income quartiles (blocking variable) with robust standard errors (HC2). Full results are presented in Table G.10 in the Appendix.

We expected that one prime would be dominant and have a stronger effect on each of the outcomes (Hypotheses 4 and 5).[19] Instead, we find that each prime affects only one of the outcomes and in only one of the expected directions. For policymakers seeking to improve tax morale, this is good news. It indicates that messaging should avoid mention of fairness, which could have a depressive effect on tax morale, and instead highlight the revenue benefits of increased enforcement, which can have a positive effect on morale.

[18] Hypotheses H11-14 in our PAP.
[19] H9 in our PAP.

4 Perceptions of Capacity and Support for Increasing Enforcement: Study 3

In the remainder of the paper, we will discuss the results from Study 3, fielded in December 2022, about ten months after Study 2. This survey was designed to gain additional insights into the mechanisms driving previous results. Specifically, we wanted to better understand how fairness concerns and their interplay with state investments in enforcement capacity affect willingness to pay and preferences over taxation. The second half of 2022 in the US provided a unique opportunity to study such questions, as Congress had passed and President Biden had just signed the Inflation Reduction Act (IRA) in the summer. This bill included an unprecedented increase in the IRS budget, initially US$80 billion over the next ten years. A majority of the additional money was allocated to improve enforcement, in particular, on high-income taxpayers (Tax Policy Center, 2024). During the debate over the bill and its passage, both Republicans and Democrats publicly spoke about the bill, and news organizations reported widely about the policy. The passage of the bill allowed us to use information about an actual, and thus realistic, increase in enforcement capacity as treatment. The questions used to measure outcomes are largely consistent with the previous two studies.

4.1 Treatments

With the third survey experiment, we aimed to investigate the interaction between the fairness concerns and investments in tax enforcement. To do so, we use a fully crossed 2x3 factorial design. The first factor consists of two levels, one in which respondents were randomly exposed to the informational treatment about unequal enforcement, priming either fairness concerns from Study 2 or a placebo condition providing information about the increase in mobile internet use. After this set of initial treatments is administered, we asked respondents about capacity beliefs and compliance rates. Next, respondents were randomly exposed to the second factor consisting of three levels: a) information about the increase in the IRS budget and the planned increase in enforcement, b) the same information as a) but with an additional quote from U.S. Secretary of the Treasury Jannet Yellen about limiting increased enforcement to high-income earners, and c) a placebo condition providing information about digital versus traditional sources of news consumption. Table 3 shows the six unique treatment arms in the factorial design, with the first

factor spanning the columns and the second factor across the rows.[20]

Table 3: Description of Treatments (Study 3)

	Factor 1	
	Fairness concerns	Placebo
Factor 2	IRS Enforcement Increase	IRS Enforcement Increase
	IRS Enforcement Increase + Yellen Quote	IRS Enforcement Increase + Yellen Quote
	Placebo	Placebo

4.2 Sample

As in both previous studies, this survey experiment was administered through Lucid Marketplace, with the same sampling strategy used previously, but ensuring that respondents did not participate in either of the previous two studies. Once again, we also employed two attention checks (Aronow et al., 2020), both administered prior to the treatments. Due to the factorial design and pre-specified interaction of treatments, we increased the sample size significantly. After dropping respondents who failed the in-survey attention checks, there are 3033 respondents in Study 3.

4.3 Manipulation Check

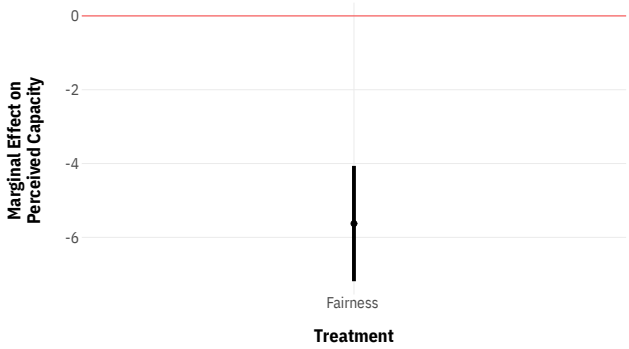
After the first treatment, we asked respondents about their beliefs about capacity and taxpayer compliance. Overall average capacity beliefs are slightly higher than in Studies 1 and 2. However, importantly, comparing responses in the Control groups across the three studies, the average capacity beliefs are quite similar, which adds additional credence to our results. The mean response in the Control group is 72 in Study 3, quite close to the average capacity beliefs for the Control groups in Studies 1 and 2, which were 71.2 and 69.2, respectively. Moreover, as Figure 6 shows, the effects of the *Fairness* treatment in Study 3 on beliefs about IRS capacity are comparable to the results in Study 2. Figure 6 shows the marginal effect and Table G.12 in the Appendix the full regression results.[21] As in Study 2, we find a strong negative effect of the fairness treatment on capacity beliefs. Even though the marginal effect of the *Fairness* treatment is smaller in Study 3 compared to Study 2, the result replicates. As in Studies 1 and 2, capacity

[20] Table G.3 shows the balance across relevant covariates for the different treatment groups.

[21] Given that the question on perceived capacity is asked before the second treatment in Study 3, we only include the first treatment variable in the regression and control for income quartiles to estimate the model used to plot Figure 6.

beliefs once again increase for higher income quartiles, though the differences are generally not statistically significant.

Figure 6: Effect of Treatments on Perceived Capacity (Study 3)



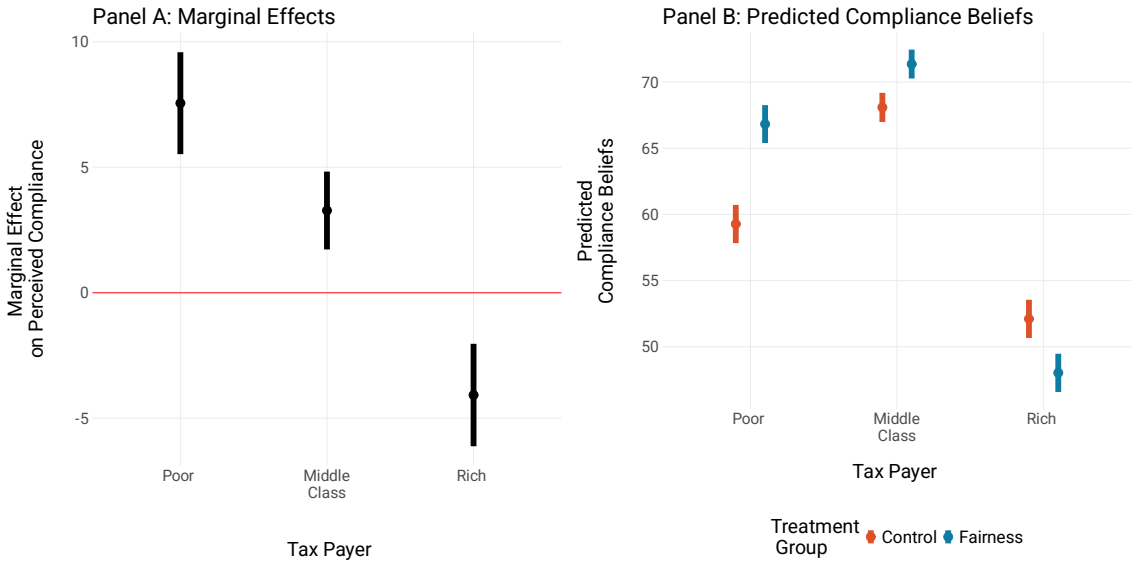
Notes: Figure shows the marginal effect of the fairness prime treatment compared to the control on beliefs about capacity. The treatment significantly decreases respondents' beliefs about the IRS's capacity to "collect taxes and enforce tax laws."

In addition to the expected result on perceived capacity, we find that the first treatment has the expected effects and replicates the findings from Study 1 when it comes to compliance beliefs.[22] As in Study 1, we ask respondents their expectations for the poor, middle class, and rich: What percentage of US residents with those household incomes fully comply with paying taxes?" For each of the three income target groups, we then separately estimate the effect of treatment on compliance beliefs in a simple OLS model, controlling for respondents' income quartiles.

The left plot in Figure 7 shows the marginal effect of the first treatment on respondents' compliance beliefs for the three different groups of taxpayers. Each marginal effect is based on a separate regression model. The full results for each model are shown in Table G.13 in the Appendix. As one can see, the *Fairness* treatment causes respondents to positively update their compliance beliefs for the poor and middle class. In contrast, and as expected, compliance beliefs about rich taxpayers significantly decrease. The right plot, in turn, shows the predicted compliance beliefs for the three groups of taxpayers for both the Control group and those receiving the *Fairness* treatment. In general, compliance beliefs are highest for the middle class and lowest for the rich. While the estimated treatment effect is in the expected direction for all three groups, especially the rich, it is notable that the average ranking of compliance beliefs across the three target income groups is the same for treatment and control respondents.

[22] H16 in the PAP.

Figure 7: Effect of Treatment 1 on Compliance Beliefs (Study 3)



Notes: The left plot shows the marginal effect of the first treatment in Study 3 (*Fairness*) on respondents' compliance beliefs for the three different groups of taxpayers: the poor, the middle class, and the rich. The right panel shows the predicted compliance beliefs across treatment groups (Control and *Fairness* treatment) and groups of taxpayers.

5 Results from Study 3

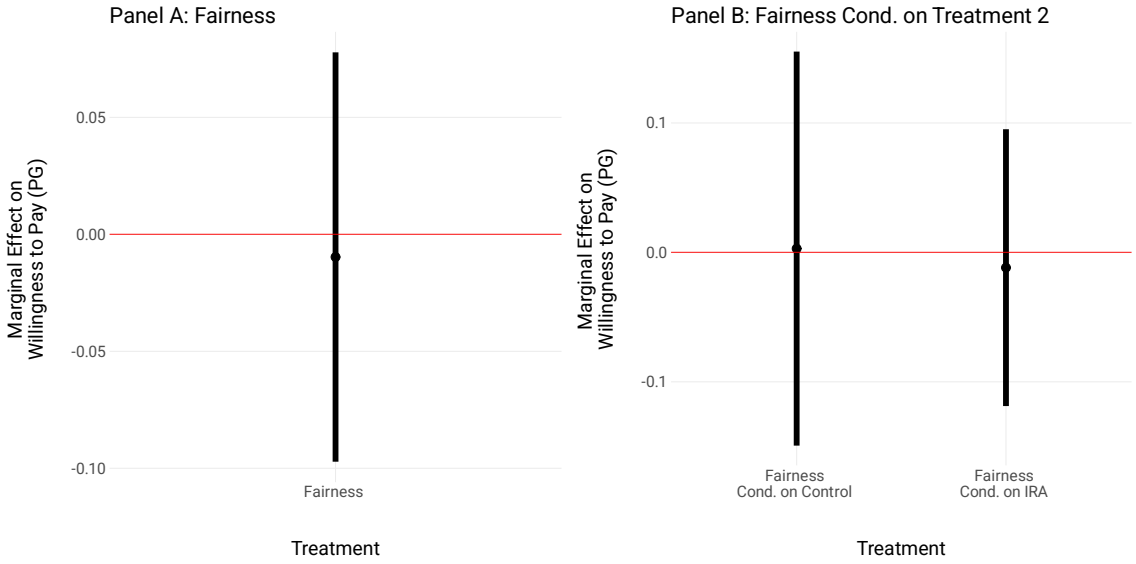
Our main outcome variables of interest in Study 3 were respondents' willingness to pay tax to finance public goods and willingness to pay tax for better enforcement. In Study 2, we found countervailing effects on tax morale across the *Losses* and *Fairness* treatments, wherein the latter – priming differential tax compliance (or enforcement) – lowers respondents' willingness to pay for public goods. In Study 3, we aimed to test whether another informational treatment could reverse that negative effect on tax morale (Hypothesis 6). In particular, we test whether credibly informing respondents about increased enforcement (i.e., the IRA bill) would undo or even reverse the negative effect on morale of priming unfairness due to asymmetric enforcement. We expect that information about the IRA would be particularly effective in doing so when its asymmetric enforcement quality – targeting the rich (*IRA + Yellen quote*) – was made salient (Hypothesis 7).[23]

We do not find strong evidence in favor of either hypothesis (Figure 8). Seemingly in favor of our first prior, we do see that the effect of the *Fairness* treatment on willingness to pay for public goods is no longer negative as it was in Study 2 – it is null as shown in Panel A of Figure 8, which

[23] H19-H20 in Pre-analysis plan.

plots the marginal effect of the *Fairness* treatment averaged over all levels of factor 2.[24] However, in the right plot, we observe that this null effect is apparent in both the IRA condition and the Control condition. So we cannot attribute this less negative effect of the *Fairness* treatment to the IRA condition. Instead, it appears that the Fairness treatment simply yielded a different effect on willingness to pay for public goods in Study 3 than in Study 2 for reasons we explore further below.

Figure 8: Effect of Treatments on Willingness to Pay for Public Goods

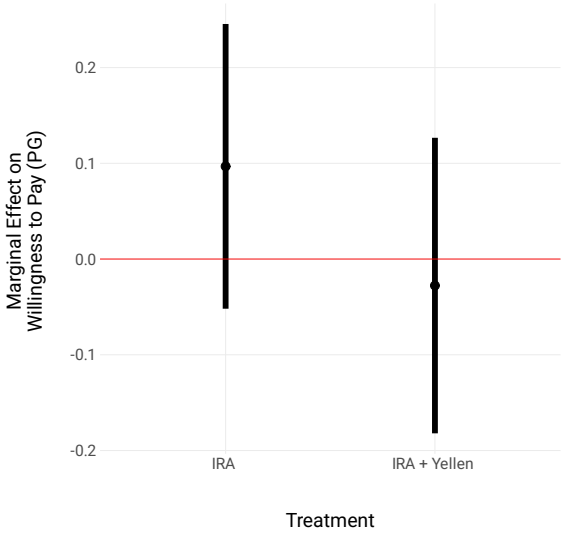


Notes: Panel A shows the marginal effect of the *Fairness* treatment on willingness to pay tax for public goods. The right plot shows the marginal effect of the *Fairness* treatment on willingness to pay for public goods, conditional on being in one of the *IRA* treatments relative to the Control group.

Figure 9 shows the marginal effect of the *IRA* and *IRA + Yellen* treatments on the willingness to pay for education conditional on receiving the *Fairness* treatment in factor 1. While slightly positive, the marginal effect of the *IRA* treatment on willingness to pay for more education is not statistically different from zero. Additionally, and contrary to our expectations, those respondents who received the *IRA + Yellen* treatment did not express a higher willingness to pay tax for public goods. Thus, specifying that enforcement would be limited to the rich did not have the expected differential effect; respondents did not increase their willingness to pay for education. Additionally, the two marginal effects are not significantly different from each other.

[24] Table G.14 in Appendix shows the full results from the underlying regression.

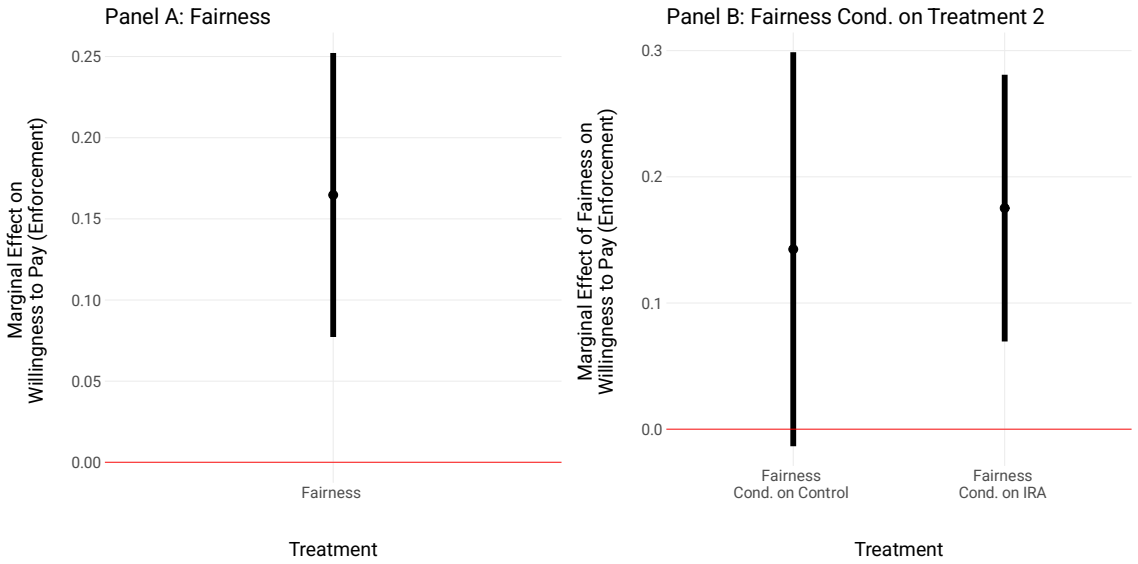
Figure 9: Marginal Effect of IRA Treatment on Willingness to Pay for Public Goods



Notes: The plot shows the marginal effect of the IRA treatments on willingness to pay tax for public goods conditional on receiving the *Fairness* treatment in factor 1. The *IRA* treatment slightly increased respondents' willingness to pay compared to the Control condition, whereas the *IRA + Yellen* treatment did not. Neither marginal effect is statistically significant, however.

Turning to willingness to pay for enforcement as the dependent variable, we see a similar pattern across the different treatments, but with some important differences. First, as with education spending, the effect of the *Fairness* treatment is more positive in Study 3 than it was in Study 2, consistent with expectations. But inconsistent with our priors, this does not seem to be due to the interaction with the *IRA* treatment. Recall that the *Fairness* treatment has no effect on willingness to pay tax for better enforcement in Study 2. In contrast, the same treatment, informing respondents about differences in compliance across income groups, has a strong positive effect on willingness to pay tax for enforcement in Study 3. The left plot of Figure 10 shows the marginal effect of the *Fairness* treatment relative to the Control group averaged over all levels of factor 2. Contrary to expectations, however, while the marginal effect is slightly larger for those respondents who received the *IRA* treatment, the difference in marginal effects conditional on the aggregated *IRA* treatments or Control in factor 2 is not statistically significant (see Figure 10, Panel B). The full results for regressions with willingness to pay tax for enforcement as the dependent variable are presented in Table G.15 in the Appendix.

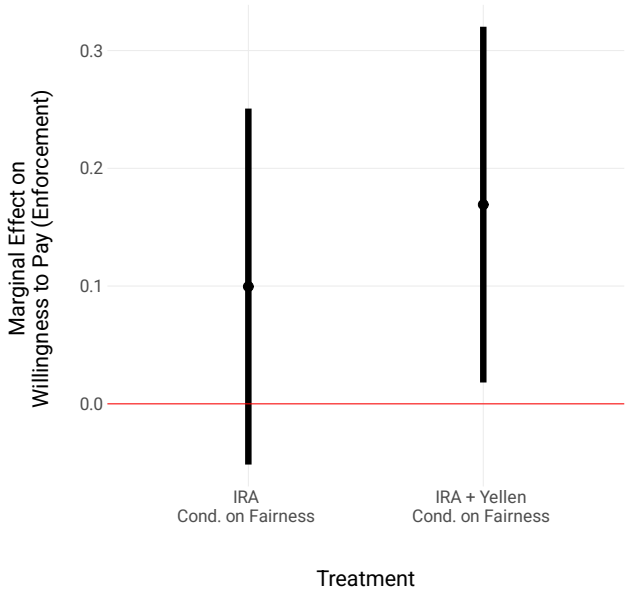
Figure 10: Effect of Treatments on Willingness to Pay for Enforcement (Study 3)



Notes: The left plot shows the marginal effect of the *Fairness* treatment relative to Control on willingness to pay tax for better enforcement across all groups in treatment 2. The right plot shows the marginal effects of the *Fairness* treatment relative to the Control, separated by the collapsed IRA vs Control group in treatment 2.

We also expected that for those respondents who received the *Fairness* treatment, the *IRA* + *Yellen* treatment would have a stronger positive effect than the standalone *IRA* treatment. Figure 11 shows the marginal effects of the two *IRA* treatments relative to the Control, conditional on receiving the *Fairness* treatment first. In contrast to willingness to pay for education, here both *IRA* treatments have a positive effect; however, the *IRA* treatment by itself is not statistically different from zero. As we expected, the treatment that includes the Yellen quote has a slightly larger effect, although differences between the two treatments are not statistically significant. Unfortunately, we are slightly underpowered for this comparison. When aggregated, the two *IRA* treatments with the *Fairness* treatment in factor 1 have a positive and significant effect on willingness to pay for enforcement.

Figure 11: Effect of *IRA* Treatments Conditional on *Fairness* on Willingness to Pay for Enforcement (Study 3)



Notes: The plot shows the marginal effect of the *IRA* and *IRA + Yellen* treatments conditional on receiving the *Fairness* treatment in treatment 1.

6 Discussion

In some ways, the findings from Study 3 raise more questions than they answer. Not only are many findings inconsistent with our priors, but some are also substantively different from those in Study 2. In particular, the *Fairness* treatment has a negative effect on willingness to pay for public goods in Study 2 but is null in Study 3. And the *Fairness* treatment has a null effect on the willingness to pay for enforcement in Study 2 but a positive effect in Study 3. The differences are not due to the addition of the *IRA* treatments, as we see substantively similar effects of the *Fairness* treatment in Study 3 across the Control and *IRA* treatment arms.

One explanation for the different effects across the two studies is that respondents in Study 3 were asked about their willingness to fund the IRA – an actual policy that is funding enforcement rather than funding hypothetical enforcement spending in Study 2. It is possible that respondents were more willing to voice support for an actual policy that has been signed into law than something that is purely hypothetical. But this explanation would not apply to the willingness to pay for public goods results since the question prompts are identical across surveys. Nevertheless, we find differences here too.

Table 4: Control Group Average Willingness to Pay by Party and Study.

	Study 2			Study 3			Study 3 - Study 2	
	Average			Average			Mean Difference	
Party ID	PG	Enforcement	N	PG	Enforcement	N	PG	Enforcement
Independent	2.81	2.52	93	3.41	3.02	85	0.61*	0.5*
No Response	2.95	2.62	143	3.23	2.96	151	0.28	0.33*
Democrat	3.48	3.17	162	3.75	3.81	142	0.27*	0.64*
Republican	2.5	2.36	118	3.12	2.73	122	0.62*	0.37*

Note: The table reports the average willingness to pay for the control group by party and study. The mean difference is calculated as the average willingness to pay in Study 3 minus the average willingness to pay in Study 2. * indicates that the difference is statistically significant at the 0.05 level based on independent sample Welch's t-tests.

A second possible explanation for the differences across studies is timing with respect to the tax calendar. Study 2 was conducted in January, when a lot of people begin to receive their W-2s and begin the tax filing process, whereas Study 3 was conducted in November, when tax filings are less likely to be on people's minds. We might expect people to feel more generous toward tax payments when filing deadlines are far off. Indeed, we see some consistent patterns in the willingness to pay tax outcomes in the Control group (which should not be affected by the different treatment conditions across studies). Table 4 shows the average willingness to pay for public goods and enforcement for the Control group for Study 3 compared to Study 2 across different partisanship responses. Across all groups and both outcomes, the average willingness to pay is substantially higher in Study 3. Except for the *No Response* group in willingness to pay for education, the differences in means are statistically significant based on Welch's independent sample t-tests.

A third explanation for the different treatment effects across studies could be different political contexts in which respondents – especially respondents of particular partisan groups – were more or less malleable in their views. For instance, the IRA was a particularly salient news item in the latter half of 2022, with partisan politics potentially polarizing views on the issue such that informational primes would be less effective when inconsistent with partisan views and more effective when consistent with partisan views. As we show in the Online Appendix, we observe partisan differences consistent with this when we examine changes across studies by partisanship. Self-identified Democrats increased their willingness to pay for enforcement in Study 3 by considerably more than their willingness to pay for public goods (see panels in

Figure G.1). At the same time, Republicans increased their willingness to pay for public goods considerably more than for enforcement. If partisanship affected the way respondents answered these willingness to pay questions differently across studies, it may also have yielded a differential treatment effect across studies. There is some suggestive evidence that the difference in the *Fairness* treatment's effect on willingness to pay for enforcement between Studies 2 and 3 – from null to positive – is being driven by Republicans and those refusing to provide information on partisanship. These are exactly the two groups where the increases in the willingness to pay for enforcement in the Control group were smallest, so this could be consistent with the other groups experiencing a ceiling effect. However, this is still surprising as we might have expected Republicans to be less malleable in their support for the IRA after months of news coverage. The effect of the *Fairness* treatment is close to zero for Democrats, consistent with a possible ceiling effect due to partisanship.

Similarly, we observe some partisan heterogeneity when it comes to the effects of the *IRA* treatment on willingness to pay tax in Study 3. As we show in the Online Appendix (Figure G.2), the *IRA* treatment had significant effects on the willingness to pay for both public goods and enforcement, regardless of the condition in treatment 1 for respondents unwilling to provide their partisanship. In addition, we see the hypothesized differential effect of the *IRA* treatment conditional on receiving the Fairness treatment in the first arm on willingness to pay for public goods for independent partisans. In contrast, we see no discernible effect for either Democrats or Republicans. While not pre-registered and mostly suggestive, these results provide some evidence that our treatments may have been overwhelmed by partisan priors.

A final possible explanation for the different effects of the *Fairness* treatment between Study 3 and 2 is that people in the Control condition in December 2022 were already primed to think about the IRA since it had been in the news since its passage in August and again during the midterm elections in November, right before the survey was fielded. If people were already primed to think that the unfairness of asymmetric non-compliance by the rich was in the process of being rectified by the planned improvements to IRS enforcement capacity, then reminding them about prior unfairness could be less demoralizing. This could also explain why the IRA treatment was not a moderator of the Fairness treatment as we expected: people in the Control condition may already have been treated by news of the IRA's passage. These are speculations that should be tested in future research, especially now that the IRA's effects are in the process of being rewound.

7 Conclusion

Together, our studies show that tax morale can be shaped by citizens' perceptions of the state's fiscal capacity and its ability to enforce tax laws equitably across socioeconomic groups. Through three survey experiments conducted in the United States, we demonstrate that perceptions of asymmetric tax enforcement – especially the belief that high-income earners evade taxes more easily – are malleable and may affect citizens' willingness to pay taxes, especially for improving tax enforcement. These findings question the common assumption in the literature that state capacity is uniformly perceived and stable. Perceptions of a state's tax capacity change with new information, and those changes can be consequential for tax morale.

We also identify two mechanisms through which perceptions of asymmetric tax enforcement and compliance can influence willingness to pay tax – fairness and material concerns – and show that they can have different effects on tax morale. Importantly, the results from Study 2 suggest that emphasizing the unfairness of asymmetric tax enforcement tends to lower willingness to pay for public goods (education), whereas focusing on revenue losses stemming from asymmetric compliance can increase willingness to pay for tax enforcement improvements. As such, informing taxpayers about revenue losses from asymmetric compliance is a potential strategy to address, through tax morale, fiscal capacity weaknesses that lead to asymmetric enforcement.

References

Alm, James, Kim M Bloomquist, and Michael McKee. 2017. "When you know your neighbour pays taxes: Information, peer effects and tax compliance." *Fiscal Studies* 38 (4): 587–613.

Anson, Ian G, and John V Kane. 2023. "Ought It Audit? Information, Values, and Public Support for the Internal Revenue Service." *Journal of Experimental Political Science* 10 (2): 209–220.

Aronow, Peter Michael, Joshua Kalla, Lilla Orr, and John Ternovski. 2020. "Evidence of Rising Rates of Inattentiveness on Lucid in 2020.".
URL: *http://dx.doi.org/10.31235/osf.io/8sbe4*

Bachas, Pierre, Anders Jensen, and Lucie Gadenne. 2024. "Tax equity in low-and middle-income countries." *Journal of Economic Perspectives* 38 (1): 55–80.

Benedek, Ms Dora, Juan Carlos Benítez, and Charles Vellutini. 2022. *Progress of the personal income tax in emerging and developing countries*. International Monetary Fund.

Bergeron, Augustin, Gabriel Tourek, and Jonathan L Weigel. 2024. "The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the drc." *Econometrica* 92 (4): 1163–1193.

Besley, Timothy, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton, NJ: Princeton University Press.

Besley, Timothy, and Torsten Persson. 2014. "Why Do Developing Countries Tax So Little?" *Journal of Economic Perspectives* 28 (4): 99–120.

Bird, Richard M., and Eric M. Zolt. 2004. "Redistribution Via Taxation: The Limited Role of the Personal Income Tax in Developing Countries." *UCLA Law Review* 52: 1627–1695.

Boning, William C, Nathaniel Hendren, Ben Sprung-Keyser, and Ellen Stuart. 2025. "A welfare analysis of tax audits across the income distribution." *The Quarterly Journal of Economics* 140 (1): 63–112.

Bullock, John G., Alan S. Gerber, Seth J. Hill, and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs About Politics." *Quarterly Journal of Political Science* 10 (4): 519–578.
URL: *http://dx.doi.org/10.1561/100.00014074*

Campos-Vazquez, Raymundo M, Camilo A Cepeda-Francese, Alice Krozer, and Aurora A Ramírez-Álvarez. 2023. "The impact of tax evasion by the rich on the willingness to pay taxes." *Applied Economics Letters* pp. 1–4.

Castañeda, Néstor. 2024. "Fairness and tax morale in developing countries." *Studies in Comparative International Development* 59 (1): 113–137.

Castro, Lucio, and Carlos Scartascini. 2015. "Tax compliance and enforcement in the pampas evidence from a field experiment." *Journal of Economic Behavior & Organization* 116: 65–82.

Choudhary, Rishabh, Franz U Ruch, and Emilia Skrok. 2024. *Taxing for Growth*. World Bank.

Cobham, Alex, and Petr Janskỳ. 2018. "Global distribution of revenue loss from corporate tax avoidance: re-estimation and country results." *Journal of International Development* 30 (2): 206–232.

de Mesquita, Bruce Bueno, George Downs, and Alastair Smith. 2017. "A Political Economy of Income Tax Policies." *Political Science Research and Methods* 5 (1): 1–29.

Del Carpio, Lucia. 2013. "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru." Working Paper.
URL: *https://irs.princeton.edu/sites/irs/files/event/uploads/DelCarpioOct2013.pdf*

Fernández-Albertos, José, and Alexander Kuo. 2015. "Income Perception, Information, and Progressive Taxation: Evidence From a Survey Experiment." *Political Science Research and Methods* 6 (1): 83–110.
URL: *https://doi.org/10.1017/psrm.2015.73*

Filippin, Antonio, Carlo V. Fiorio, and Eliana Viviano. 2013. "The effect of tax enforcement on tax morale." *European Journal of Political Economy* 32: 320–331.
URL: *http://dx.doi.org/10.1016/j.ejpoleco.2013.09.005*

Fjeldstadt, Odd-Helge, and Mick Moore. 2008. "Tax Reform and State-Building in a Globalised World." In *Taxation and State-Building in Developing Countries: Capacity and Consent*, ed. Deborah Br/"autigam, Odd-Helge Fjeldstadt, and Mick Moore. Cambridge, U.K.: Cambridge University Press chapter 10, pp. 235–260.

Frank, Robert. 2022. "Just 2% of the richest Americans had their taxes audited in 2019, down from 16% in 2010." Last Accessed August 8, 2022.
URL: *https://tinyurl.com/3xusd2uv*

Frey, Bruno S., and Benno Torgler. 2007. "Tax Morale and Conditional Cooperation." *Journal of Comparative Economics* 35 (1): 136–159.
URL: *http://dx.doi.org/10.1016/j.jce.2006.10.006*

Gordon, Roger, and Wei Li. 2005. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." Working Paper. http://www.econ.ucsd.edu/~rogordon/puzzles16.pdf (accessed June 21, 2011).

Guyton, John, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman. 2021. Tax Evasion at the Top of the Income Distribution: Theory and Evidence. Working Paper 28542.

Luna, Juan Pablo, and Hillel David Soifer. 2017. "Capturing Sub-National Variation in State Capacity: A Survey-Based Approach." *American Behavioral Scientist* 61 (8): 887–907.
URL: *http://dx.doi.org/10.1177/0002764217720964*

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Rubolino, Enrico. 2023. "Does weak enforcement deter tax progressivity?" *Journal of Public Economics* 219: 104833.

Slemrod, Joel. 2019. "Tax compliance and enforcement." *Journal of economic literature* 57 (4): 904–954.

Tax Policy Center. 2024. "The Tax Policy Briefing Book: How Did the Inflation Reduction Act of 2022 Affect the IRS's Budget?" https://www.taxpolicycenter.org/briefing-book. Accessed: 2025-06-16.

U.S. Government Accountability Office. 2022. Tax Compliance: IRS Audit Trends for Individual Taxpayers Vary by Income. Technical Report GAO-22-106032 U.S. Government Accountability Office. Accessed: 2025-06-16.
URL: *https://www.gao.gov/products/gao-22-106032*

Wilson, Nicole E, and Leah R Rosenzweig. 2024. "Paying for growth or goods: Tax morale among property owners in Lagos." *Journal of Experimental Political Science* 11 (3): 277–287.

Supplementary Online Appendix

Additional Results Studies 1 & 2

Table G.1: Balance Table - Study 1

		Control (N=500)		Low (N=462)		Asymmetric (N=460)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education		4.74	1.96	4.79	1.99	4.77	1.98
Age		47.67	17.16	47.84	17.83	47.90	17.39
		N	Pct.	N	Pct.	N	Pct.
Pol. Party	Democrat	225	45.0	194	42.0	183	39.8
	Republican	155	31.0	166	35.9	134	29.1
	Independent	120	24.0	102	22.1	143	31.1
Income Quartiles	1st quartile	142	28.4	119	25.8	122	26.5
	2nd quartile	149	29.8	136	29.4	138	30.0
	3rd quartile	137	27.4	138	29.9	136	29.6
	4th quartile	72	14.4	69	14.9	64	13.9

Means and Standard Devations for continuous variables and group sizes for categorical variables by treatment group for Study 1.

Table G.2: Balance Table - Study 2

		Control (N=516)		Fairness (N=500)		Losses (N=505)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education		4.62	1.87	4.32	1.90	4.42	1.88
Age		46.71	17.37	47.90	17.87	48.35	17.94
		N	Pct.	N	Pct.	N	Pct.
Pol. Party	Independent	93	18.0	67	13.4	80	15.8
	No Response	143	27.7	138	27.6	155	30.7
	Democrat	162	31.4	168	33.6	164	32.5
	Republican	118	22.9	127	25.4	106	21.0
Income Quartiles	1st quartile	146	28.3	147	29.4	148	29.3
	2nd quartile	162	31.4	157	31.4	164	32.5
	3rd quartile	149	28.9	138	27.6	136	26.9
	4th quartile	59	11.4	58	11.6	57	11.3

Means and Standard Devations for continuous variables and group sizes for categorical variables by treatment group for Study 2.

Table G.3: Balance Table - Study 3

Treatment 1		Control						Fairness					
Treatment 2		Control		IRA		IRA + Yellen		Control		IRA		IRA + Yellen	
		(N=500)		(N=531)		(N=530)		(N=492)		(N=519)		(N=469)	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Education		4.61	2.01	4.58	1.97	4.41	1.98	4.39	1.91	4.54	1.96	4.41	1.99
Age		47.15	18.59	45.88	18.28	47.28	18.45	46.99	17.93	47.16	18.86	48.56	17.62
		N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
Pol. Party	Independent	85	17.0	79	14.9	98	18.5	82	16.7	95	18.3	89	19.0
	No Response	151	30.2	152	28.6	135	25.5	128	26.0	136	26.2	119	25.4
	Democrat	142	28.4	168	31.6	167	31.5	165	33.5	159	30.6	146	31.1
	Republican	122	24.4	132	24.9	130	24.5	117	23.8	129	24.9	115	24.5
Income Quartiles	1st quartile	147	29.4	166	31.3	158	29.8	158	32.1	149	28.7	136	29.0
	2nd quartile	152	30.4	148	27.9	195	36.8	165	33.5	164	31.6	150	32.0
	3rd quartile	137	27.4	146	27.5	122	23.0	125	25.4	130	25.0	131	27.9
	4th quartile	64	12.8	71	13.4	55	10.4	44	8.9	76	14.6	52	11.1

Means and Standard Devations for continuous variables and group sizes for categorical variables by treatment groups for Study 3.

Table G.4: Descriptives: Perceived Capacity

Outcome:	Perceived IRS Capacity		
	Study 1	Study 2	Study3
Intercept	61.033***	55.059***	63.736***
	(6.168)	(6.079)	(2.232)
Democrat	-0.491	3.442	1.571
	(2.143)	(2.317)	(1.276)
Republican	0.601	7.428**	2.757*
	(2.291)	(2.486)	(1.337)
Education	-0.852+	-0.200	-0.066
	(0.497)	(0.434)	(0.233)
White	-3.695	4.027	-0.551
	(5.661)	(5.258)	(1.414)
Black	0.384	11.580*	-0.108
	(6.185)	(5.528)	(1.840)
Asian	-0.936	5.388	-2.649
	(7.857)	(6.382)	(2.518)
Hispanic	-4.916	4.431	
	(6.005)	(5.653)	
Age	-0.094+	-0.083+	0.072**
	(0.053)	(0.048)	(0.024)
2nd Income Quartile	4.948*	0.552	1.282
	(2.225)	(1.935)	(1.106)
3rd Income Quartile	5.656*	-1.599	1.277
	(2.411)	(2.189)	(1.216)
4th Income Quartile	9.990***	1.526	-0.312
	(2.993)	(3.015)	(1.592)
Audit Experience	9.811***	3.960*	3.340**
	(1.995)	(1.890)	(1.077)
Voted	3.585	-1.991	1.013
	(2.444)	(2.106)	(1.177)
Num. Federal Benefits	0.756*	0.492	-0.208
	(0.383)	(0.373)	(0.214)
Num. Observations	1066	1264	2817
Adj. R-squared	0.043	0.016	0.006

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001 Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.5: Perceived IRS Capacity (Studies 2 & 3)

Outcome:	Perceived IRS Capacity	
	Study 1	Study 2
Intercept	66.250***	69.682***
	(1.528)	(1.558)
Low	-19.885***	
	(1.599)	
Asymmetric	-13.447***	
	(1.561)	
Fairness		-12.131***
		(1.611)
Losses		-16.538***
		(1.600)
2nd Income Quartile	5.483**	0.547
	(1.785)	(1.728)
3rd Income Quartile	6.983***	-2.204
	(1.765)	(1.795)
4th Income Quartile	9.738***	0.181
	(2.209)	(2.381)
Num. Observations	1422	1521
Adj. R-squared	0.110	0.066

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.6: Perceived Compliance across Income Groups

Outcome:	Perceived Compliance		
	Rich	Middle Class	Poor
Intercept	49.881***	65.125***	59.519***
	(1.800)	(1.408)	(1.820)
Low	-3.891*	-3.739**	-3.925*
	(1.857)	(1.404)	(1.934)
Asymmetric	-10.725***	0.730	5.266**
	(1.842)	(1.376)	(1.834)
2nd Income Quartile	1.488	5.480***	1.654
	(1.984)	(1.553)	(2.022)
3rd Income Quartile	3.965*	6.456***	0.441
	(2.009)	(1.574)	(2.067)
4th Income Quartile	12.118***	9.008***	3.968
	(2.505)	(1.852)	(2.521)
Num. Observations	1422	1422	1422
Adj. R-squared	0.038	0.025	0.015

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.7: Willingness to Pay for PG - Study 1

Outcome:	Willingness To Pay Tax for PG	
	Binary	Continuous
Intercept	0.563***	2.532***
	(0.034)	(0.074)
Low	-0.040	-0.099
	(0.035)	(0.076)
Asymmetric	0.022	0.031
	(0.035)	(0.077)
2nd Income Quartile	-0.100*	-0.207*
	(0.040)	(0.086)
3rd Income Quartile	-0.000	0.031
	(0.039)	(0.085)
4th Income Quartile	0.066	0.131
	(0.046)	(0.103)
Num. Observations	1179	1179
Adj. R-squared	0.011	0.010

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.8: Willingness to Pay for PG – Mediated by Compliance Beliefs – Study 1

Outcome:	Willingness To Pay Tax for PG			
	Treatment:			
	Low Enforcement		Asymmetric Enforcement	
	Mediator:			
	Compliance Rich	Compliance Poor	Compliance Rich	Compliance Poor
ACME	-0.015+	-0.027*	-0.047***	0.014
	[-0.039, 0.000]	[-0.062, -0.003]	[-0.086, -0.019]	[-0.002, 0.036]
ADE	-0.090	-0.068	0.082	0.023
	[-0.228, 0.033]	[-0.211, 0.081]	[-0.063, 0.225]	[-0.120, 0.192]
Total Effect	-0.105	-0.096	0.035	0.037
	[-0.245, 0.024]	[-0.236, 0.047]	[-0.105, 0.175]	[-0.103, 0.205]
Prop. Mediated	0.135	0.225	-0.441	0.083
	[-0.521, 1.887]	[-1.428, 3.052]	[-8.541, 7.293]	[-2.664, 1.878]

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the mediation package in R. Robust standard errors are calculated.

Table G.9: Fairness Tax System - Study 1

Outcome:	Perceived Fairness of:	
	Tax Collection	Tax Rates
Intercept	3.041***	2.971***
	(0.073)	(0.073)
Low	-0.204**	-0.122
	(0.076)	(0.075)
Asymmetric	-0.698***	-0.523***
	(0.079)	(0.079)
2nd Income Quartile	-0.022	-0.059
	(0.082)	(0.083)
3rd Income Quartile	0.154+	0.138
	(0.085)	(0.084)
4th Income Quartile	0.444***	0.409***
	(0.111)	(0.109)
Num. Observations	1398	1397
Adj. R-squared	0.069	0.047

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.10: Willingness to Pay - Study 2

Outcome:	Willingness To Pay Tax for			
	PG		Enforcement	
	Binary	Continuous	Binary	Continuous
Intercept	0.401***	3.023***	0.282***	2.709***
	(0.029)	(0.077)	(0.027)	(0.077)
Fairness	-0.049	-0.171*	0.025	0.050
	(0.031)	(0.083)	(0.029)	(0.083)
Losses	0.010	0.036	0.067*	0.218**
	(0.031)	(0.083)	(0.030)	(0.084)
2nd Income Quartile	0.023	0.008	0.051	0.068
	(0.032)	(0.085)	(0.031)	(0.087)
3rd Income Quartile	0.015	-0.081	0.006	-0.066
	(0.034)	(0.090)	(0.032)	(0.089)
4th Income Quartile	0.045	-0.135	0.083+	0.058
	(0.045)	(0.125)	(0.044)	(0.124)
Num. Observations	1506	1506	1507	1507
Adj. R-squared	0.000	0.003	0.004	0.003

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.11: Willingness to Pay by Partisan ID - Study 2

Outcome:	Willingness To Pay Tax for:	
	Public Goods	Enforcement
Intercept	2.823***	2.495***
	(0.136)	(0.134)
Fairness	-0.093	0.202
	(0.201)	(0.196)
Losses	0.237	0.341+
	(0.199)	(0.200)
Party ID: No Response	0.152	0.104
	(0.172)	(0.172)
Party ID: Democrat	0.678***	0.657***
	(0.158)	(0.158)
Party ID: Republican	-0.294	-0.157
	(0.180)	(0.185)
Fairness x Party ID: No Response	-0.016	-0.181
	(0.256)	(0.252)
Fairness x Party ID: Democrat	-0.108	-0.066
	(0.239)	(0.238)
Fairness x Party ID: Republican	-0.182	-0.350
	(0.263)	(0.261)
Losses x Party ID: No Response	-0.182	0.075
	(0.253)	(0.257)
Losses x Party ID: Democrat	-0.435+	-0.318
	(0.240)	(0.238)
Losses x Party ID: Republican	-0.095	-0.252
	(0.269)	(0.272)
2nd Income Quartile	-0.002	0.060
	(0.084)	(0.084)
3rd Income Quartile	-0.041	-0.025
	(0.088)	(0.087)
4th Income Quartile	-0.089	0.105
	(0.118)	(0.118)
Num. Observations	1506	1507
Adj. R-squared	0.063	0.067

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.12: Perceived IRS Capacity (Study 3)

Outcome:	Perceived IRS Capacity
	(1)
Intercept	70.486***
	(0.855)
Fairness	-5.612***
	(0.797)
2nd Income Quartile	1.854+
	(1.030)
3rd Income Quartile	2.407*
	(1.069)
4th Income Quartile	2.100
	(1.362)
Num. Observations	3041
Adj. R-squared	0.017

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Additional Results Study 3

Table G.16: Willingness to Pay by Partisan ID - Study 3

Outcome:	Willingness To Pay Tax for:	
	Public Goods	Enforcement
Intercept	3.444***	3.075***
	(0.127)	(0.131)
Fairness	-0.307+	0.168
	(0.174)	(0.182)
IRA	-0.328+	-0.060
	(0.175)	(0.179)
IRA + Yellen	-0.222	0.095

Continued on next page

Table G.13: Perceived Compliance across Income Groups (Study 3)

Outcome:	Perceived Compliance		
	Rich	Middle Class	Poor
Intercept	50.086***	63.698***	58.575***
	(1.117)	(0.865)	(1.087)
Fairness	-4.078***	3.275***	7.552***
	(1.041)	(0.791)	(1.036)
2nd Income Quartile	0.562	5.094***	1.694
	(1.342)	(1.038)	(1.317)
3rd Income Quartile	2.449+	6.862***	0.024
	(1.415)	(1.067)	(1.402)
4th Income Quartile	10.133***	8.166***	1.287
	(1.730)	(1.307)	(1.768)
Num. Observations	3041	3041	3041
Adj. R-squared	0.015	0.023	0.017

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.16: Willingness to Pay by Partisan ID - Study 3 (Continued)

	(0.168)	(0.179)
Party ID: No Response	-0.181	-0.084
	(0.164)	(0.165)
Party ID: Democrat	0.334*	0.779***
	(0.152)	(0.152)
Party ID: Republican	-0.297+	-0.310+
	(0.167)	(0.175)
Fairness x IRA	0.823***	0.244
	(0.242)	(0.254)
Fairness x IRA + Yellen	0.340	0.119
	(0.239)	(0.248)

Continued on next page

Table G.14: Willingness to Pay for PG - Study 3

Outcome:	Willingness To Pay Tax for Public Good			
	Binary	Continuous	Binary	Continuous
Intercept	0.493***	3.416***	0.493***	3.416***
	(0.027)	(0.065)	(0.027)	(0.065)
Fairness	0.014	0.003	0.015	0.003
	(0.032)	(0.078)	(0.032)	(0.078)
IRA	0.073*	0.133+		
	(0.031)	(0.078)		
IRA + Yellen	0.013	-0.028		
	(0.032)	(0.077)		
Collapsed IRA			0.043	0.052
			(0.027)	(0.067)
2nd Income Quartile	-0.003	-0.052	-0.005	-0.056
	(0.023)	(0.056)	(0.023)	(0.056)
3rd Income Quartile	0.004	-0.062	0.005	-0.061
	(0.025)	(0.060)	(0.025)	(0.060)
4th Income Quartile	0.078*	0.016	0.080**	0.023
	(0.031)	(0.079)	(0.031)	(0.079)
Num. Observations	2978	2978	2978	2978
Adj. R-squared	0.003	0.001	0.002	-0.001

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.15: Willingness to Pay for Enforcement - Study 3

Outcome:	Willingness To Pay Tax for Enforcement			
	Binary	Continuous	Binary	Continuous
Intercept	0.420***	3.206***	0.420***	3.207***
	(0.027)	(0.066)	(0.027)	(0.066)
Fairness	0.056+	0.143+	0.056+	0.143+
	(0.032)	(0.080)	(0.032)	(0.080)
IRA	0.068*	0.135+		
	(0.032)	(0.079)		
IRA + Yellen	0.040	0.065		
	(0.032)	(0.079)		
Collapsed IRA			0.054*	0.100
			(0.027)	(0.069)
2nd Income Quartile	-0.017	-0.117*	-0.017	-0.118*
	(0.023)	(0.055)	(0.023)	(0.055)
3rd Income Quartile	0.006	-0.108+	0.007	-0.107+
	(0.024)	(0.060)	(0.025)	(0.060)
4th Income Quartile	0.120***	0.087	0.119***	0.086
	(0.031)	(0.080)	(0.031)	(0.080)
Num. Observations	2970	2970	2970	2970
Adj. R-squared	0.010	0.008	0.010	0.008

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Table G.16: Willingness to Pay by Partisan ID - Study 3 (Continued)

Fairness x Party ID: No Response	0.388+	0.012
	(0.235)	(0.240)
Fairness x Party ID: Democrat	0.324	-0.192
	(0.214)	(0.215)
Fairness x Party ID: Republican	0.304	-0.066
	(0.236)	(0.245)
IRA x Party ID: No Response	0.687**	0.425+
	(0.233)	(0.232)
IRA x Party ID: Democrat	0.494*	-0.021
	(0.214)	(0.215)
IRA x Party ID: Republican	0.389	0.208
	(0.238)	(0.242)
IRA + Yellen x Party ID: No Response	0.432+	0.301
	(0.226)	(0.234)
IRA + Yellen x Party ID: Democrat	0.157	-0.221
	(0.210)	(0.211)
IRA + Yellen x Party ID: Republican	0.069	-0.281
	(0.235)	(0.243)
Fairness x IRA x Party ID: No Response	-0.860**	-0.223
	(0.322)	(0.332)
Fairness x IRA x Party ID: Democrat	-1.077***	-0.180
	(0.298)	(0.302)
Fairness x IRA x Party ID: Republican	-0.999**	-0.434
	(0.335)	(0.341)
Fairness x IRA + Yellen x Party ID: No Response	-0.426	-0.251

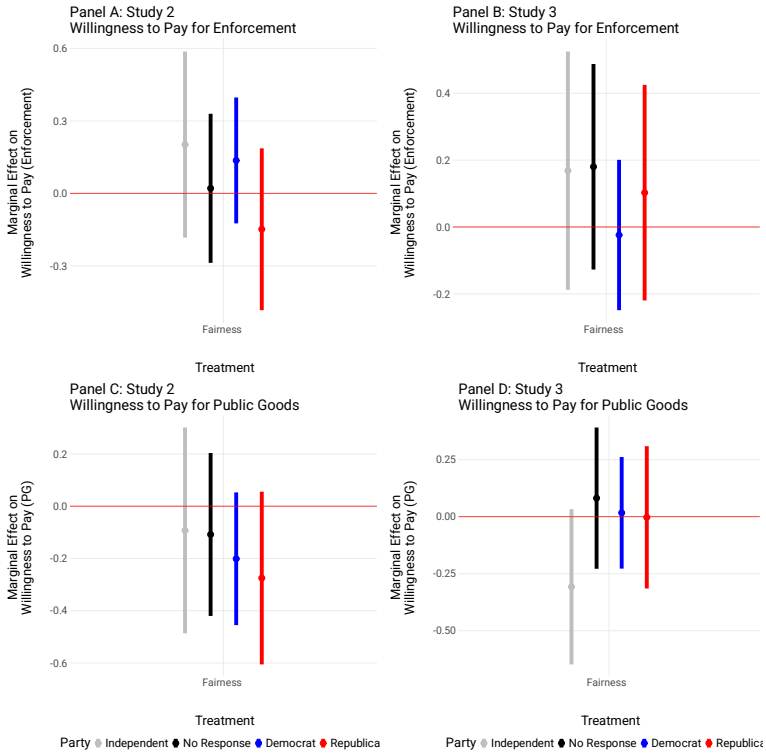
Continued on next page

Table G.16: Willingness to Pay by Partisan ID - Study 3 (Continued)

	(0.325)	(0.332)
Fairness x IRA + Yellen x Party ID: Democrat	-0.433	0.051
	(0.301)	(0.295)
Fairness x IRA + Yellen x Party ID: Republican	-0.234	0.338
	(0.335)	(0.338)
2nd Income Quartile	-0.043	-0.092+
	(0.055)	(0.053)
3rd Income Quartile	-0.049	-0.073
	(0.059)	(0.057)
4th Income Quartile	-0.014	0.070
	(0.078)	(0.076)
Num. Observations	2978	2970
Adj. R-squared	0.046	0.098

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Figure G.1: Marginal Effect of Fairness by Respondent Party (Studies 2 & 3)



Notes: Panels (A) and (B) show the marginal effect of the fairness treatment on willingness to pay for enforcement by party identification in Studies 2 and 3, respectively. The second row – panels (C) and (D) – shows the marginal effects of the fairness treatment on the willingness to pay for education, again separated by party ID. For Study 3, the marginal effects of fairness are limited to the pure control group in factor 2.

Table G.17: Willingness to Pay by Partisan ID (Collapsed IRA Treatment) - Study 3

Outcome:	Willingness To Pay Tax for:	
	Public Goods	Enforcement
Intercept	3.443***	3.074***
	(0.126)	(0.131)
Fairness	-0.308+	0.168
	(0.174)	(0.182)
Collapsed IRA	-0.270+	0.026
	(0.148)	(0.155)
Party ID: No Response	-0.183	-0.085

Continued on next page

Table G.17: Willingness to Pay by Partisan ID (Collapsed IRA Treatment) - Study 3 (Continued)

	(0.164)	(0.165)
Party ID: Democrat	0.333*	0.779***
	(0.152)	(0.152)
Party ID: Republican	-0.299+	-0.310+
	(0.167)	(0.175)
Fairness x Collapsed IRA	0.580**	0.172
	(0.210)	(0.219)
Fairness x Party ID: No Response	0.391+	0.012
	(0.235)	(0.240)
Fairness x Party ID: Democrat	0.325	-0.192
	(0.214)	(0.215)
Fairness x Party ID: Republican	0.306	-0.066
	(0.236)	(0.245)
Collapsed IRA x Party ID: No Response	0.558**	0.354+
	(0.199)	(0.202)
Collapsed IRA x Party ID: Democrat	0.320+	-0.129
	(0.184)	(0.185)
Collapsed IRA x Party ID: Republican	0.224	-0.046
	(0.205)	(0.212)
Fairness x Collapsed IRA x Party ID: No Response	-0.642*	-0.224
	(0.283)	(0.290)
Fairness x Collapsed IRA x Party ID: Democrat	-0.752**	-0.058
	(0.260)	(0.260)
Fairness x Collapsed IRA x Party ID: Republican	-0.617*	-0.045

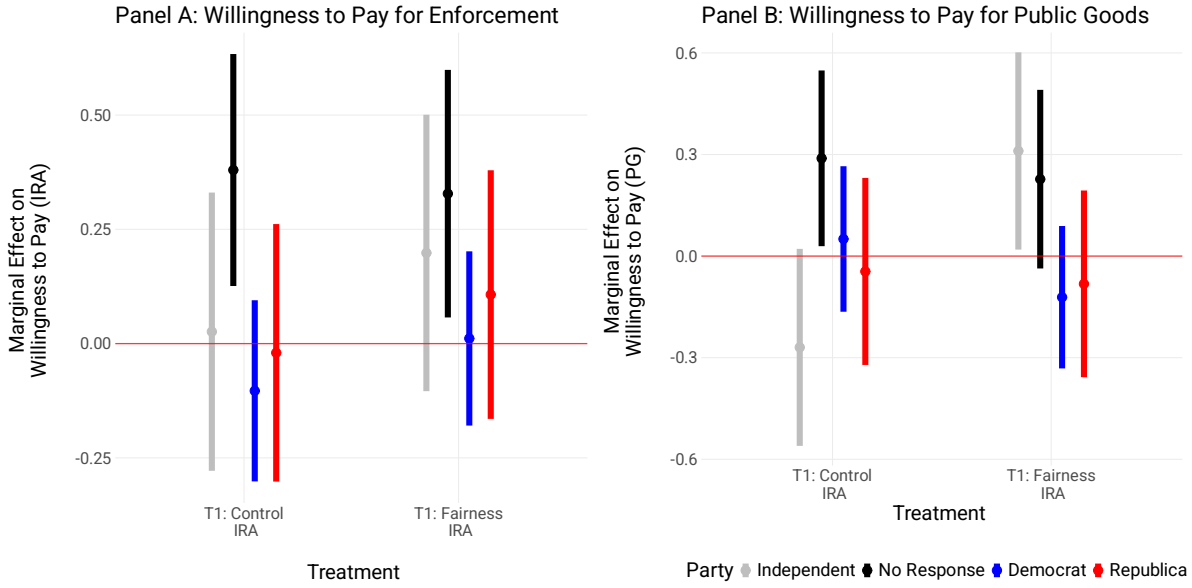
Continued on next page

Table G.17: Willingness to Pay by Partisan ID (Collapsed IRA Treatment) - Study 3 (Continued)

	(0.289)	(0.296)
2nd Income Quartile	-0.046	-0.092+
	(0.055)	(0.053)
3rd Income Quartile	-0.045	-0.068
	(0.059)	(0.057)
4th Income Quartile	-0.001	0.073
	(0.078)	(0.076)
Num. Observations	2978	2970
Adj. R-squared	0.045	0.097

Note: + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Models estimated using the estimatr package in R. Robust standard errors (HC2) are calculated.

Figure G.2: Marginal Effect of IRA Treatment & Treatment 1 by Respondent Party (Study 3)



Notes: Panel (A) and (B) show the marginal effects of the combined *IRA* treatment conditional on treatment arm 1 on willingness to pay for enforcement (A) and public goods (B) separated by party identification. As we can see, the IRA treatment effect on willingness to pay for enforcement conditional on the *Fairness* treatment (see Figure 11) is driven by Independents and respondents unwilling to state their partisanship. Similarly, we see a strong effect of the *IRA* treatment on willingness to pay for education, but again, only for independents and those without a partisanship response.

Treatments

Study 1

Weak Enforcement

> This week, the IRS released its annual disclosure of enforcement statistics. Every year, it's an opportunity to measure how effectively the U.S. government has sabotaged its own ability to enforce its tax laws. This year's report, along with other data ProPublica has collected on the state of the IRS, is full of evidence that the IRS has hit a passel of historic lows... Overall, **2019 brought the lowest audit rate in generations.** ProPublica searched back to the 1950s and could not find a lower audit rate of individual returns... Fewer audits has meant less revenue.
>
> And then there are collections — in which the issue isn't auditing taxpayers, but collecting what they've agreed to pay. One clear measure is how often the IRS simply lets tax debts evaporate because it doesn't have the personnel to pursue them. This used to happen only rarely, but in 2019, the IRS let $6.7 billion in tax debt expire, an increase of over 1,000% from 2010.
>
> If the IRS had maintained a level of enforcement similar to that of the years from 2004 to 2010, it would have collected about $115 billion more. Since 2010, the IRS lost about 5,900 agents. The total cost of hiring that number of agents would be less than $1 billion.

Asymmetric Enforcement

> For the past couple years, ProPublica has been cataloging the descent of the IRS. We've watched as audits of the rich and the largest corporations have plummeted and become less aggressive, while **audits of poor taxpayers have remained comparatively high.**
>
> The abysmal audit rate of the nation's richest was no surprise to the ranking member of the Senate Finance Committee and a consistent critic of budget cuts to the agency. "It's a national scandal that the wealthy are stealing tens of billions from American taxpayers," he said. "Paying taxes has become increasingly voluntary for those at the top." In past years, members of Congress have pointed to our coverage and pressured IRS Commissioner Charles Rettig to increase audits of the rich; one of Rettig's responses [was] auditing the poor was simply much easier.
>
> Last month, the agency was embarrassed by a scathing inspector general report that hundreds of thousands of wealthy taxpayers had been allowed to skip filing taxes altogether without consequence. This was because the agency decided at one point in 2015 to all but stop chasing such [wealthy] nonfilers because of inadequate staffing. If the IRS had maintained a level of enforcement similar to that of the years from 2004 to 2010, it would have collected about $115 billion more. Since 2010, the IRS lost about 5,900 agents. The total cost of hiring that number of agents would be less than $1 billion.

Study 2

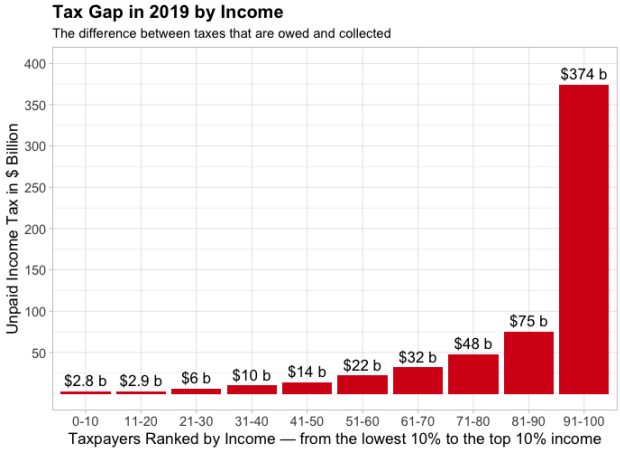
Losses Prime

The excerpt below was adapted from an article published by the U.S. Department of the Treasury last year. Please closely read the excerpt below before moving on to the next question. You will be able to move to the next question after carefully reading the text below.

A well-functioning tax system requires that everyone pays the taxes they owe. Today, the "**tax gap**"—the difference between taxes that are owed and collected—totals around **$600 billion annually** and will mean approximately $7 trillion of lost tax revenue over the next decade. The sheer magnitude of lost revenue is striking: it is equal to 3 percent of GDP, or all the income taxes paid by the lowest-earning 90 percent of taxpayers. As Figure 1 (next page) demonstrates, estimates from academic researchers suggest that **more than $370 billion lost annually is from taxes that the top 10 percent choose not to pay.**

Currently, **an under-staffed IRS**, with outdated technology, **is unable to collect 15 percent of taxes that are owed**, and a lack of resources means that audit rates have fallen across the board. In particular, **audit rates for high-income earners have fallen the most.**

The treatment also includes the following figure:



Fairness Prime

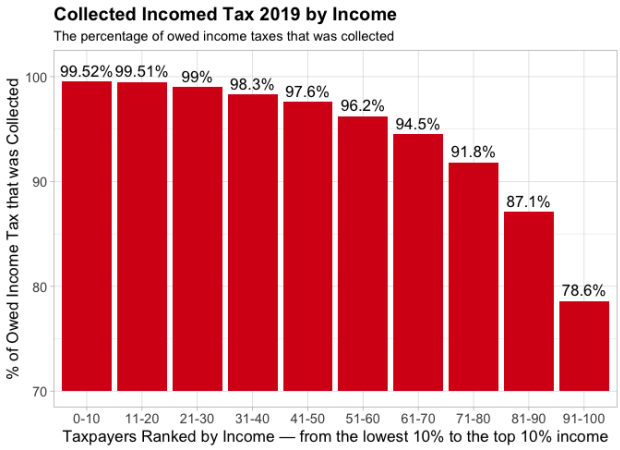
The excerpt below about **how the limited capacity to collect taxes affects income levels differently** was adapted from an article published by the U.S. Department of the Treasury last year. Please closely read the excerpt below before moving on to the next question. You will be able to move to the next question after carefully reading the text below.

The **tax gap** — the difference between taxes that are owed and collected—can be a major source of **inequity**. The taxes for regular wage and salaried workers are

generally automatically deducted from their income. Most regular people report their full income to the IRS. For the rich, the difference between taxes owed and taxes paid is often large. As Figure 1 (next page) demonstrates, estimates from academic researchers suggest that **the richest 10% are often able to avoid a large share of the taxes and choose not to pay almost 22% of the taxes they owed in 2019.**

Currently, **an under-staffed IRS**, with outdated technology, **is unable to collect 15 percent of taxes that are owed**, and a lack of resources means that **audit rates have fallen** across the board. In particular, **audit rates for high-income earners have plummeted the most.**

The treatment also includes the following figure:



Control

To simulate the time respondents spend reading the treatment but without the attendant affect on tax beliefs, we include a placebo text for the Control group. The text describes the growth in internet usage in recent years among Americans – information that is meant to be relevant for the internet survey population but is not expected to affect mood or attitudes related to our outcomes of interest.

American adults spent about 3 hours and 30 minutes a day using the mobile internet in 2019, an increase of about 20 minutes from a year earlier, according to measurement company Zenith. The firm expects that time to grow to over four hours in 2021. We're spending more time online because pastimes like socializing that used to happen offline are shifting online, and we're generally ceding more of our days to digital activities. The overall time Americans spend on various media is expected to grow to nearly 11 hours per day this year, after accounting for declines in time spent with other media like TV and newspapers that are increasingly moving online, according to Zenith. Mobile internet use is responsible for the entirety of that growth.

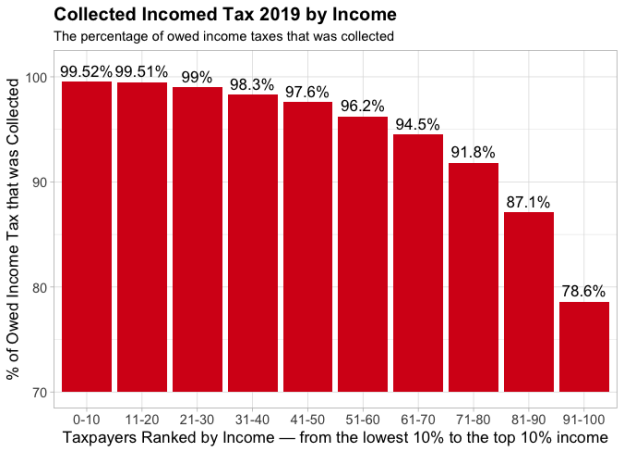
Study 3

Treatment 1

Fairness Prime

> The excerpt below about **how the limited capacity to collect taxes affects income levels differently** was adapted from an article published by the U.S. Department of the Treasury last year. Please closely read the excerpt below before moving on to the next question. You will be able to move to the next question after carefully reading the text below.
>
> The **tax gap** — the difference between taxes that are owed and collected—can be a major source of **inequity**. The taxes for regular wage and salaried workers are generally automatically deducted from their income. Most regular people report their full income to the IRS. For the rich, the difference between taxes owed and taxes paid is often large. As Figure 1 (next page) demonstrates, estimates from academic researchers suggest that **the richest 10% are often able to avoid a large share of the taxes and choose not to pay almost 22% of the taxes they owed in 2019.**
>
> Currently, **an under-staffed IRS**, with outdated technology, **is unable to collect 15 percent of taxes that are owed**, and a lack of resources means that **audit rates have fallen** across the board. In particular, **audit rates for high-income earners have plummeted the most.**

The treatment also includes the following figure:



Control

To simulate the time respondents spend reading the treatment but without the attendant affect on tax beliefs, we include a placebo text for the Control group. The text describes the growth in internet usage in recent years among Americans – information that is meant to be relevant

for the internet survey population but is not expected to affect mood or attitudes related to our outcomes of interest.

> American adults spent about 3 hours and 30 minutes a day using the mobile internet in 2019, an increase of about 20 minutes from a year earlier, according to measurement company Zenith. The firm expects that time to grow to over four hours. We're spending more time online because pastimes like socializing that used to happen offline are shifting online, and we're generally ceding more of our days to digital activities. The overall time Americans spend on various media is expected to grow to nearly 11 hours per day this year, after accounting for declines in time spent with other media like TV and newspapers that are increasingly moving online, according to Zenith. Mobile internet use is responsible for the entirety of that growth.

Treatment 2

IRS Enforcement Increase

> The excerpt below about **the Inflation Reduction Act** was adapted from an article published by the U.S. Department of the Treasury last year. Please closely read it before moving on to the next question. You will be able to move to the next question after carefully reading the text below.
>
> The "Inflation Reduction Act" (**IRA**), signed into law on August 16, includes an **$80 billion appropriation in additional IRS funding** that is projected to increase federal tax receipts by $204 billion over 10 years. Approximately half of **the additional funding is allocated to enforcement**; the other half is allocated to services **and systems modernization**. The Act states that these appropriated funds are to remain available until September 30, 2031.

IRS Enforcement Increase & Yellen Quote

> The excerpt below about **the Inflation Reduction Act** was adapted from an article published by the U.S. Department of the Treasury last year. Please closely read it before moving on to the next question. You will be able to move to the next question after carefully reading the text below.
>
> The "Inflation Reduction Act" (**IRA**), signed into law on August 16, includes **an $80 billion appropriation in additional Internal Revenue Service (IRS) funding** that is projected to increase federal tax receipts by $204 billion over 10 years. Approximately half of the additional funding is allocated to **enforcement**; the other half is allocated to services and systems **modernization**.
>
> U.S. Treasury Secretary Janet Yellen told the IRS that these **additional IRS resources should not** be used to **increase audit rates on taxpayers** making **under $400,000 a year**. In a letter released Wednesday, Yellen said that
>
> "Specifically, I direct that **any additional resources** — including any new personnel or auditors that are hired — **shall not be used to increase the share of small business**

or households below the $400,000 threshold that are audited relative to historical levels. This means that **small business or households** earning $400,000 per year or less **will not see an increase in the chances that they are audited.**"

Control

We again include a placebo text for the Control group to simulate the time spent on the other treatments.

> The excerpt below was published in a reputable US news outlet in the last year. Please closely read the excerpt below before moving on to the next question. You will be able to move to the next question after carefully reading the text below.
>
> More than eight-in-ten U.S. adults (86%) say they get news from a smartphone, computer or tablet "often" or "sometimes," including 60% who say they do so often. This is higher than the portion who get news from television, though 68% get news from TV at least sometimes and 40% do so often. Americans turn to radio and print publications for news far less frequently, with half saying they turn to radio at least sometimes (16% do so often) and about a third (32%) saying the same of print (10% get news from print publications often).
>
> Underneath these numbers lie stark differences by age, with those under 50 showing very different news use patterns than their elders. Americans ages 50 and older use both television and digital devices for news at high rates, while the younger age groups have almost fully turned to digital devices as a platform to access news.