Positive versus Normative Disagreement in Optimal Income Taxation

Matthias Rodemeier and Gregory Sun

October 2025

Abstract

Optimal income taxes depend on how taxpayers respond to taxation and how society normatively values redistribution. We quantify uncertainty in both components by combining a meta-analysis of 1,205 elasticity estimates of the Elasticity of Taxable Income (ETI) with incentivized experiments with 7,500 U.S. adults making real transfers to households at different income levels. We find that normative disagreement about redistribution explains up to 85% of uncertainty in optimal tax rates, substantially more than positive uncertainty about the ETI. Society values transfers to $20,000-earning households 50 times more than to $100,000-earning households, implying tax rates far above current U.S. policy. Yet, even conservatives' redistributive preferences imply 60% average tax rates for high earners, conflicting with the policies they vote for. We show that this is explained by *government aversion*: many people have strong redistributive preferences but aversion to government-mediated redistribution impedes support for progressive taxation. Our findings demonstrate that uncertainty over optimal income taxes stems primarily from normative disagreement about redistribution and the role of government, rather than from behavioral responses to taxes.

Rodemeier: Bocconi University. Sun: Olin Business School, Washington University. We thank Sara Bernardoni for her exceptional dedication as a research assistant. We are grateful to Benjamin Moll for very helpful comments. Rodemeier thanks BAFFI and IGIER for generous support.

1 Introduction

Mirrlees (1971)'s theory of optimal income taxation has provided a powerful framework for understanding how governments should balance efficiency and equity when deciding how much to tax people's income. At its core, it posits that policymakers should balance two key forces: the efficiency costs arising from behavioral responses to taxation (*the positive dimension*), and the gains from redistribution according to society's equity preferences (*the normative dimension*). While economists have made substantial progress measuring behavioral responses through decades of empirical work, the normative dimension has only received traction more recently. Traditionally, economists relegated normative judgment to the domain of policymakers and refrained from systematic empirical inquiry.

This paper's objective is to provide a comprehensive empirical assessment of both ingredients for the optimal income tax schedule, efficiency and equity. We proceed in two steps. First, we implement a large meta-analysis of empirical studies that estimate the Elasticity of Taxable Income (ETI). We develop a novel methodology that allows us to construct a flexible elasticity function, giving an ETI for every income level together with a confidence interval. This confidence interval quantifies the degree of positive disagreement about the ETI at different parts of the income distribution. Positive disagreement tells us how much we can scientifically disagree about the efficiency costs of taxation.

Second, we develop a series of experiments with a representative sample of the US population and elicit redistributive preferences via an incentive-compatible mechanism. We let subjects choose to transfer money to real US households and exogenously vary both the recipient's income and the transfer size. This variation identifies implied social marginal welfare weights (SMWW) for each income level. The distribution of these preferences gives us a measure of normative disagreement about redistribution among US citizens. Normative disagreement tells us how differences in redistributive values translates into disagreement over optimal tax policy.

Our central insight is that optimal taxes are to a large extent determined by normative disagreement rather than positive uncertainty about the ETI. We find that up to 85% of the variance around the optimal marginal tax rate of households earning less than $100k (two-third of the US population) is driven by variance in the SMWW rather than variance in the ETI. The story is more nuanced at the very top: disagreements on how much society should value the welfare of the rich and scientific uncertainty over how the rich adjust their income play similar roles in determining uncertainty over how to tax them. Importantly, we also find that progressive welfare weights do *not* result in support for progressive income taxation: although most people value the welfare of the poor much more than the welfare of the rich, they are averse to *government-mediated* redistribution via higher taxes. This aversion creates a gap between people's strong

generosity toward the poor in transfer choices and their limited support for progressive income taxation.

Bayesian Meta-Analysis of the ETI In the first part of the paper, we develop a novel Bayesian meta-analysis framework using Gaussian Process methods to synthesize the disparate elasticity estimates from the literature into a continuous elasticity function $\varepsilon(z)$. This approach not only produces a best estimate of the elasticity at each income level but, crucially, quantifies the *uncertainty* around these estimates. We can reject with high statistical certainty the common assumption in the literature that the ETI is constant across the income distribution.[1] Our meta-analysis covers 1,205 hand-coded elasticity estimates across the income distribution from 74 studies, and produces posterior distributions that capture both sampling uncertainty and heterogeneity across contexts. We find that the ETI is around 0.3 at the bottom of the income distribution, decreases to below 0.15 in the middle where roughly 60% of U.S. households are located, and rises modestly again to about 0.2 in the upper tail. Uncertainty about the true ETI is lowest in the middle of the distribution, where 90% Bayesian credibility intervals are only about ± 0.04 wide. In the tails, however, uncertainty explodes: the median elasticity for the poorest households is 0.3, but 90% credibility intervals include values as large as 0.5 and as low as 0.19. For the richest households, the median elasticity is 0.2, but could plausibly range from 0.05 to 0.37.

Experiments In the second part of the paper, we conduct a series of large-scale experiments with a representative sample of the US population (N =7,432) to directly elicit redistributive preferences. Strikingly, political ideology shapes not just the level but also the nature of redistributive concerns: liberals' transfer choices are most responsive to exogenous variation in the recipient's income and least responsive to the transfer size. Conservatives show the opposite pattern, caring most about transfer amounts and least about recipient income. This points to fundamental philosophical differences about whether transfers should address inequality or efficiency. We then use this data to estimate the function of SMWW: On average, society values a dollar transferred to a family earning $20,000 annually 2.5x as much as to a family earning $40,000; 8x as much compared to a family earning $80,000; and almost 50x as much compared to a family earning $100,000. Families earning above $300,000 receive a SMWW of zero, meaning they should receive the revenue-maximizing marginal tax rate (i.e., the rate that lies on top of the Laffer curve).

[1] However, we show theoretically that the constant-elasticity specifications used in individual studies can be interpreted as local approximations to a global elasticity function, justifying our synthesis approach.

Optimal Income Taxes Combining the elicited welfare weights from the experiments with the ETI function, we estimate the optimal tax schedules and the relative importance of behavioral versus normative uncertainty. Comparing the 10th to 90th percentile of redistributive preferences, we find that optimal marginal tax rates differ by, on average, 15 percentage points. Disagreement over these tax rates is especially important at the bottom: For the 67% of Americans earning below $100,000, normative disagreement accounts for up to 85% of the variance in optimal marginal tax rates. These rates change by approximately 20 percentage points due to differences in redistributive preferences. Within this same income range, uncertainty about the ETI is not large and generates tight confidence bands of only ±5 percentage points. These patterns reverse at the top: above $150,000 in income, the confidence intervals from behavioral uncertainty become very wide. The most redistributive individuals would implement marginal rates reaching around 90% for top earners, while the least redistributive prefer rates closer to 70%. Both rates are substantially more progressive than current US tax rates.[2]

Progressive Behavior versus Preferences for Taxation While the implied income tax schedule is more progressive than the current US tax code, we find that it also conflicts sharply with hypothetical tax rates people say they consider as ideal in survey questions. Even Republicans' transfer choices implicitly suggest optimal average tax rates of 60% for high earners, yet when directly asked about ideal top tax rates, their average response is merely 18%. A follow-up experiment uncovers *government aversion* as the source of this divergence between transfer choices and political tax views. In the experiment, participants choose transfers to poor recipients but a randomized share truthfully goes to the U.S. Treasury. We make choices consequential via a payment portal offered by the Treasury that accepts donations. Strikingly, people are either indifferent between or explicitly prefer giving less money directly to poor recipients over choosing options that deliver more to the poor but also more to the Treasury. For example, conservatives would rather transfer $50 directly to a recipient than choose an option delivering $60 to recipients while sending $40 to the government. Liberals show no such aversion to Treasury-generating options, though they also do not value government revenue positively, perhaps because the current government is Republican. Allowing for this government aversion in the estimation of the optimal income tax schedules yields much less progressive tax rates, and closes most of the gap between redistributive preferences and preferences for progressive taxation. These results point to an important insight: people across the political spectrum generally put much higher weight on the welfare of the poor than the rich, but an aversion to government-mediated redistribution

[2]In a second experimental wave, we show that information about efficiency reduces redistribution by shifting attention to costs, while information about others' redistributive tastes increases it by correcting belief misperceptions, in line with some earlier results in the literature (Stantcheva, 2021). The effects of information are not large enough to yield substantially different income taxes.

impedes more progressive income taxes.

Main Takeaways These findings deliver three main insights for the future of tax policy research and debate. First, for the vast majority of the income distribution, disagreements about tax rates are fundamentally disagreements about norms rather than tax elasticities. Second, the concentration of behavioral uncertainty at the income extremes reveals a critical irony: we know the least about how taxes affect those who should contribute the most to tax revenues (the rich) and those who most depend on the tax-and-transfer system (the very poor). This illustrates where future empirical work on the ETI could have the greatest policy impact. Third, many people are generous towards the poor, but do not consider taxation as the solution to reduce inequality.

Literature Our work relates to two strands of literature studying behavioral responses to taxation and redistributive preferences. Since Saez (2001), there has been a large set of studies that quantify the ETI in order to identify the optimal tax schedule. Saez et al. (2012) provides an overview of this literature. A particularly influential approach has been to identify local elasticities of taxable income by studying behavior around discontinuous changes in the tax rate (Saez, 2010; Kleven and Waseem, 2013). The maturity of the literature on the ETI calls for a systematic synthesis of local estimates into one global elasticity function. Our Bayesian approach contributes this synthesis to the literature and also identifies parts of the income distribution where estimates are missing and uncertainty is largest.[3]

Tax elasticities have also been used to understand the normative component of optimal income taxation in the "inverse optimum" approach (e.g., Bargain et al., 2011, 2014; Lockwood and Weinzierl, 2016; Hendren, 2020). The idea is to find the set of social welfare weights that would make the status quo tax policy welfare-maximizing. However, researchers rarely interpret resulting estimates as society's true preferences, as the approach relies on the strong assumption that tax policy has been set optimally, which is likely violated due to political economy constraints and other frictions. Saez and Stantcheva (2016) introduce a general theory of Social Marginal Welfare Weights that allows for ideals like equality of opportunity or poverty alleviation to shape weights alongside traditional utilitarian calculus. Complementary works has used surveys to understand motives behind support for redistribution (e.g., Alesina and La Ferrara, 2005; Alesina et al., 2018; Kuziemko et al., 2015; Mengel and Weidenholzer, 2023; Cappelen et al., 2024). Stantcheva (2021) provides qualitative evidence that normative views are more

[3]Different from Chetty (2012) who studies a small set of selected elasticities in the context of optimization frictions, we consider the entire literature with temporal and regional variation. We do not just focus on US estimates as US estimates can differ substantially along the temporal dimension, with some estimates coming from the 1960s. It is unclear whether two studies are more comparable when they are, say, 30 years apart from each other but come from the same country, or when they happened at the same time but in two similar high-income countries (e.g., the US versus the UK).

important for the support of redistributive taxes than efficiency concerns. Our approach complements this evidence by quantifying the importance of social welfare weights relative to tax efficiency. Different from prior work, this allows our results to be used as direct inputs into optimal tax formulas.

The complementary working paper by Capozza and Srinivasan (2025) uses a similar elicitation method to quantify social marginal welfare weights from transfer choices. Our paper advances this work on two important dimensions. First, we combine elicited welfare weights with an estimated ETI function from over 1,200 estimates and propagate both sources of uncertainty into optimal marginal tax rates. This quantifies how much uncertainty exists about optimal taxes and whether it is *positive or normative*—the unique questions of our paper. Second, our experiments show that transfer-choice–based welfare weights overstate support for progressive tax schedules because people are averse to government. Another unique contribution of our paper is, therefore, the finding that estimated welfare weights do not generally produce *politically feasible* income tax schedules.

Section Overview In the next Section, we briefly refresh the reader's knowledge of the Mirrleesian theory of optimal income taxation that motivates our work. In Section 3, we present our Bayesian Meta-Analysis of the ETI and estimate the entire function of elasticities across the income distribution. We then present experiments in Section 4 that elicits social welfare weights among a representative US sample. Section 5 estimates the optimal income tax schedule together with a measure of uncertainty, leveraging our elasticity function and welfare weight function. Section 6 presents a final experiment that uncovers why people's behavior is much more progressive than the policies they vote for. We conclude in Section 7.

2 Revisiting the Mirrleesian Framework of Income Taxation

To motivate the rest of this paper, it is useful to briefly revisit Mirrlees (1971)'s canonical optimal income taxation framework. Individuals in the population differ by a type $\theta > 0$. Each individual chooses a pre-tax income $z \geq 0$. For a given nonlinear tax schedule $T(z)$ with corresponding marginal rate $T'(z)$, after-tax income is $c = z - T(z)$. Individuals trade off consumption utility and labor disutility, so utility is given by $U(c, z; \theta) = u\big(c\big) - \theta v(z)$, where $u'(c) > 0$, $u''(c) \leq 0$ captures diminishing marginal utility of consumption, and $v'(z) > 0$, $v''(z) > 0$ encodes the increasing marginal disutility of labor. Fixing the tax schedule, these choices generate a distribution of pre-tax incomes, denoted by the density $h(z; T(\cdot))$ on $[0, \infty)$. When obvious from the context, we will not explicitly include the dependence of h on the underlying tax rule in our notation.

The social planner sets a nonlinear tax schedule $T(z)$ to maximize a social welfare function $W(T(z))$ subject to an exogenous tax revenue requirement E:

$$\max_{T(\cdot)} W\big(T(\cdot)\big) = \int_0^\infty \Gamma\big(u\big(z - T(z)\big) - \theta v(z)\big)\, h(z; T(\cdot))\, dz, \quad \text{s.t.} \quad \int_0^\infty T(z)\, h(z; T(\cdot))\, dz \geq E,$$

Redistributive preferences are captured by a social marginal welfare weight $G(z) \equiv \Gamma'(u\big(z - T(z)\big) - \theta v(z))u'(z - T(z)) \geq 0$ of individuals with pre-tax income z. Intuitively, given an existing tax schedule $T(\cdot)$, $G(z)$ quantifies the marginal benefit to the social planner of giving an extra dollar to someone who currently has pre-tax income z. To make the dual dependence on normative weights and behavioral elasticities transparent, define the *average* social marginal welfare weight above income z as:

$$\bar{G}(z) \equiv \frac{\int_z^\infty G(s)\, h(s)\, ds}{1 - H(z)}.$$

It is well known that at an optimum, the marginal tax rates satisfy

$$T'(z) = \frac{1 - \bar{G}(z)}{1 - \bar{G}(z) + \alpha(z) \cdot \varepsilon(z)}, \tag{1}$$

where $\alpha(z) = \frac{zh(z)}{1-H(z)}$ is the Pareto parameter measuring how thick the tail of the income distribution is. Both positive and normative considerations enter into Equation (1).

The *positive considerations* enter through the elasticity of taxable income,

$$\varepsilon(z) \equiv \frac{\partial z}{\partial(1 - T'(z))} \frac{1 - T'(z)}{z},$$

which governs how z adjusts when the net-of-tax rate $1 - T'(z)$ is perturbed. Higher elasticities imply that taxes are distorting labor supply more. The optimal marginal tax rate is consequently decreasing in the elasticity.

Normative concerns enter the optimal taxation problem through $\bar{G}(z)$: In a *utilitarian* benchmark, $G(z)$ is constant; in *Rawlsian* or *generalized egalitarian* frameworks, $G(z)$ is strictly decreasing. Lower values of $G(z)$ for higher values of income z mean that society places less weight on the welfare of the wealthy. This implies a higher marginal tax as income increases.

Crucially, scientific uncertainty about $\varepsilon(z)$ or disagreement about redistribution $G(z)$ translate directly into uncertainty about optimal policy. Our empirical strategy separately quantifies the variation in each component to understand their relative importance.

3 Measuring Positive Uncertainty: A Bayesian Meta-Analysis

The behavioral responsiveness of taxable income to marginal tax rates plays a central role in optimal tax policy. To inform the elasticity term $\varepsilon(z)$ in our tax formula, we conduct a comprehensive Bayesian meta-analysis of the existing empirical literature. Our contribution is to synthesize hundreds of scattered point estimates into a coherent, smooth elasticity function $\varepsilon(z)$ defined over the income distribution, together with a full quantification of uncertainty.

3.1 Data Collection

We compile a hand-coded database of 1,205 elasticity estimates from 74 empirical studies. The reason one study has multiple estimates is both because they can come from different parts of the income distribution and because many papers analyze different subsamples (e.g., self-employed workers). Studies span both developed and developing countries. We first instructed research assistants to collect the universe of studies that exploit nonlinearities in tax schedules—tax kinks and notches—for identification of the ETI. These estimates are often well-identified and can be directly linked to a position in the income distribution. Of our total sample, 871 elasticities are identified of these nonlinearities. Next, we instructed research assistants to search for papers that exploit tax reforms for identification. This added 334 elasticity estimates. Many of those estimates add important information about the ETI to the far right part of income distribution where there is no change in the marginal tax rate (i.e., no kink or notch). All studies had to report an elasticity estimate $\hat{\varepsilon}_{si}$, a standard error $\hat{\sigma}_{si}$, and a clear mapping to a point or bin in the income distribution z_{si}.

For each estimate, we record contextual variables including the country, tax system, policy reform exploited, year, estimation method, and whether the elasticity refers to taxable income or broader income measures. We also record a variable indicating which of the estimates at a given income bin is the authors' preferred estimate, which we later use as our main estimate for that bin within a study. We will make this database of tax elasticities publicly available to the academic community once this paper is published, and aim to enrich it further with additional detailed information that can be used by other scholars in future work.

3.2 Raw Data

Panel A of Figure 1 presents a boxplot for each income quartile, together with the raw point estimates of the ETI. Using the income percentile rather than the level of income allows us to easier compare elasticities from different countries with different levels of wealth. We harmonize income measures by converting all estimates into U.S. dollars at Purchasing Power Parity

using World Bank exchange rates and mapping reported incomes to percentiles of the US income distribution. As a robustness check, in Appendix B, we also re-estimate results locating each observation by its percentile within the national income distribution rather than the U.S. distribution and obtain similar patterns. The majority of our estimates come from the US and the Euro zone. The other large group of countries are developed economies (Denmark, the UK, and New Zealand). The only emerging economies we have in the data are China, Cyprus and Pakistan; the latter coming from the seminal study by Kleven and Waseem (2013). In Appendix B, we study the robustness of our results to excluding developing countries and reach very similar conclusions.

The first notable takeaway from this figure is that the number of point estimates changes substantially with the income quartile. The first quartile has 621 estimates, the second 279, the third only 97, and the fourth 208 estimates. Median elasticities in each quartile follow a U-shape and are, respectively 0.2, 0.14, 0.05, and 0.29. While there is substantial heterogeneity in the ETI within quartile, most values are far below 1.

Panel B is a binscatter plot of the main elasticity estimate of each study against the location in the income distribution. The main estimate is the one that authors state as their preferred estimate given a point in the income distribution. If authors state multiple estimates at the same location on the income distribution without specifying which estimate is the preferred one, we plot the average of those estimates. We bin estimates into 10 bins of equal width in percentile terms. One bin (from the 70th to the 80th percentile) did not have any observations.

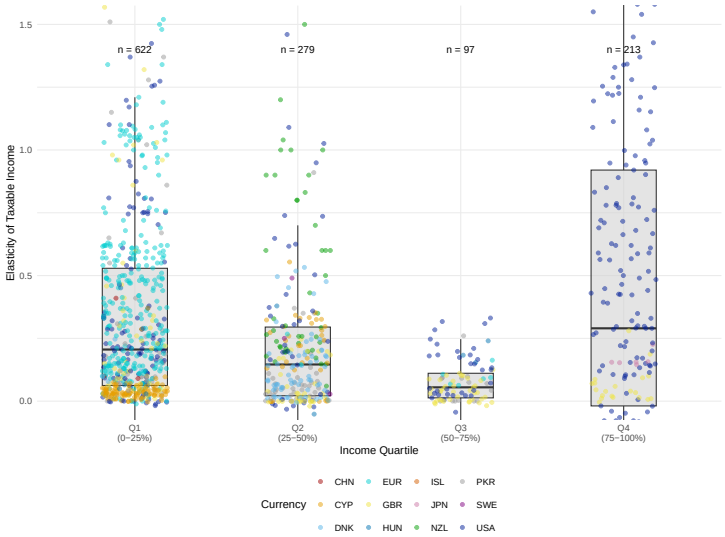
Uncertainty tends to be smallest between the 30th and 50th percentile and relatively large in the tails. High uncertainty around the 70th percentile is explained by the lack of sufficient studies in this range of the income distribution. By contrast, high uncertainty in the far right tail exists despite a large number of studies and is rather explained by more noisy estimates. Almost all studies in the right tail exploit *tax reforms* to identify behavioral responses to taxation. These studies are often constructed on small sample sizes and therefore provide much less precise ETI estimates than studies exploiting local behavioral responses around kinks and notches.

We now show how we translate this collection of estimates into continuous elasticity function and quantify uncertainty around this function.

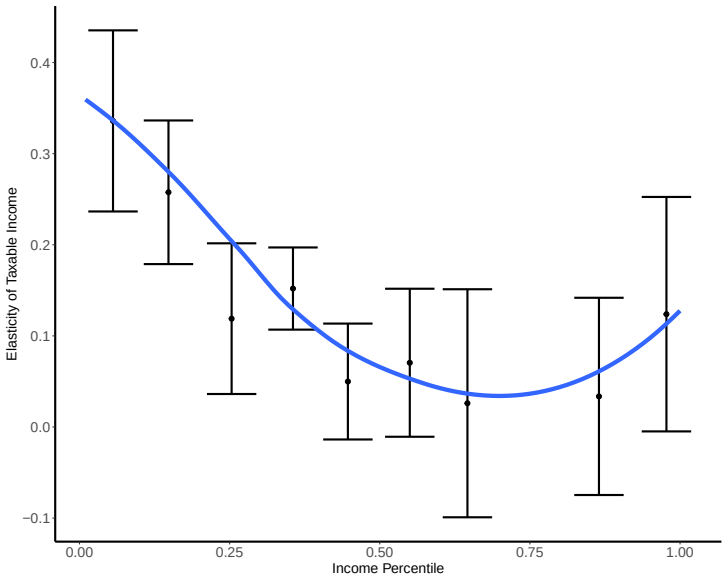
3.3 From Scattered Elasticities to a Global Elasticity Function

The end product of our meta-analysis is a function $\varepsilon(\cdot)$, where for each fixed income level z, $\varepsilon(z)$ represents the best guess of the ETI conditional on an individual having income z, given estimates from the data. Unlike a more standard meta-analysis, the object of interest that we aim to recover here is not just a *number*, but instead, an entire *continuous function*. In order to

Figure 1: Raw Data



(a) Elasticity Estimates across the Income Distribution



(b) Binscatter of Main Elasticity Estimates

Notes: This figure summarizes the sample of elasticity estimates. Panel A presents boxplots of all point estimates for each income quartile. Panel B presents the estimates that authors state as their main or preferred estimate. Each point represents a point estimate of elasticity on the vertical axis and the percentile rank of the income under consideration on the horizontal axis. Each point also is plotted with an associated 95% confidence interval. Colors represents the country under study.

accomplish this, we adopt a Bayesian approach that borrows from an extensive machine learning literature on Gaussian processes (GP).[4]

[4]For more background on the statistical theory behind Gaussian Processes, see, e.g., the review article by MacKay et al. (1998). There is precedent outside of economics for using Gaussian processes as a basis for specifying a prior for the purpose of conducting a meta-analysis. For example, see Salimi-Khorshidi et al. (2011) for an

Here, we provide intuition about what our Bayesian meta-analysis approach does and defer a full discussion of technical details to Appendix A. Just like more standard Bayesian frameworks allow the researcher to place a prior on random variables/vectors, Gaussian processes parameterize priors over random functions. Importantly, GP priors have hyperparameters that encode prior knowledge that the underlying function we wish to estimate is *smooth*. Intuitively, smoothness of the underlying elasticity function $\varepsilon(\cdot)$ implies that getting a signal about $\varepsilon(z)$ is also informative about $\varepsilon(z')$ for a point z' close to z. In our dataset, we observe ETIs at specific income levels, so this idea gives us a principled way to "glue together" evidence from a discrete set of income levels to form conclusions about the global elasticity function.

Our basic approach is as follows. First, we specify a diffuse GP prior on the space of possible ETI functions. Second, we derive the likelihood function of a given observation in our dataset. With a specification of the prior and the likelihood, we then use a standard software implementation of GP models to (numerically) apply Bayes Rule and sample draws from the posterior distribution over possible ETI functions. The average elasticity function across these draws provides us with a point estimate of the underlying ETI function, while the variance in the resulting elasticity function across different draws allows us to quantify uncertainty.

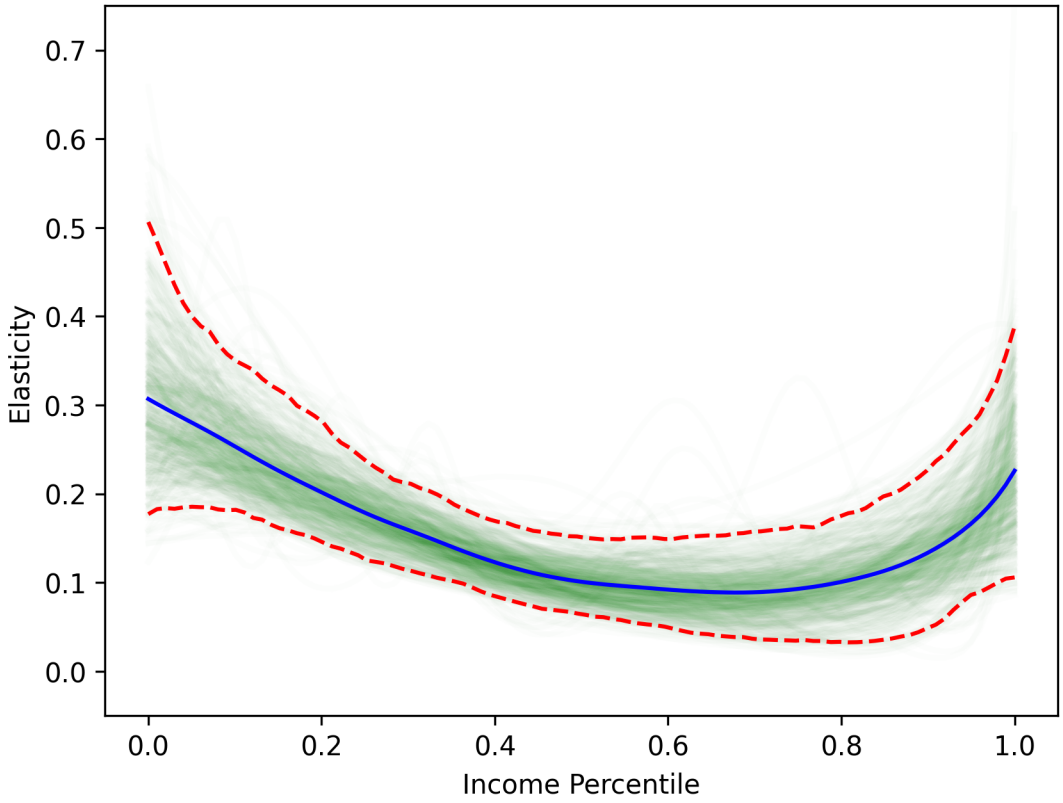
The results from doing this are contained in Figure 2. To construct this figure, we draw 100 realizations of the ETI function from the posterior distribution given our sample of ETI estimates. Each draw from the posterior is depicted as a green curve. The blue curve is the point-wise mean of the elasticity function across the posterior sample. Red dotted lines show 90% Bayesian credible intervals of this mean function. Reassuringly, the mean posterior agrees well pointwise with the reduced-form smoothed regression point estimates in Panel B of Figure 1. Moreover, the Bayesian approach provides us with additional information that the reduced-form summary of the data in the previous section cannot: it allows us to make inferences about the pattern of elasticities across the entire income distribution, not just at a single point. This will be important in later sections, when we use the inferences derived in this section about optimal tax policies.

At the bottom of the income distribution, the elasticity is around 0.3 and decreases slowly to 0.1 as we reach the median income. Above the median, the function increases once again to 0.21. The Bayesian credible intervals reject the typical assumption in the literature that the elasticity is constant.

Uncertainty about the elasticity is largest in the tails. For instance, we cannot exclude with 90% certainty that the elasticity of 0.4 at the bottom could also be 0.19 or 0.5. In the right tail, above the 75th percentile, elasticities might be as low as 0.05 or as large as 0.37. The right tail is particularly noisy because i) there are no studies around the 75th percentile, and ii) the studies

example from neuroscience.

Figure 2: Posterior of Elasticity of Taxable Income Function



Notes: This figure summarizes the posterior distribution of the elasticity function $\varepsilon(r)$ given the full sample. The light green lines represent 100 draws from the posterior distribution. The solid blue corresponds to the pointwise mean of the elasticity function across our posterior sample while the dotted red lines represent pointwise 90% Bayesian credible intervals.

that exist further up in the tail provide very noisy estimates. By contrast, we have substantially more certainty over elasticities between the 20th and 60th percentile of the income distribution, which also represents the majority of the US population. Here, the credibility bounds move by only +/- 0.04 around the mean.

In order to obtain the optimal income tax schedule, we further need to identify society's welfare weights for each income level, which we do next.

4 Measuring Normative Disagreement: Experimental Evidence

Conceptual Framework We model subjects in our experiments as making choices about their preferred redistribution according to marginal social welfare weights. Following the canonical Mirrlessian framework we parameterize welfare weights by a CARA function in other people's baseline consumption. We also explore a CRRA function in the Appendix E for robustness.

Let i index an individual subject. Then, the marginal social welfare weights of individual i are assumed to be of the form

$$g_i(C_j) = \exp(-\gamma_i C_j) \tag{2}$$

where C_j is baseline consumption of *another* household in the population, denoted household j. γ_i is the absolute risk aversion parameter of individual i, measuring how preferences for household j change as j's consumption level increases. If $g_i(\cdot)$ is concave, the subjects puts larger welfare weights on poorer households.

Identifying Welfare Weights from Choice Data We designed three experiments with a total of 7,432 subjects to elicit social welfare weights from a representative sample of the U.S. population. Data collection started in April 2025 and was done through Prolific, a commonly used platform in economics for running experiments and surveys. Below, we discuss the first experiment and elaborate on the follow-up experiments later. The core of our methodology relies on pairwise choice tasks that reveal how individuals trade off transfers to households at different income levels.

Each participant completed 10 pairwise choice scenarios. In each scenario, respondents were presented with two potential transfers to two different U.S. households and asked to choose which transfer they preferred to implement. The scenarios varied along two dimensions:

1. **Transfer amounts**: T_{iq1} and T_{iq2} for households 1 and 2, respectively. We randomly draw transfers from the interval \$100 to \$2,000 with equal probability.

2. **Recipient income levels**: The current after-tax income C_{iq1} and C_{iq1}. These values are drawn from the interval \$5,000 to \$500,000 per year, in \$5,000 increments.

An example of a choice scenario is given in Figure 3 where a subject could either allocate a transfer of \$1,700 to a household with net income of \$80,000 per year, or a smaller transfer of \$900 to a richer household earning \$150,000 per year. Transfers and income were randomly drawn for each choice set and could also include larger transfers to richer households.

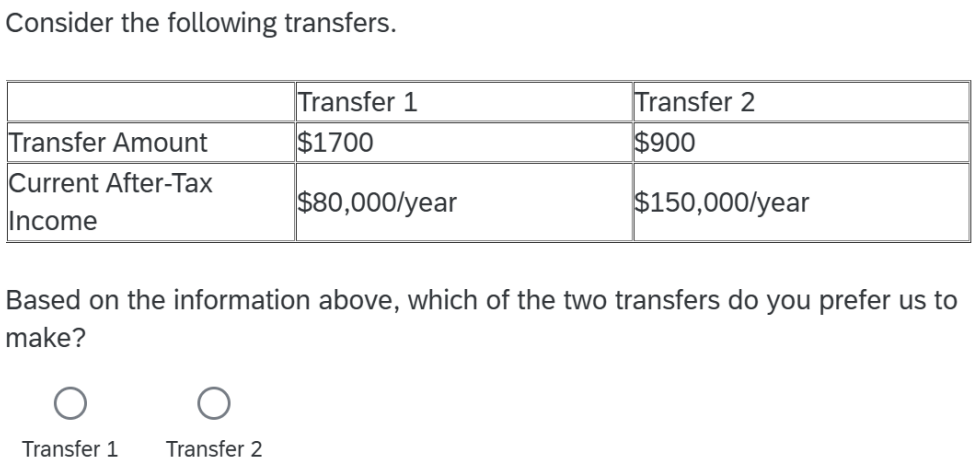
If the subject answers this question according to her redistributive preferences, she would transfer to household 1 over household 2 if and only if

$$T_1 \exp(-\gamma_i C_1) - T_2 \exp(-\gamma_i C_2) > 0$$

We allow for the possibility that there is some cognitive noise in choices, so in reality, the subject chooses redistribution to household 1 if and only if

$$T_1 \exp(-\gamma_i C_1)\varepsilon_{iq1} > T_2 \exp(-\gamma_i C_2)\varepsilon_{iq2}.$$

Figure 3: Example of Choice Screen


Consider the following transfers.

	Transfer 1	Transfer 2
Transfer Amount	$1700	$900
Current After-Tax Income	$80,000/year	$150,000/year

Based on the information above, which of the two transfers do you prefer us to make?

Transfer 1 Transfer 2

We assume that for each individual i, $\log \varepsilon_{iq1}, \log \varepsilon_{iq2} \overset{i.i.d.}{\sim} T1EV(\sigma_i)$. Here, σ_i is a scale parameter on the T1EV distribution, which may vary across individuals.

Taking logs and dividing both sides of the inequality by σ_i, the above is equivalent to individuals choosing to redistribute to family 1 if and only if

$$-\frac{\gamma_i}{\sigma_i}(C_1 - C_2) + \frac{1}{\sigma_i}[\log(T_1) - \log(T_2)] + \frac{1}{\sigma_i}[\log \varepsilon_{iq1} - \log \varepsilon_{iq2}] \geq 0.$$

We divide by σ_i in deriving the above expression because the distributional assumption on the ε's implies that the last term above follows a standard logistic distribution, hence our model boils down to a fairly standard random coefficients logit model without an intercept. The absolute value of the ratio between the random coefficient in front of $(C_1 - C_2)$ and the random coefficient in front of $[\log(T_1) - \log(T_2)]$ represents individual-level redistributive preference, γ_i.

Incentive Compatibility To make choices consequential, participants were informed that at the end of the study, we would randomly select two families from the U.S. Prolific sample (not participating in the experiment) with income levels matching those in one of the scenarios. One of these families would receive the exact transfer amount chosen by a randomly selected participant, while the other would receive nothing (see also Capozza and Srinivasan, 2025, for a similar design). To assure that subjects understood that their choices had real consequences in expectation, we included i) an extensive explanation of the incentive mechanism, and ii) a comprehension check about the consequential nature of the transfer choices (see Appendix I).[5]

[5]We constructed a comprehension check that was relatively difficult to pass as it involved four possible answers that were long to read and where some answers only differed slightly. While this may cause some subjects that understood instructions but were too impatient to read through all options to fail the test, our conservative test

Additional Questions The experiment further elicited a wide range of characteristics such as income, employment, education, age, and political preferences.

We also included attention checks throughout the experiment. Subjects who failed these checks were disqualified from the experiment and are not included in the analysis. As is common for experiments on Prolific, the data quality is high and only 2 subjects failed any attention check.[6]

4.1 Results

Our sample size of the first experimental wave consists of 2,421 subjects. Before illustrating the estimation of social welfare weights, we provide a set of reduced-form results.

Sensitivities to Recipient Income and Transfer Size Our first question is to what extent differences in income of and transfers to families matter to subjects. To answer this, we estimate the linear probability model:

$$Y_{i1} = \alpha + \beta \Delta I_i + \beta_2 \Delta T_i + \epsilon_i \tag{3}$$

where $Y_{i1} \in \{0, 1\}$ is one if subject i chose the transfer for family 1. The variables $\Delta I_i = I_{i1} - I_{i2}$ and $\Delta T_i = T_{i1} - T_{i2}$ are the differences in income and transfers between the two families.

Panel A of Figure 4 reports the coefficient on the transfer amount, β_1, for the average subject and also across the political spectrum. On average, a \$1,000 increase in the transfer amount to family 1 increases the likelihood for that family to receive the transfer by 4 percentage points. The most liberal survey participants respond the least to transfer variation, with a sensitivity of 2 percentage points. As we move further to the political center, transfer sensitivity increases to 4.8pp for Liberals, and 3.8 both for Moderate and Conservatives. Subjects identifying as very conservative are most responsive to transfer size, with a sensitivity coefficient of 6.7pp.

Panel B shows transfer sensitivity, β_2. The likelihood for family 1 to receive a transfer decreases on average by 0.12pp if their income increases by \$1,000. Interestingly, we observe the reversed pattern across the political spectrum that we found in Panel A: Sensitivity decreases as we move from the political left to the right. The most liberal subjects have a sensitivity of -0.16pp. Moving further to the right politically reduces this negative income sensitivity to -0.13 (Liberals), -0.12 (Moderate), -0.1 (Conservative), and -0.08 (Very Conservative). These results

allows us to condition on the subsample that truly engaged and understood the instructions. 73% of all subjects chose the correct answer, which is typical on these platforms for comprehension (rather than attention) checks.

[6] We also included an open-text question asking participants to describe the main factor influencing their transfer decisions. We do not report these results here, as they provide limited additional insight beyond the choice data. Most participants cited either the transfer size or income as the primary determinant of their decisions.

imply that the political left cares relatively little about transfer size while being very sensitive to income. The opposite is true for the political right, whose choice is mostly determined by transfer size rather than income. Philosophically, this insight means that those on the political left prioritize redistributive concerns based on need (as measured by income levels) over efficiency considerations (transfer size), and view transfers primarily as a tool for addressing inequality. This fundamental difference reflects competing visions of the welfare state between Democrats and Republicans.

These results provide the first confirmation that choices in the experiment are meaningful. Subjects choose broadly in line with the political agenda of the party they support. Another takeaway is that redistribution is preferred across *all* political positions, as every income coefficient is statistically different from zero.

Implied Welfare Weights We now show how these reduced-form results translate into implied social welfare weights. We first derive the sampling distribution of the data conditional on individual level parameters γ_i and σ_i from our CARA specification in equation 2. The experiment randomly samples income and transfer amounts for family 1 and 2, (C_1, C_2, T_1, T_2). Since we have 10 questions per subject, let the corresponding 10 draws be denoted by

$$\mathbf{X}_i \equiv (C_{1i}^1, C_{2i}^1, T_{1i}^1, T_{2i}^1, \ldots, C_{1i}^{10}, C_{2i}^{10}, T_{1i}^{10}, T_{2i}^{10}).$$

In each scenario, the responent decided whether to transfer to family 1, Y_{iq}, yielding a vector of dummy variables

$$\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{i10}),$$

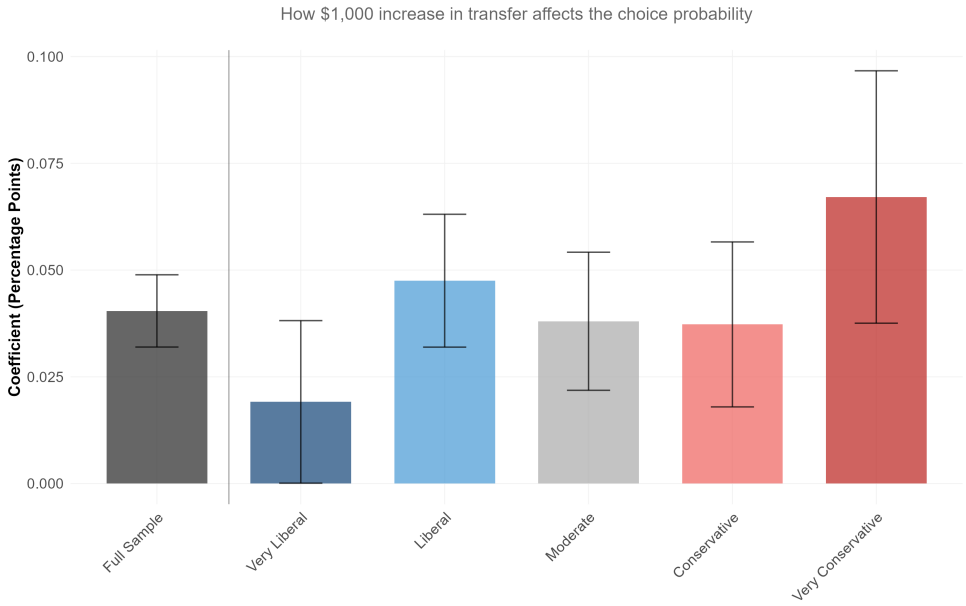
The likelihood function conditional on $\mathbf{X}_i$ and the individual-level parameter values is given by

$$\mathcal{L}(\mathbf{Y}_i|\mathbf{X}_i, \gamma_i, \sigma_i) = \prod_{q=1}^{10} \frac{\exp\left(\frac{-\gamma_i(C_{1i}^q - C_{2i}^q) - \log(T_{1i}^q - T_{2i}^q)}{\sigma_i}\right)}{1 + \exp\left(\frac{-\gamma_i(C_{1i}^q - C_{2i}^q) - \log(T_{1i}^q - T_{2i}^q)}{\sigma_i}\right)}$$

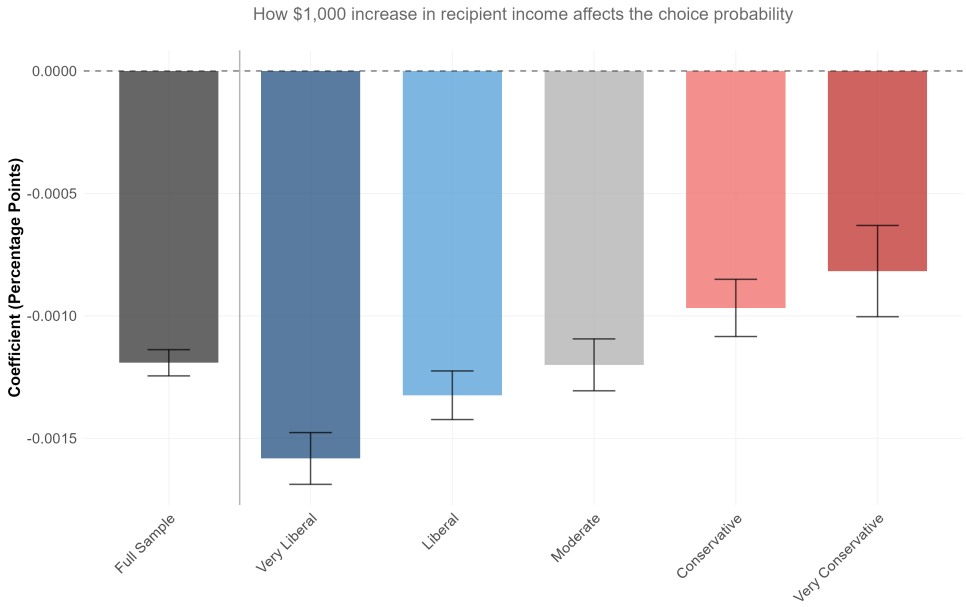
To estimate the distribution of γ_i as flexibly as possible, we approximate the nonparametric joint distribution of (γ_i, σ_i) as being supported on a fine grid of values, and estimate this model using a variant of the approach proposed in Fox et al. (2011) (FKRB).

The basic idea behind the FKRB estimator is to parameterize the joint distribution of (γ_i, σ_i) as being supported on some grid of the form $\boldsymbol{\gamma} \times \boldsymbol{\sigma}$. Given this support restriction, the joint distribution can be represented by the probability mass placed on each support point. Specifically, let $\boldsymbol{\theta}$ be a $|\boldsymbol{\gamma}| \times |\boldsymbol{\sigma}|$, where each entry $\theta_{a,b}$ represents the probability $\Pr[\gamma_i = \boldsymbol{\gamma}_a, \sigma_i = \boldsymbol{\sigma}_b]$. The laws of probability imply that $\theta_{a,b} \geq 0$ and $\sum_{a,b} \theta_{a,b} = 1$, but no further constraints on the

Figure 4: Redistributive Preferences



(a) Transfer Sensitivity



(b) Income Sensitivity

Notes: This figure plots two dimensions of redistributive preferences. In Panel A, we plot the effect of raising family 1's income by $1,000 on the probability that subjects in the experiment choose the transfer for family 1. We plot regression coefficients in percentage points and plot both the effect on the full sample (purple), as well as on the political spectrum. Panel B, does the same exercise for an exogenous increase in the income of family 1 by $1,000.

entries of $\boldsymbol{\theta}$. For a given guess of $\boldsymbol{\theta}$, the likelihood function for a given participant is thus given by

$$\mathcal{L}(\mathbf{Y}_i|\mathbf{X}_i, \boldsymbol{\theta}) = \sum_a \sum_b \theta_{ab} \mathcal{L}(\mathbf{Y}_i|\mathbf{X}_i, \gamma_a, \sigma_b).$$

The basic idea behind our estimation procedure is thus as follows. First, using a procedure detailed below, we construct a data-driven set of gridpoints, $\boldsymbol{\gamma}$ and $\boldsymbol{\sigma}$. Second, we estimate $\boldsymbol{\theta}$ by maximizing a penalized log-likelihood, subject to the constraints on $\boldsymbol{\theta}$ described above:[7]

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta},\sigma} \sum_{i=1}^{N} \log \mathcal{L}(\mathbf{Y}_i|\mathbf{X}_i, \boldsymbol{\theta}) + \lambda \sum_{a=2}^{41} \left(\sum_b \theta_{a,b} \right)^2 \quad \text{s.t.} \quad \theta_{a,b} \geq 0, \forall a, b, \sum_{a,b} \theta_{a,b} = 1.$$

Choice of Grid To apply an FKRB-style estimator, we must choose the grid of points which (γ_i, σ_i) is supported on. There is little formal guidance in the literature on how to do this, so we pick a data-dependent procedure based on what worked well in a pilot version of our study. In the first step, we fit a standard logistic regression without an intercept. The outcome variable is an indicator for whether or not redistribution to family 1 was chosen, and the regressors are $(C_1 - C_2)$ and $[\log(T_1) - \log T_2]$.

Let $\bar{\gamma}$ be the ratio of the coefficients from this regression, and let $\bar{\sigma}$ be the reciprocal of the coefficient in front of $[\log(T_1) - \log T_2]$. Let $\tilde{\boldsymbol{\gamma}}$ be a grid of 40 points that is evenly spaced in log units between $\bar{\gamma}/10$ and $40\bar{\gamma}$. Further, let $\boldsymbol{\gamma} = (0, \tilde{\boldsymbol{\gamma}}, 100)$. Following our pre-registration we restrict our analysis to $\gamma_i \notin \{0, 100\}$. The case of $\gamma = 0$ corresponds to "efficiency maximizing" preferences, and someone with $\gamma = 0$ will always choose the largest transfer. The case of $\gamma = 100$ corresponds to Rawlsian redistributive preferences in that the respondent will always transfer to the poorer household. An issue with interpreting these choices as true preferences is that they are also observationally equivalent to choices corresponding to simple "rules of thumb" rather than preferences. The $\gamma = 0$ case corresponds to only paying attention to transfer amount, while the $\gamma = 100$ case corresponds to only paying attention to income. Our downstream analysis takes the distributional of γ_i conditional on $\gamma_i \notin \{0, 100\}$ as our measure of true preferences. This will *understate* the true dispersion in redistributive preferences. We discuss robustness to this pre-registered sample restriction in Appendix D.

Finally, let $\boldsymbol{\sigma} = (\bar{\sigma}/3, (\bar{\sigma}/3 + 2\bar{\sigma})/2, 2\bar{\sigma})$ be a grid of error variances. Then we assume that the joint distribution of (γ_i, σ_i) is supported on $\boldsymbol{\gamma} \times \boldsymbol{\sigma}$. This discrete distribution can be

[7] The intuition behind the penalty term is that the standard FKRB specification has difficulty distinguishing random coefficients which are too similar to one another. Heiss et al. (2022) propose the above penalty term to help smooth over these difficulties by encouraging the estimator to pick estimated distributions of γ_i that places similar mass on similarl γ's. This is akin to the use of kernels to smooth out estimated densities when performing kernel density estimation. In line with this intuition, we pick our penalty parameter to decay with sample size according to $\lambda \propto N^{-1/5}$.

represented by a matrix $\boldsymbol{\theta}$ where $\theta_{ab} \equiv \Pr[\gamma_i = a, \sigma_i = b]$, where $\theta_{ab} \geq 0$ for all a, b and $\sum_{a,b} \theta_{ab} = 1$.

Recovering the Welfare Weight Function From the estimated distribution of γ_i, we can recover the average SMWW at any income level z:

$$\hat{G}(z) = \mathbb{E}[\exp(-\gamma_i z)] \tag{4}$$

Figure 5 plots this function together with a 95% confidence interval. Welfare weights are convex in income: People place lower weights on the welfare of richer families, but weights fall at a decreasing rate. The implication is that differences in income among two relatively poor households matter more for welfare than differences in income among two rich families.

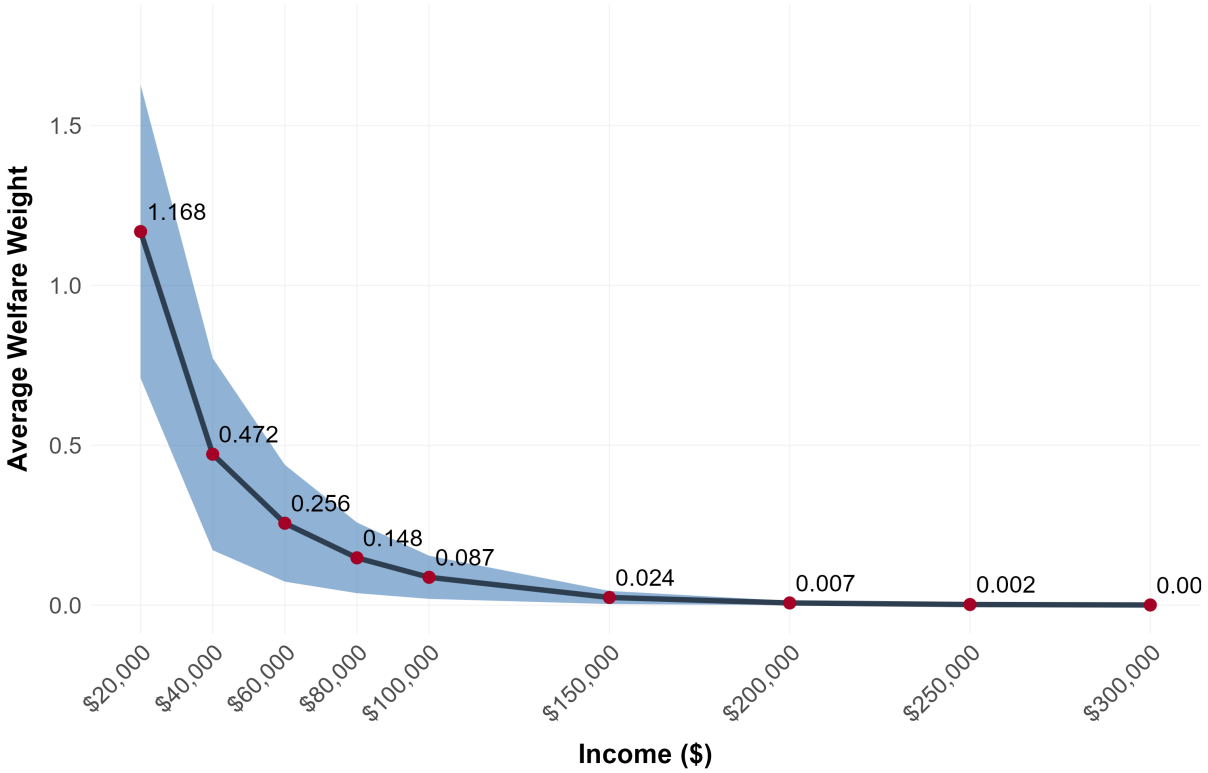
At the bottom of the income distribution with an annual pre-tax income of $20,000, the average social welfare weight is 1.17 and falls to 0.47 as income increases to $40,000. Consequently, people care roughly twice as much about the welfare of the poorer household. As we further increase income to $60,000 and $80,000, weights drop to 0.26 and 0.15, respectively. These weights imply that the welfare of a family earning $20,000 has almost 8 times more import than one that earns $80,000. Annual incomes of $100,000 already have only a weight of 0.024, which is 2% of the welfare weight for the poorest family in the plot. Above $150,000 the function becomes very flat and reaches 0 at $300,000. A welfare weight of zero means that the optimal marginal tax rate is revenue maximizing (at the maximum of the Laffer curve) and proportional to the inverse income elasticity (recall equation 1). Thus, normative preferences have a large effect on the marginal tax rate at the bottom and middle of the income distribution, but matter little at the top.

In Appendix G, we illustrate that variance in the welfare weights is barely driven by sampling variance and almost entirely driven by true disagreement in norms regarding redistribution. This suggests that gathering more data from experiments like ours would not meaningfully decrease variance in welfare weights. However, it could be that normative preferences are malleable, for instance, because people have miscalibrated beliefs about behavioral responses to taxation or because they misestimate other people's social preferences. We explore these possibilities in a second experiment next.

4.2 The Role of Beliefs about Efficiency and Equity for Welfare Weights

In a second experimental wave (N=2,456), we test whether information about behavioral responses and social norms affects redistributive preferences. We randomly provide subjects with information about either the ETI from our meta-analysis or the redistributive preferences re-

Figure 5: Social Marginal Welfare Weights



Notes: This figure plots social marginal welfare weights across the income distribution, together with 95% confidence interval.

vealed in our first experiment. We find that information about the ETI reduces redistributive preferences, likely by shifting attention to efficiency costs. Information about others' redistributive tastes increases redistributive behavior, mostly because people underestimate others' strong progressive tastes. These findings are qualitatively in line with earlier evidence by Stantcheva (2021). The effects of the informational interventions have no economically meaningful effects on our optimal tax schedules, which we estimate in the next section. We refer the curious reader to Appendix F.1 where we discuss the information experiment in detail.

5 Uncertainty about The Optimal Income Tax

In this section, we combine our meta-analytic results about elasticities of taxable income with our experimental results about heterogeneity in redistributive preferences in order to understand the relative roles of positive versus normative uncertainty in driving disagreements over tax policy. Specifically, we compute optimal income tax schedules within a standard Mirleesian framework,

and show how variation in preference and beliefs about elasticity drive the resulting optimal tax schedules.

We are interested in solving for a tax schedule $T(z)$ mapping pre-tax income z to the amount of taxes owed by the individual. Let $v(\cdot)$ be a fixed value of the individual-level utility function, and let F_θ be the corresponding distribution of θ types in the population. Given a tax schedule $T(z)$, an individual of type θ endogenously chooses labor supply to solve $z^*(T;\theta,v) \equiv \arg\max_z z - T(z) - \theta v(z)$, which yields utility $U(T;\theta,v) \equiv \max_z z - T(z) - \theta v(z)$. The social planner has redistributive preferences of the same CARA form used to parameterize participants' preferences: $\mathcal{W}(U;\gamma) = -\frac{1}{\gamma}\exp\left(-\gamma(U-\bar{z})\right)$, where $\bar{z}$ is the average level of income at current tax levels.[8] Following the standard setup, we assume that the social planner wishes to maximize average welfare, subject to raising revenue at least equal to $\underline{R}$, where $\underline{R}$ is the level of revenue raised by the current income tax schedule. Formally, the social planner's objective is

$$\max_T \int \mathcal{W}(U(T;\theta,v);\gamma)\,\mathrm{d}F(\theta) \quad \text{s.t.} \quad \int T(z^*(T;\theta,v))\,\mathrm{d}F(\theta) \geq \underline{R} \tag{5}$$

We calibrate $\underline{R}$ to be equal to the current level of revenue raised by the government. In practice, $\underline{R} \approx$\$15,000, which constitutes the current government proceeds per household in the US *after transfers*. Thus, this can be thought of as the amount currently collected for the purpose of financing public goods, and other government projects.

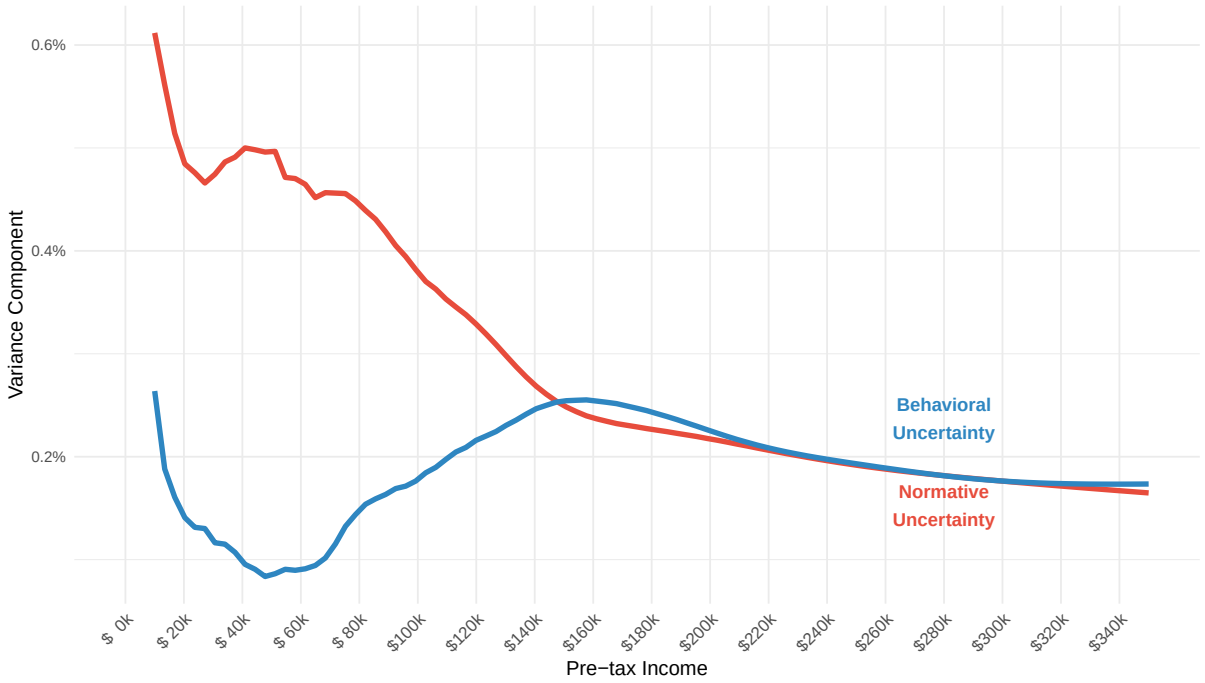
As in Section 2, let $H(z)$ denote the distribution of endogenously chosen incomes under the optimal tax schedule, let $\bar{G}(z) \equiv \int_{\theta:z(T;\theta)\geq z} \mathcal{W}'(U(T;\theta,v);\gamma)\,\mathrm{d}F(\theta)$ be the social marginal welfare weight for individuals with income above z, and let $\alpha(z) \equiv zh(z)/(1-H(z))$ be the local Pareto parameter. Then standard results (e.g., Piketty and Saez, 2013) show that the solution to the optimal tax problem in equation (5) can alternatively be characterized in terms of its well-known first-order condition which we had stated in Section 2:

$$T'(z) = \frac{1-\bar{G}(z)}{1-\bar{G}(z)+\alpha(z)\varepsilon(z)}. \tag{6}$$

While this formula is useful in gaining intuition about the nature of the optimal income taxation, it cannot be directly implemented, because the distribution of incomes $H(z)$, as well as the social marginal welfare weights $\bar{G}(z)$, are endogenous to the tax schedule. In Appendix C, we describe in detail how, for each given draw of the elasticity function $\varepsilon(z)$ and redistributive preference γ, we compute the optimal income tax schedule characterized by Equations (5) and (6).

[8] Note that subtracting $\bar{z}$ from U does not affect the optimal policy, but is useful for maintaining numerical stability of the exponential function.

Figure 6: Variance Decomposition of Uncertainty about the Optimal Income Tax Schedule



Notes: This figure decomposes the total variance of the optimal marginal tax rate into two components: i) the part of the variance that results from variation in the behavioral response to taxation (in blue), and ii) the part that results from variation in normative preferences about redistribution (in red). Behavioral uncertainty accounts for a larger part of the total variance at very low incomes and at large incomes above $150,000. Normative uncertainty is more important for the large share of the population that earns between $20,000 and $150,000.

Variance Decomposition In our first exercise, we show how the optimal tax rate varies because of i) uncertainty about behavioral responses, and ii) variance in redistributive preferences. For each pre-tax income level z, we compute the following decomposition of the variance in the optimal marginal tax rate:

$$Var(T'(z)) = \underbrace{Var(\mathbb{E}[T'(z)|\gamma])}_{\text{Normative Uncertainty}} + \underbrace{\mathbb{E}[Var(T'(z)|\gamma)]}_{\text{Positive Uncertainty}} \tag{7}$$

The first part of the equation computes the expected optimal tax rate on z for a given redistributive preference γ, and then takes the variance of this conditional expectation over the distribution of γ. This tells us by how much the expected optimal tax rate varies *because people have different redistributive preferences.* The second part of the variance decomposition computes the variance of the optimal tax rate for a *given* redistributive preference γ, and then takes the expectation of these conditional variances. This quantity represents the average amount of variation in the optimal tax that is left after redistributive preferences are fixed. This is precisely the amount of variation coming from uncertainty about how people respond to taxes at income

level z.

Figure 6 plots the two components for each part of the income distribution: the red curve is the variance from normative uncertainty, the blue curve the one from behavioral uncertainty.

For most of the income distribution, heterogeneity in redistribution preferences contributes more to uncertainty in optimal tax rates than uncertainty in behavioral elasticities. At the very bottom of the income distribution, the contribution of positive uncertainty is relatively large and then decreases steeply as income increases and only becomes economically relevant again in the right tail of the income distribution. This pattern reflects the fact from our meta-analytical results, seen in Figure 4, that uncertainty about the elasticity of taxable income is large at the very bottom of the income distribution and to the right of the 75th percentile. By contrast, the contribution of normative uncertainty to uncertainty about optimal tax policy is monotonically declining in income, mirroring Figure 5 where welfare weights are more precisely measured for high relative to low incomes.

The net implication of these two patterns is that for the two-thirds of Americans who earn below $100K, normative uncertainty accounts for up to 85% of the overall variance in uncertainty about optimal tax policy. Disagreement about how to tax these households is mostly about values, not about efficiency. For incomes above $150K, the contribution of normative and positive uncertainty are roughly equivalent. More precise evidence on the ETI of high-earners can be particularly useful here.

Taken together, normative uncertainty is the dominating force in driving uncertainty about the optimal tax rate for a large part of the income distribution. This result is striking considering that most of the empirical literature on optimal taxation tends to focus on measuring the ETI rather than people's norms (with some important exceptions mentioned in the introduction). As we show next, the results underscore the importance of integrating normative heterogeneity into tax policy analysis, alongside the traditional focus on behavioral elasticities.

5.1 The Optimal Income Tax Schedule

We discuss how positive and normative uncertainty translate into the optimal income tax schedule. To provide an intuitive comparison, we do the following exercise. We compute two sets of optimal tax schedules: one that would be optimal if redistributive preferences were all low, equal to the 10^{th} percentile of the distribution of γ; and another schedule that would be optimal if redistributive preferences were all high, equal to the 90^{th} percentile. For each of these two extreme values of γ, we compute a 90% confidence interval that represent the remaining uncertainty solely from variation in the estimated posterior distribution of ETIs. Panel A of Figure 7 plots the results for the entire income distribution. The thick blue curve represents median optimal tax policy among the set of people with strong redistributive tastes, as represented by

the 90^{th} percentile of γ (the "political left"). The thick red curve represents median optimal tax policy for the subset of people with low redistributive preferences, given by the 10^{th} percentile γ (the "political right"). Meanwhile, the thin dotted lines demarcate the 10^{th} and 90^{th} confidence intervals due to behavioral uncertainty arising from variation in the ETI.

The differences between the thick lines represent the extent to which *normative* considerations cause different people to have different beliefs about optimal tax policy. This difference is large throughout the entire income distribution. The most redistributive people would want marginal tax rates to be around 10-20 percentage points higher everywhere compared to the least redistributive people. Behavioral uncertainty matters little for incomes below $70,000 as confidence intervals are tightly estimated. These intervals start becoming much larger around $100,000. At around $150,000, the red and blue confidence intervals are nearly overlapping. Policy Uncertainty is now equally driven by uncertainty over the ETI as by normative disagreement about how to tax the relatively wealthy.

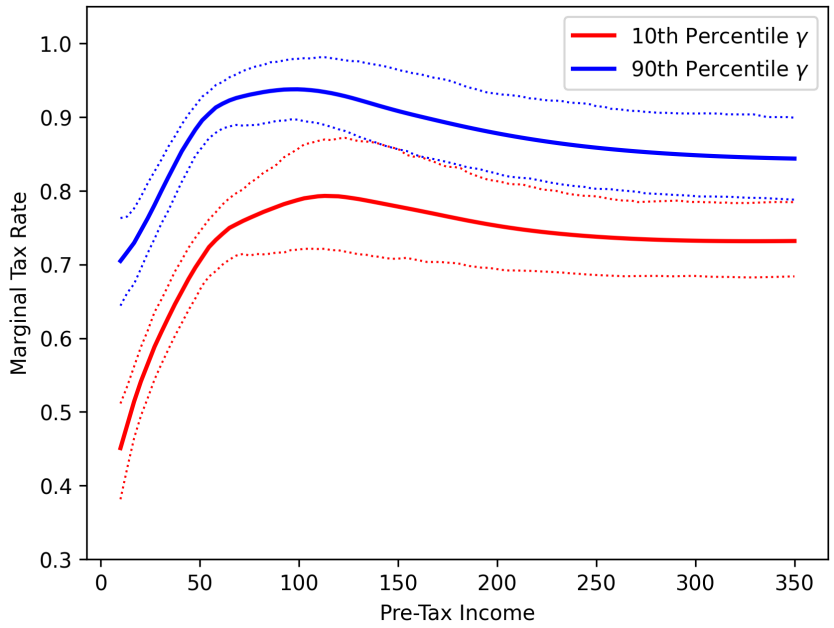
Note further that marginal tax rates peak at around $100k and then slightly decrease again. This pattern is driven by efficiency concerns: Consulting Figure 2, the ETI is much lower for the 75th percentile ($\hat{\approx}$ $100k) than for percentiles further to the right of the distribution. While society assigns greater normative weight to $100k earners than to top earners, this is outweighed by the efficiency advantage of imposing higher rates where labor supply responses are more limited.

In Panel B, we try to understand how the marginal tax rate schedules from Panel A affect transfer policies at the bottom. We plot the overall tax burden for the subset of households earning less than $60,000. Negative values represent net transfers from the government to the taxpayer. We find that across the distribution of redistributive preferences, respondents are in favor of a substantial level of support for low-income households. Even the least redistributive households support transfers of $27,000 to the poorest households. However, the most redistributive subjects prefer to redistribute about $8,000 more annually to the lowest-income household.

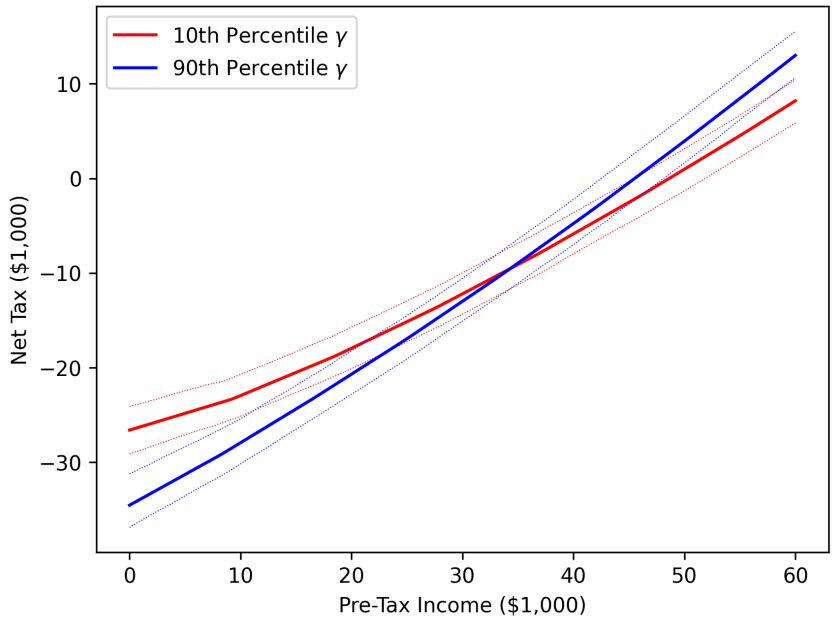
Democratic Preference Aggregation Finally, we compare the social planner's optimal tax policy against the one that would arise if tax policy was decided on through direct democracy. Specifically, for each posterior draw, we compare the policy which maximizes *average* welfare (averaging the welfare function across the distribution of γ) to the policy which maximizes the preferences of the median voter (the median γ's welfare). In Figure 8, we plot the optimal marginal tax rates. We find that the two policies are essentially identical: maximizing the median voter's preferences gets us to the first-best policy that would maximize average welfare.

Optimal marginal tax rates are substantially higher than in the current US tax system, which has maximal marginal rates of 37% in 2025. In Appendix D, we show that the current income

Figure 7: The Optimal Income Tax Schedule



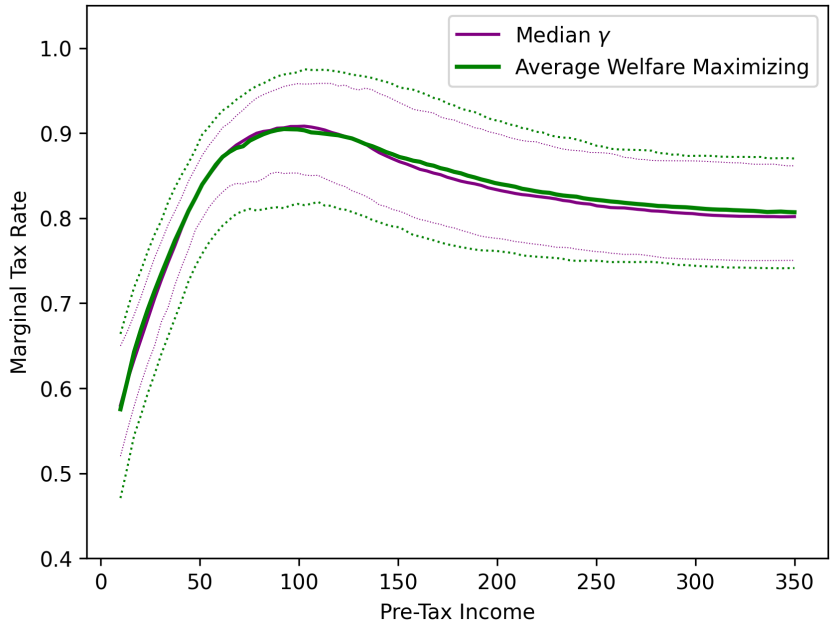
(a) Marginal Tax Rates



(b) Net-of-Transfer Tax Burden at the Bottom

Notes: Panel A shows optimal marginal tax rates as a function of pre-tax income. In red, we illustrate the optimal tax schedule using the 10th percentile of the distribution of CARA parameters. These are the parameters preferred by people with the least redistributive preferences. In blue we show the 90th percentile of CARA parameters, reflecting the most redistributive preferences. The difference between the red and blue lines gives our measure of *normative* disagreement about redistribution. The dotted lines are confidence intervals of the optimal tax rate, resulting from *positive* uncertainty about the elasticity of taxable income. At the bottom of the income distribution, normative disagreement is large while positive uncertainty is small. At the top of the distribution, confidence intervals become wide, implying that positive uncertainty becomes more important. Panel B illustrates how the marginal tax rates translate into the implied total tax burden for households earning less than $60,000 annually. Negative values are transfers.

Figure 8: Optimal Tax Policy versus Median Voter Preferences



Notes: This Figure compares the optimal tax schedule if the social planner either maximizes welfare using the *average* welfare weights of the population or the welfare weight of the *median* preference in the experiment. 90% confidence intervals are depicted by the dotted lines.

tax schedule in the US would be close to optimal if society assigned *equal* welfare weights to each income level (i.e., the utalitarian case). This is puzzling because, given our welfare weights, even Republicans should support much higher taxes than are currently in place. We next explore why our estimated tax rates are so much higher than the policies people vote for.

6 Why are Optimal Tax Rates so Much Higher than Current US Income Taxes?

Redistributive preferences in our data mirror standard ideological divisions, with conservatives favoring lower redistribution than liberals. Despite these qualitative patterns, even the 10^{th} percentile least redistributive individual in our sample should prefer a tax policy that is far more progressive quantitatively than the status quo in the United States. Clearly, this does not reflect the typical political debate in which Democrats want more redistribution than the status quo and Republicans prefer lower taxes throughout. What drives this gap between stated preferences that we observe in debates and preferences implied from transfer choices in our experiment?

To answer this question, we run a third wave of our experiment with 2,555 subjects. We randomize subjects into one of two groups. In the first group, subjects make the same transfer

choices as in wave 1 and 2 but, in addition, we ask them directly about their preferences for income tax rates in a separate question. This question is shown either before or after the transfer choices, in random order. Specifically, for incomes in $\{\$20K, \$40K, \$60K, \$100K, \$500K, \$1M\}$, we ask subjects what, in their ideal tax policy, would be the *average* tax rate of individuals at those income. We elicit average tax rates because they are cognitively easier to process than marginal tax rates. The inclusion of this question allows us to understand if our transfer-choice-based preferences for redistribution systematically diverge from the tax schedule subjects can directly select in a hypothetical scenario.

One reason why stated and revealed preferences may differ, is that people dislike giving money to the government and our transfer choices do no incorporate any notion of taxation: subjects choose to allocate windfall gains to one of two families, but there is no income tax raised by the government.[9] This motivates our second experimental group. In this group, we design transfer choices that involve giving money to the government through a new incentive-compatible mechanism. We modify our 10 transfer choices by telling subjects that there is a baseline transfer to each family, and a tax on the transfer which goes to the government. Let B_{iqj} denote the baseline transfer in the experiment, and let τ_{iqj} be the tax rate. Then the recipient household j in question q of decision-maker i gets an after-tax transfer of $T_{iqj} = (1 - r_{iqj})B_{iqj}$, while the remaining amount $\tau_{iqj} = r_{iqj}B_{iqj}$ goes to the government. We randomize $B_{iqj} \in \{400, 800, \ldots, 4000\}$ and $r \in \{0\%, 25\%, 50\%, 75\%\}$ across choices and, as in previous waves, also the income of each family.

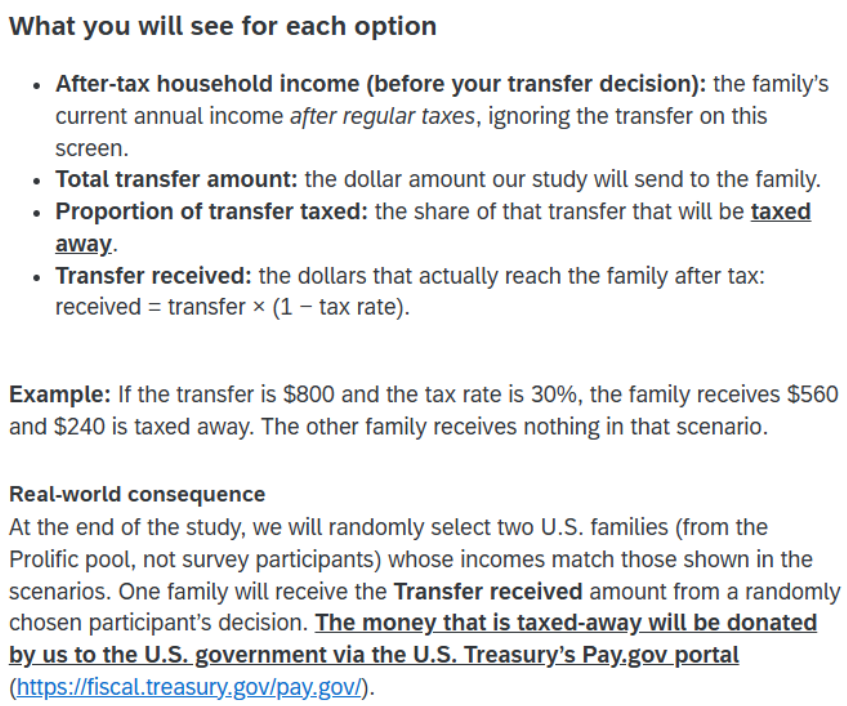
Incentive Compatibility Importantly, choices are fully incentivized: the transfer amount of a randomly selected choice goes to the selected family and the corresponding tax revenue is *donated to the US government* via the U.S. Treasury's www.Pay.gov portal. This portal allows anyone to give money to the Treasury. We tell subjects that if their choice is randomly selected, we will donate the collected tax revenue to the US government. Figure 9 provides a screenshot of the relevant instructions.

6.1 Results

We start by comparing stated versus revealed preferences of the group that made our standard transfer choices without taxes to the Treasury. In Figure 10, dotted lines are stated preferences, and solid lines are revealed preferences from transfer choices; the black line is the current US

[9] Another reason why stated and revealed preferences diverge could be because hypothetical choices are often more pro-social than true behavior, e.g. due to social desirability bias (Rodemeier, 2025; Bursztyn et al., 2025). However, this would mean that hypothetical tax rates would be even more progressive than our revealed preference estimates, which is not what we find.

Figure 9: Instructions Explaining Real Transfers to the U.S. Treasury



What you will see for each option

- **After-tax household income (before your transfer decision):** the family's current annual income *after regular taxes*, ignoring the transfer on this screen.
- **Total transfer amount:** the dollar amount our study will send to the family.
- **Proportion of transfer taxed:** the share of that transfer that will be **taxed away**.
- **Transfer received:** the dollars that actually reach the family after tax: received = transfer × (1 − tax rate).

Example: If the transfer is $800 and the tax rate is 30%, the family receives $560 and $240 is taxed away. The other family receives nothing in that scenario.

Real-world consequence

At the end of the study, we will randomly select two U.S. families (from the Prolific pool, not survey participants) whose incomes match those shown in the scenarios. One family will receive the **Transfer received** amount from a randomly chosen participant's decision. **The money that is taxed-away will be donated by us to the U.S. government via the U.S. Treasury's Pay.gov portal** (https://fiscal.treasury.gov/pay.gov/).

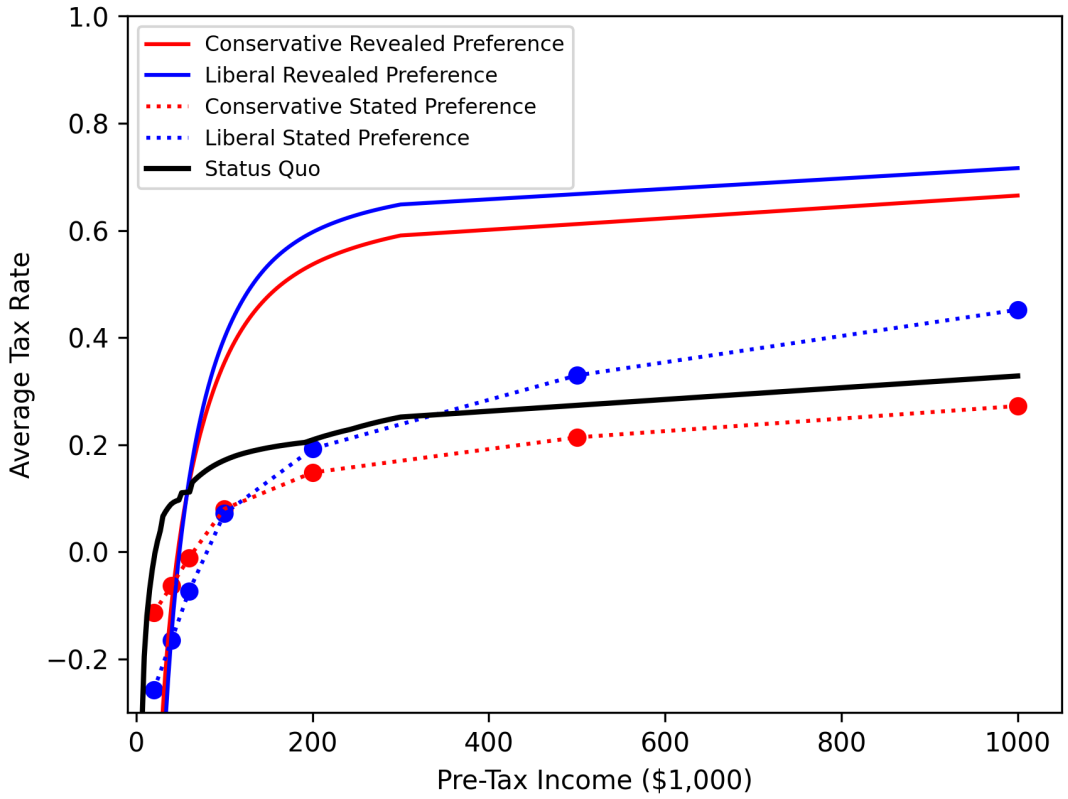
Notes: This Figure illustrates the relevant excerpt of the experimental instructions for subjects that faced transfer choices with taxes to the US Treasury.

federal tax code. Red and blue lines represent the subsamples that self-identify as Republican and Democrats; moderates are excluded from the plot for simplicity.[10] Revealed preference estimates that are implicit in the transfer choices of subjects are again close to our previous results, and, dramatically more progressive than stated preferences. For incomes of $100k annually, implied *average* tax rates are around 35%, around 15% above the current US tax code. By contrast, both political camps state that they want lower taxes (of around 10%) for this income group in their hypothetical elicitation. Generally, we find that, stated preferences across the political spectrum imply lower taxes on families with incomes below $300k than the current US tax code implements. Above that level, Republicans prefer lower tax rates for richer households than are currently in place, while Democrats prefer higher taxes. For instance, at an income of $500k annually, the US tax code implements an average tax rate of around 29%, but Democrats say they prefer a rate of 33% and conservatives a rate of 21%. Yet, revealed preferences would imply much higher rates of around 60% for both political camps.

These results illustrate that when subjects directly choose income taxes in hypothetical elicitations, the resulting preferences are much closer to the current US tax code, and far below the rates we obtain from incentivized transfer choices.

[10] Stated tax rates by moderates lie, as expected, between the blue and red dotted lines.

Figure 10: Revealed vs Stated Tax Preferences



Notes: This Figure illustrates revealed preference estimates of optimal average income taxes (colored solid lines), stated preferences (dashed lines), and the current US income tax schedule black line). Blue and red indicate people's self-reported political affiliation to either Democrats or Republicans.

To explore this puzzle, we turn to our second group whose incentivized transfer choices were tied to giving money to the government. As before, let C_{iqj} be the post-tax income of the recipient household, and let Y_{iq} be an indicator for giving the transfer to household 1. We run a regression of the form

$$Y_{iq} = \underbrace{\beta_1[\tau_{iq1} - \tau_{iq2}]}_{\text{effect of tax difference}} + \underbrace{\beta_2 \frac{C_{iq1} + C_{iq2}}{2}}_{\text{average effect of income}} + \underbrace{\beta_3[C_{iq1} - C_{iq2}]}_{\text{effect of income difference}} + \underbrace{\beta_4 \frac{T_{iq1} + T_{iq2}}{2}}_{\text{average effect of transfer}} + \underbrace{\beta_5[T_{iq1} - T_{iq2}]}_{\text{effect of transfer difference}} + \varepsilon_{iq}. \quad (8)$$

In this regression, β_1 measures the isolated effect of higher tax amounts paid to the government on subjects' choice, while holding constant income and net-of-tax transfer size across receiving families. Coefficients β_2 and β_4 represent, respectively, the effect of the average income level and the average transfer level across the two families on the subject's choice. Coefficients β_3 and β_5 tell us how the *difference* in income and transfer size between the two families matter for the choice. A comparison of β_1 with β_5 quantifies how much more subjects value giving money to the receiving family than to the government.

We run the regression separately for conservatives, moderates, and liberals. The results from this regression are given in Table 1. For conservatives, the first row shows that holding all else constant, increasing the amount received by the government by $1,000 *decreases* the probability of making that transfer choice by 4.7 p.p. This means that conservatives actively dislike when the Treasury receives money. It is interesting to compare this to the last row, which shows that increasing the amount given to a household *increases* the probability of making the transfer choice by 7.2 p.p. To put this number into perspective, this implies that a $1,500 transfer, taxed at 33% (so the household receives $1,000 and the government receives $500) is, on average, as attractive as making a tax-free $700 transfer to the household. In other words, *conservatives in our sample are willing to reduce net generosity to recipients in order to avoid giving money to the government*. The results illustrates a strong aversion against providing funds to the government generally, and are even more striking in the light of the current US federal government being Republican.

The point estimate for taxes amongst liberals is small and statistically insignificant. Thus, while liberals in our sample do not seem to actively avoid giving the government money, they derive no incremental utility from doing so. One explanation for why Democrats do not positively value government revenue is that the current administration is Republican and may use the funds for a political agenda they oppose politically. Moderates appear about halfway between conservative and liberals, and while not statistically significant, appear to directionally also dislike transferring money to the government.

These results are consistent with the idea that observed tax policy deviates from what would be implied by revealed redistributive preferences because people in general, are skeptical of redistribution mediated through government taxation.

6.2 Optimal Taxes with Government Aversion

We quantitatively assess the importance of our estimated measures of government aversion in shaping optimal tax policies. For a given tax policy $T(\cdot)$, define $\mathcal{T}_+(T) = \{z : T(z) > 0\}$ as the set of incomes where individuals face a net tax burden. We parameterize the idea of government aversion by modifying the planner's problem so that the social planner trades off the utility effects of tax reform against an explicit penalty on the overall level of tax revenue collected. Formally, we compute optimal tax policies of the form

$$\max_{T(\cdot)} \int_0^\infty \Gamma\big(u\big(z - T(z)\big) - \theta v(z)\big)\, h(z)\, dz - \lambda \int_{\mathcal{T}_+(T)} T(z) h(z)\, \mathrm{d}z,$$
$$\text{s.t.} \quad \int_0^\infty T(z)\, h(z)\, dz \geq E.$$

Table 1: Transfer Sensitivities to Taxes, Income, and Transfer Size

	Subgroup:		
	Conservative	Moderate	Liberal
	(1)	(2)	(3)
$\tau_{iq1} - \tau_{iq2}$	-0.047192^{***}	-0.011121	0.000401
	(0.011578)	(0.010731)	(0.008202)
$\frac{C_{iq1}+C_{iq2}}{2}$	0.000951^{***}	0.001209^{***}	0.001184^{***}
	(0.000085)	(0.000078)	(0.000061)
$[C_{iq1} - C_{iq2}]$	-0.000568^{***}	-0.000407^{***}	-0.000805^{***}
	(0.000062)	(0.000061)	(0.000044)
$\frac{T_{iq1}+T_{iq2}}{2}$	0.138473^{***}	0.111931^{***}	0.111746^{***}
	(0.014865)	(0.013512)	(0.010656)
$[T_{iq1} - T_{iq2}]$	0.072480^{***}	0.060652^{***}	0.048438^{***}
	(0.009414)	(0.008790)	(0.006619)
N	1,440	1,760	2,840

Notes: This Table reports regression results from specification 8, separately for self-reported conservatives, moderates, and liberals. Statistical significance at **p<0.1; **p<0.05; ***p<0.01*

Here, $\lambda \geq 0$ reflects the degree of government aversion, and the case where $\lambda = 0$ corresponds to the standard Mirrlees problem. This represents an additional deadweight loss from government involvement itself, not included in the standard Mirrleesian framework.

To identify λ, we use the coefficients from Table 1. Note that if preferences over tax policy conform to the standard Mirrlees model, a single dollar of tax raised for the government should be valued equivalently to a transfer, given to a randomly chosen member of society. Thus, the coefficient on $\tau_{iq1} - \tau_{iq2}$ should be positive and equivalent in magnitude to the coefficient on $T_{iq1} - T_{iq2}$. Since liberals do not actively value tax revenue (at least not to a Republican government), their level of government aversion must be just enough to offset the positive benefits of the transfer. Thus, $\lambda_{lib} \approx \mathbb{E}[G_{lib}(z)]$ for liberals, where $G_{lib}(z)$ is the welfare weight placed on an individual with income z, given the redistributive preferences of the average liberal. Conservatives value a \$1 increase in transfers at 0.072 but disvalue a \$1 increase in taxes at –0.047. Thus, $\lambda = 0.047/0.072 \times \mathbf{E}[G_{cons}(z)] \approx 1.65 \times \mathbf{E}[G_{cons}(z)]$, where analogously, $G_{cons}(z)$ is the welfare weight on an individual with income z, given the redistributive preferences of the

average conservative.

With these values of λ we re-compute optimal tax functions, and produce an analogue of Figure 10 in Figure 11. Incorporating the aversion parameter drastically transforms the optimal tax schedules. Without government aversion, revealed preferences from our transfer experiments imply average tax rates of approximately 60% for households earning $500,000. With government aversion incorporated, these optimal rates fall to approximately 46% for liberals and 36% for conservatives, much closer to their stated preferences of 33% and 21% respectively. The convergence is even more striking at lower income levels. For households earning $100,000, revealed preferences without government aversion implied average tax rates are 35 and 40% for liberals and conservatives, but this falls to 10% for conservatives and 13% for liberals once we account for their respective levels of government aversion.

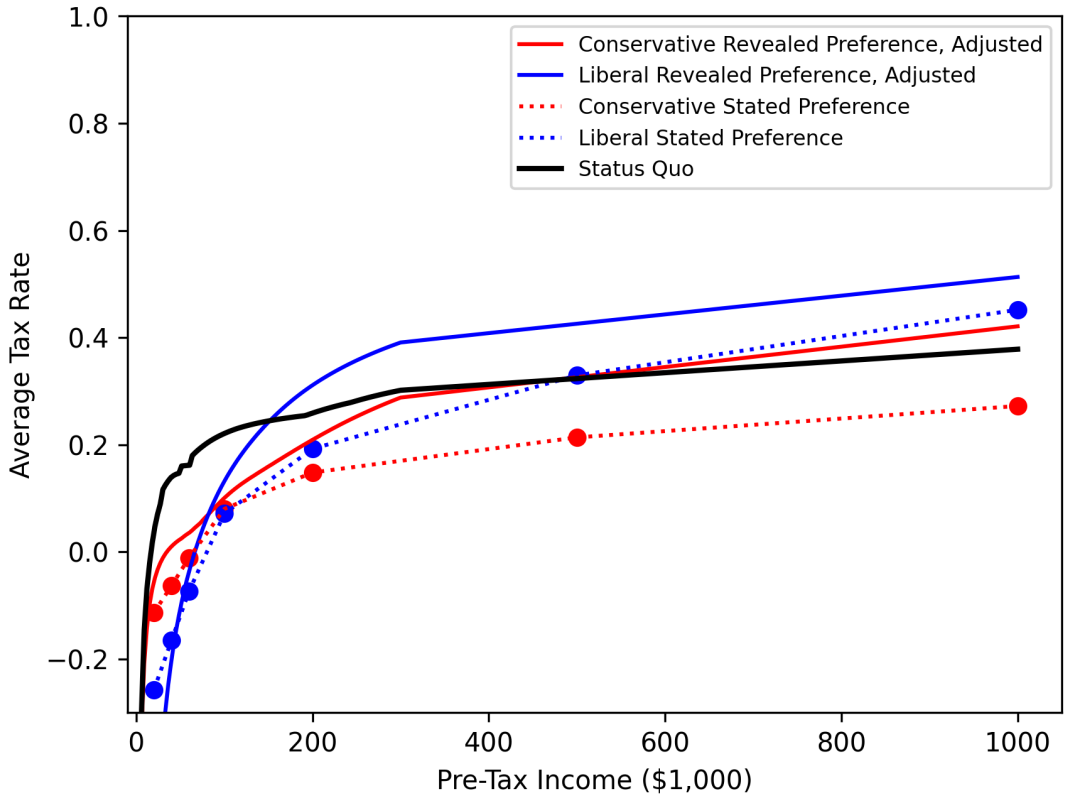
Our revised revealed preference rates now align much more closely with people's stated preferences for taxation. Government aversion can explain approximately 70-80% of the gap between transfer-based revealed preferences and people's reported ideal income tax schedule. The remaining 20-30% gap may reflect additional factors we can only speculate about, such as misunderstanding of the tax system or political economy constraints not captured in our framework.[11]

The documented partisan difference in government aversion helps explain why tax policy remains contentious even when both political camps have high redistributive preferences in private choice settings.

Interpreting Government Aversion These findings suggest that the large gap between choice-implied optimal tax rates and actual policy support may stem from a deep skepticism about government-mediated redistribution. Our experimental design leaves some questions for future work about two competing explanations for this aversion: First, respondents may distrust that the government efficiently allocates tax revenue. Under this interpretation, people care about helping lower-income households but believe that government programs waste resources, create perverse incentives, or fail to reach intended beneficiaries. The aversion to government taxation would then reflect instrumental concerns about implementation rather than fundamental opposition to redistribution. Second, respondents may have principled objections to the coercive nature of taxation itself, regardless of efficiency considerations. This view emphasizes property rights and the moral significance of earned income, suggesting that taking money from those

[11] For example, in our data, we see that liberals state preferred top income tax rates that are above the current US federal tax rate but below the average US tax rate once state and local taxes are also included. The remaining difference between stated and revealed income taxes could be largely explained if stated tax preferences from our survey correspond to preferences about *federal* policy, while revealed tax preferences are more relevant to understanding preferences about the total tax burden, including state and local taxes.

Figure 11: Revealed vs Stated Tax Preferences



Notes: This Figure illustrates revealed preference estimates of optimal average income taxes (colored solid lines), stated preferences (dashed lines), and the current US income tax schedule black line). Blue and red indicate people's self-reported political affiliation to either Democrats or Republicans.

who earned it violates important ethical principles even when used for beneficial purposes. Our results show clear partisan differences in these attitudes but we cannot determine which mechanism drives these patterns. This is important because the policy implications differ substantially. If government aversion stems from efficiency concerns, institutional reforms might unlock latent support for redistribution. If instead it reflects deeper philosophical commitments about property and coercion, such reforms would likely prove ineffective.

Implications for Optimal Tax Theory These findings add a new dimension of uncertainty to the optimal tax literature: even when we can measure redistributive preferences through revealed choice, the translation to tax policy preferences depends critically on attitudes toward government that may override pure redistributive motives. This suggests that the standard optimal tax framework, which typically assumes a benevolent social planner implementing transfers costlessly, may miss a first-order consideration in real-world policy design. The implementation mechanism itself—whether redistribution occurs through government taxation, private charity,

or some alternative institution—appears to fundamentally alter people's willingness to support transfers.

7 Conclusion

This paper provides a comprehensive empirical assessment of the relative importance of positive uncertainty versus normative disagreement in optimal income tax design. By combining a Bayesian meta-analysis of elasticity estimates with experimental evidence on redistributive preferences, we quantify how much of the disagreement about tax policy stems from scientific uncertainty about behavioral responses versus fundamental differences in values about redistribution.

Our central finding is that, given the well-established evidence on behavioral elasticities, a large part of the remaining uncertainty about optimal taxation is driven by variance in normative preferences. For the two-thirds of American households earning below $100,000, up to 85% of the uncertainty about optimal marginal tax rates stems from disagreement about redistributive values. This pattern only changes in the right tail of the income distribution, where normative and behavioral uncertainty become equally important.

Our final experiment reveals an additional layer of complexity: stated preferences for income taxes are much less progressive than what people's redistributive behavior in transfer choices would imply. We trace this discrepancy back to people's aversion against government taxation and find that implementation mechanisms of transfers through taxes via the government fundamentally reduces support for progressive taxation.

These findings have profound implications for both research and policy. First, disagreement about tax policy is less about technocratic disagreements about elasticities, and more about normative values and deep skepticism about government. This three-way tension between behavioral uncertainty, normative disagreement, and government aversion may better explain why tax reforms remain politically contentious despite extensive empirical research and apparent public support for helping lower-income households.

Second, our results challenge the standard approach of inferring optimal policy directly from revealed redistributive preferences. The same preferences that imply highly progressive taxation in abstract choice settings may not translate to support for government-implemented tax policy. This may stem from concerns about government efficiency or philosophical objections to coercive redistribution.

For future research, these findings point to several priorities. In the tails of the income distribution, where behavioral uncertainty is large, we need more well-identified evidence on how the very poor and especially the wealthy respond to taxation. For understanding tax policy prefer-

ences, we need to go beyond measuring redistributive values to understand how implementation mechanisms shape policy support. This includes distinguishing between efficiency-based and philosophical sources of government aversion, and exploring whether alternative mechanisms for redistribution might better align with both redistributive preferences and institutional skepticism.

By acknowledging all sources of disagreement about income taxation, we can work toward tax systems that are not only efficient and aligned with social values but also politically feasible.

References

Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of public Economics*, 89(5-6):897–931.

Alesina, A., Stantcheva, S., and Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, 108(2):521–554.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2011). Tax-benefit systems in europe and the us: between equity and efficiency. Technical report, EUROMOD Working Paper.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2014). Comparing inequality aversion across countries when labor supply responses differ. *International Tax and Public Finance*, 21:845–873.

Bottan, N. and Perez-Truglia, R. (2025). Betting on the house: Subjective expectations and market choices. *American Economic Journal: Applied Economics*, 17(1):459–500.

Bursztyn, L., Haaland, I. K., Röver, N., and Roth, C. (2025). The Social Desirability Atlas. Technical report, National Bureau of Economic Research.

Capozza, F. and Srinivasan, K. (2025). Who should get money? estimating welfare weights in the us. Working Paper, August 2025.

Cappelen, A., Liu, Y., Nielsen, H., and Tungodden, B. (2024). Fairness in a society of unequal opportunities.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.

DQYDJ (2020). Average, median, top 1%, and all united states household income percentiles in 2020. https://dqydj.com/

2020-average-median-top-household-income-percentiles/. Accessed June 16, 2025.

Fox, J. T., Kim, K. I., Ryan, S. P., and Bajari, P. (2011). A simple estimator for the distribution of random coefficients. *Quantitative Economics*, 2(3):381–418.

Hajargasht, G. and Griffiths, W. E. (2013). Pareto–lognormal distributions: Inequality, poverty, and estimation from grouped income data. *Economic Modelling*, 33:593–604.

Hawkins, D. M. (1991). The convolution of the normal and lognormal distributions. *South African Statistical Journal*, 25(2):99–128.

Heathcote, J., Storesletten, K., and Violante, G. L. (2017). Optimal tax progressivity: An analytical framework. *The Quarterly Journal of Economics*, 132(4):1693–1754.

Heiss, F., Hetzenecker, S., and Osterhaus, M. (2022). Nonparametric estimation of the random coefficients model: An elastic net approach. *Journal of Econometrics*, 229(2):299–321.

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of public Economics*, 187:104198.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8(1):435–464.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

Kuziemko, I., Norton, M. I., Saez, E., and Stantcheva, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508.

Lockwood, B. B. and Weinzierl, M. (2016). Positive and normative judgments implicit in us tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics*, 77:30–47.

MacKay, D. J. et al. (1998). Introduction to gaussian processes. *NATO ASI series F computer and systems sciences*, 168:133–166.

Matthews, A. G. d. G., van der Wilk, M., Nickson, T., Fujii, K., Boukouvalas, A., León-Villagrá, P., Ghahramani, Z., and Hensman, J. (2017). GPflow: A Gaussian process library using TensorFlow. *Journal of Machine Learning Research*, 18(40):1–6.

Mengel, F. and Weidenholzer, E. (2023). Preferences for redistribution. *Journal of Economic Surveys*, 37(5):1660–1677.

Mirrlees, J. A. (1971). An exploration in the theory of optimal taxation. *The Review of Economic Studies*, 38(2):175–208.

Piketty, T. and Saez, E. (2013). Optimal income labor taxation. *Handbook of Public Economics*, 5:391–474.

Rodemeier, M. (2025). Willingness to pay for carbon mitigation: Field evidence from the market for carbon offsets. *Review of Financial Studies*.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229.

Saez, E. (2010). Do taxpayers bunch at kink points? *American economic Journal: economic policy*, 2(3):180–212.

Saez, E., Slemrod, J., and Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, 50(1):3–50.

Saez, E. and Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(01):24–45.

Salimi-Khorshidi, G., Nichols, T. E., Smith, S. M., and Woolrich, M. W. (2011). Using gaussian-process regression for meta-analytic neuroimaging inference based on sparse observations. *IEEE transactions on medical imaging*, 30(7):1401–1416.

Stantcheva, S. (2021). Understanding tax policy: How do people reason? *The Quarterly Journal of Economics*, 136(4):2309–2369.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: journal of the Econometric Society*, pages 307–333.

Online Appendix

A Estimation Details of Bayesian Meta-Analysis

In this Appendix, we describe a Hierarchical Bayes framework for thinking about uncertainty about the elasticity function ε. The empirical challenge is to use the estimates from our raw data and construct a smooth function $\varepsilon(z)$ that gives an average ETI for each income level z. Rather than learning about one particular elasticity, we want to learn about an elasticity *function* for counterfactual policy analysis. This requires an approach that differs from the typical Bayesian meta-analysis in economics which typically learns about a single elasticity. Here we borrow from an extensive machine learning literature that uses Gaussian Processes (GP) to predict functions from data.

A.1 Specification of the Prior

Recall that our goal is to construct a function $\varepsilon(z)$ which represents the ETI at different points in the income distribution.

We model a prior over the function ε by specifying it as the realization of a *stochastic process* over $\mathbb{R}_+$ taking values in $\mathbb{R}_+$. Specifically, we model $\log \varepsilon$ as being drawn from an appropriately specified Gaussian process. We begin with a definition.

Definition 1. Let $\mathbf{m} : \mathbb{R}_+ \to \mathbb{R}$ be a *mean function* and let $\mathbf{C} : \mathbb{R}_+^2 \to \mathbb{R}$ be a *covariance kernel* satisfying the following properties:

1. $\mathbf{C}$ is symmetric, that is, for any $(z, z') \in \mathbb{R}_+^2$, $\mathbf{C}(z, z') = \mathbf{C}(z', z)$
2. $\mathbf{C}$ is positive-definite, that is, for any $z_1, \ldots, z_k$,

$$\sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{C}(z_i, z_j) \geq 0$$

Then Y is a *Gaussian process* with mean $\mathbf{m}$ and covariance $\mathbf{C}$, denoted $Y \sim \mathrm{GP}(\mathbf{m}, \mathbf{C})$, is a stochastic process on $\mathbb{R}_+ \to \mathbb{R}$ characterized by the property that for any finite set of points, $z_1, \ldots, z_k$

$$\begin{bmatrix} Y(z_1) \\ \vdots \\ Y(z_k) \end{bmatrix} \sim \mathcal{N}(\boldsymbol{\mu}(z_1, \ldots, z_k), \boldsymbol{\Sigma}(z_1, \ldots, z_k))$$

where $\boldsymbol{\mu}(z_1, \ldots, z_k)_i = \mathbf{m}(z_i)$ and $\Sigma(z_1, \ldots, z_k))_{ij} = \mathbf{C}(z_i, z_j)$.[12]

Rather than directly specifying elasticity as a function of income, we specify elasticity as a function of the percentile rank of income. This has the benefit of compressing the distribution of kink/notch locations studied in the literature into a uniformly spread-out domain. We assume that $\varepsilon : [0, 1] \to \mathbb{R}_+$ is described by the stochastic process

$$\log \varepsilon \sim \mathrm{GP}(\mathbf{m}, \mathbf{C}) \tag{9}$$

for mean function $\mathbf{m}$ and covariance kernel $\mathbf{C}$ to be specified. If $H(z)$ is the distribution of incomes, then the elasticity of taxable income at z is given by $\varepsilon(H(z))$. Our priors are specified for $\log \varepsilon$ in order to constrain the elasticity function to be positive. Our mean function $\mathbf{m}$ is chosen to be the constant function $\mathbf{m}(r) = \log \varepsilon_0$ where ε_0 is taken to be the 0.33 mean elasticity found in Chetty (2012).

Our covariance function, on the other hand, is chosen to take the form

$$\mathbf{C}(r, r') = \sigma_0^2 \exp\left(-\frac{(r - r')^2}{2l^2}\right)$$

where here, σ_0, l are hyper-parameters controlling respectively the overall prior uncertainty and smoothness of $\varepsilon(r)$. We take our prior parameter σ_0 to be sufficiently diffuse so that for each r, the plausible values of $\varepsilon(r)$ reflect the full range of elasticities considered to be reasonable in the literature. Since labor supply elasticities in the literature tend to range from a low end of nearly 0 to a high end of just under 2, we take $\sigma_0 = 0.9$. This choice of σ_0 ensures that $\exp(\mathbf{m} + 2\sigma_0) \approx 2$. The literature provides less guidance for what a reasonable value of l could be. We therefore parameterize it to come from a relatively uninformative prior. Specifically, we take $l \sim \Gamma(0.2, 1)$.

Given this prior distribution, we assume that the data from our meta-analytic study arrives as follows:

1. Nature samples a elasticity function $\log \varepsilon$ from the Gaussian Process prior (9). This is the *true* elasticity of taxable income function for the world.

2. Fix study s, measuring the elasticity of taxable income at income level z_s. The study has a standard error, $\sigma_{2s} \sim g(\cdot|H(z_s))$ and an *asymptotic estimate*

$$\mu_s \sim \log N(\log \varepsilon(H(z_s)) - \sigma_1^2/2, \sigma_1^2).$$

[12] Intuitively, a Gaussian process is a stochastic process whose finite-dimensional marginal distribution on any set of points $\{z_1, \ldots z_k\}$ are jointly normal with mean given by $\mathbf{m}$ evaluated at $z_1, \ldots, z_k$ and with covariance given by $\mathbf{C}$ evaluated at pairs of points within the set $\{z_1, \ldots, z_k\}$.

3. Given the asymptotic estimate, the study yields a *point estimate*

$$\hat{\varepsilon}_s \sim \mathcal{N}(\mu_s, \sigma_{2s}^2).$$

This DGP requires us to specify one more hyperparameter, σ_1^2. Because the sample is directly informative about σ_1^2, we can place a prior distribution on σ_1^2 in a relatively data-driven manner. Specifically, we propose to obtain a prior distribution σ_1^2 using the procedure in Algorithm 1.

Algorithm 1 Empirical Bayes Estimates for σ_1^2 Prior Distribution

1: Estimate a weighted regression of $\hat{\varepsilon}$ as a function of z.
2: Compute an estimate for the average conditional variance of the asymptotic estimate $\tilde{\sigma}_1^2$ by subtracting the mean squared standard standard error from the residual variance of the first-stage regression
3: Compute $\hat{\sigma}_1^2$ from $\tilde{\sigma}_1^2$ by as $\hat{\sigma}_1^2 = \log\left(1 + \frac{\tilde{\sigma}_1^2}{\frac{1}{S}\sum_{s=1}^{S}\hat{\varepsilon}_s^2}\right)$
4: Use bootstrap to compute standard errors from the above three-step procedure, $\widehat{se}$.
5: Parameterize the hyperparameter σ_1 to follow a Γ prior via moment matching $\beta = \frac{\hat{\sigma}_1^2}{\widehat{se}^2}$, $\alpha = \beta\hat{\sigma}_1^2$.

Before proceeding, a few remarks are in order. Within our framework, we model the existence of a single "true" elasticity of taxation function $\varepsilon(r)$ characterizing the entire world. On top of this true elasticity function, there are two sources of uncertainty in interpreting data. First, is the more standard source of uncertainty that the point estimates reported in any given study are measured with some noise, σ_2, relative to the value they would take on given an infinite study population. Second, however, we also introduce a second layer of uncertainty as captured by σ_1. This second layer of uncertainty can be thought of as modeling idiosyncratic *misspecification bias* within any given study. This misspecification bias might arise for a number of reasons. First, most studies justify their empirical specifications by taking a Taylor approximation of the density of taxable income over some non-infinitesimal interval of incomes. To the extent that the approximating interval is wide, this may lead to bias in the resulting estimates. Second, most studies in the literature use a constant-elasticity parametric approximation to v in constructing their elasticity estimates. While the results presented here imply that these parametric approximations may indeed be thought of as providing "local" approximations to the true elasticity, there is no guarantee that the approximation is exact. Third, and perhaps most importantly, most studies within the literature are identified by jointly specifying a utility model and some form of "behavioral" optimization friction justifying deviations from the rather stark predictions of a baseline "perfectly rational" model. Thus, misspecification errors may well arise in any given study due to misspecification of the behavioral component of the model. Intuitively, the em-

pirical Bayes strategy for picking a prior distribution for σ_1 described in Algorithm 1 identifies the variability of this misspecification bias by looking at how much of the variability in point estimates at similar income levels cannot be explained by standard errors alone.

This completes our specification of the prior and the data-generating process. To summarize, the set of stochastic parameters in our framework are given by $\Theta = (\varepsilon, \sigma_0, l, \sigma_1)$. There are a number of parameters to keep track of, so we collect each of the hyperparameters we need to choose and summarize our procedure for choosing their values in Table 2. Fixing the choices of hyperparameters gives us a well-defined prior distribution $\pi(\Theta)$ over our stochastic parameters. In the next section, we describe the likelihood function $\pi(\mathcal{D}|\Theta)$ and our methodology for sampling from the posterior given the data, $\pi(\Theta|\mathcal{D})$.

Table 2: Prior Hyperparameters

Parameter Name	**Parameter Description**	**How We Obtain Parameter Value**	**Details**
ε_0	Prior mean elasticity	Match typical mean value in literature	From Chetty (2012)
σ_0^2	Prior pointwise variance of elasticity	Try to match range of estimates in literature	Set so that upper end of 95% credible interval is roughly 2
l	Prior covariance of elasticity function	Uninformative Prior	$\Gamma(0.2, 1)$
σ_1	Parameter from Gamma prior for misspecification variance, σ_1^2	Empirical Bayes	$\Gamma(\alpha, \beta)$ with (α, β) from Algorithm 1

A.2 Sampling From The Posterior

We begin this section by explicitly describing the likelihood of the data fixing a value of the prior parameters, Θ. Given our assumption that point estimates are sampled i.i.d. from some population, it suffices to analyze the likelihood contribution of a single study in isolation. Furthermore, given our assumption that study standard errors are independent of asymptotic estimates conditional on income level z_s, the posterior distribution of $\varepsilon(r)$ is *orthogonal* information about σ_{2s}. As a result, to compute the posterior distribution for $\varepsilon(r)$, it suffices to compute the conditional likelihood of a single observation, which we now do.[13]

Fix a study s which estimates and elasticity $\hat{\varepsilon}_s$ with standard error σ_{2s} at income level z_s. Given a realization of $\varepsilon(H(z_s))$ and σ_1, the DGP for $(\hat{\varepsilon}_s, \sigma_{2s})$ is as follows. First, the asymptotic value of study s is drawn according to $\mu_S \sim \log \mathcal{N}(\varepsilon(H(z_s)) - \sigma_1^2/2, \sigma_1^2)$. Second, an independent error term ϵ_s is drawn $\mathcal{N}(0, \sigma_{2s}^2)$. The point-estimate of elasticity formed by the study

[13] More precisely, it suffices to compute the likelihood conditional on parameters and conditional on the standard error.

is then given by $\hat{\varepsilon}_s = \mu_s + \epsilon_s$. The distribution of a variable like $\hat{\varepsilon}_s$, composed of the sum of an independent Lognormal and Normal random variable, has been studied before by Hawkins (1991). Its log-likelihood can be expressed via the usual convolution formula as

$$\ell(\hat{\varepsilon}_s; \mu, \sigma_1, \sigma_{2s}) = \log \left\{ \int_{-\infty}^{\infty} \frac{1}{2\sigma_1 \sigma_{2s} \pi} \exp \left(-\frac{1}{2} \left[\left(\frac{t - \mu_S}{\sigma_1} \right)^2 + \left(\frac{\hat{\varepsilon}_s - e^t}{\sigma_{2s}} \right)^2 \right] \right) \mathrm{d}t \right\} \quad (10)$$

where here, μ denotes $H(z_s) - \sigma_1^2/2$. The integral in Equation (10) has no analytical form, so we must evaluate it numerically. We describe these numerical details in Appendix A.3.

Given a specification of both the prior $\pi(\Theta)$ and the likelihood function $\pi(\mathcal{D}|\Theta)$, computing the posterior distribution $\pi(\Theta|\mathcal{D})$ is simply a matter of numerical implementation. Fortunately, Gaussian processes are fairly well-studied within the machine learning literature, so standard software packages exist that allow us to both specify our model and to numerically compute the posterior distribution. We use the GPFlow package described in Matthews et al. (2017), which provides a Markov Chain Monte Carlo (MCMC) functionality allowing us to sample from the posterior distribution of our parameters given the data. We now turn to our data before implementing the model described in this section.

A.3 Numerical Implementation of Likelihood Function

In this section, discuss the numerical implementation of the log-likelihood function (10). We proceed by first using a change-of-variables to express the variable of integration t in standardized units and then by approximating the integral in Equation (10) using standard quadrature techniques. Specifically, we re-express the log-likelihood as

$$\begin{aligned} \ell(\hat{\varepsilon}_s; \varepsilon(H(z_s)), \sigma_1, \sigma_{2s}) &= \log \left\{ \int_{-\infty}^{\infty} \frac{1}{2\sigma_1 \sigma_{2s} \pi} \exp \left(-\frac{1}{2} \left[\left(\frac{t - \mu_S}{\sigma_1} \right)^2 + \left(\frac{\hat{\varepsilon}_s - e^t}{\sigma_{2s}} \right)^2 \right] s \right) \mathrm{d}t \right\} \\ &= \log \left\{ \int_{-\infty}^{\infty} \frac{1}{2\sigma_{2s} \pi} \exp \left(-\frac{1}{2} \left[t^2 + \left(\frac{\hat{\varepsilon}_s - e^{\mu + \sigma_1 t}}{\sigma_{2s}} \right)^2 \right] \right) \mathrm{d}t \right\} \end{aligned} \quad (11)$$

We evaluate the last expression in Equation (11) by evaluating the integral from -4 to 4 using the trapezoid rule and 10,000 equally spaced sub-divisions.

B Robustness to Alternative Specifications

One potential concern with the specification adopted in the main body of the paper is that elasticities in countries that are too dissimilar for the US are not informative about elasticities in

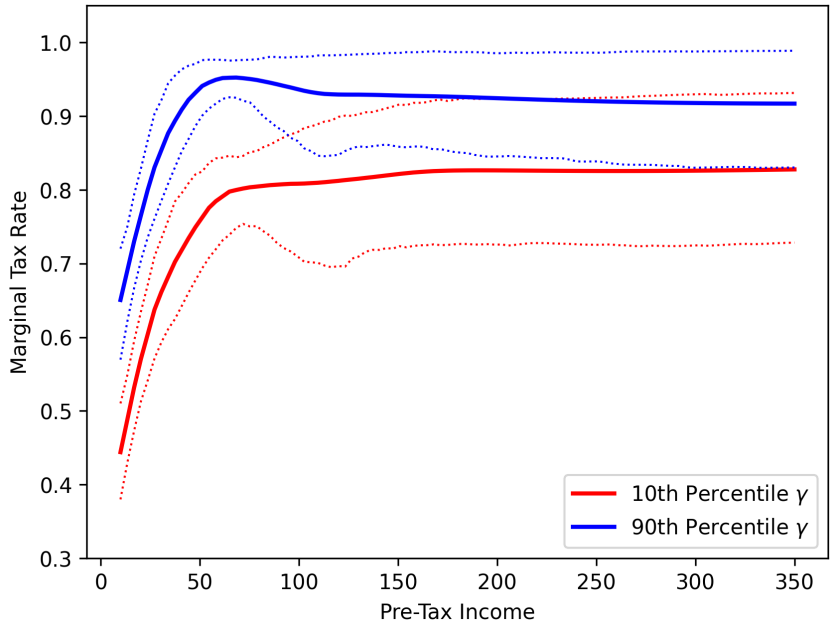
the US. In this section, we probe for robustness against this concern by reproducing our main meta-analytic results, using two alterantive specifications. First, we consider the case where we only use point estimates from high-income countries. In practice, in this first specification, we only use estimates from Denmark, Europe, Great Britain, Netherlands, New Zealand, and the United States.

Figure 12 reproduces the optimal income tax schedule using only data from high-income countries. We find that this income tax schedule looks very similar to Figure 7 in the main text that was based on estimates from all countries. The main differences are that confidence intervals in Figure 12 are wider because we have less studies and that marginal tax rates decrease *slightly* less in the far right tail of the income distribution. Panel B looks virtually identical across the two plots.

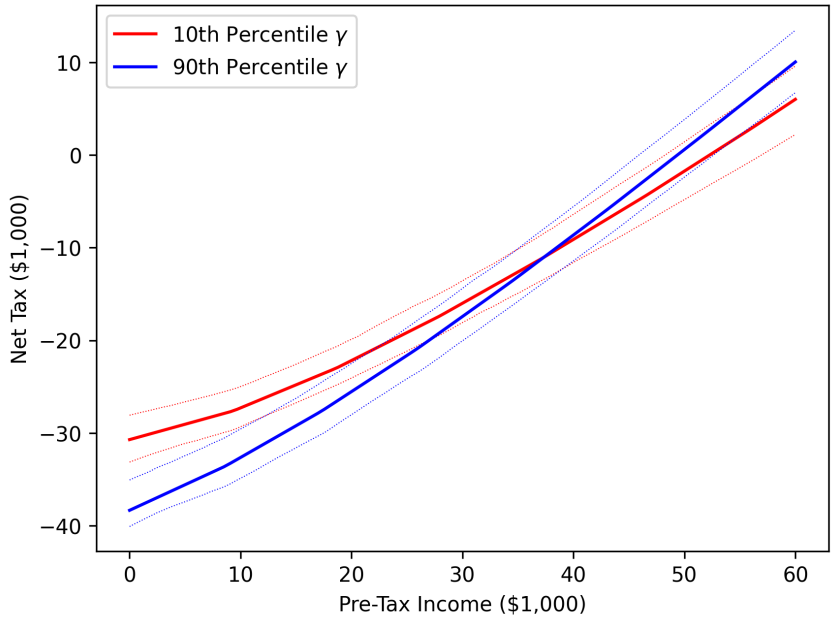
Our second robustness check addresses the concern that countries differ in their income distributions. In our main specification, we determine income percentiles through a two-step process: we first convert the income level at which each elasticity is estimated from local currency to 2020 US dollars, then calculate where this dollar amount falls within the 2020 US income distribution. This approach effectively standardizes all observations against a common benchmark—–the US income distribution. As an alternative, we now calculate percentile ranks directly within each country's own income distribution, using the original local currency values. For example, a non-US worker's income that falls into the 90th percentile of her national income distribution will now be assigned to the same group as a US worker whose income falls into the 90th percentile of the US income distribution. This approach is valid if elasticities are mostly driven by a person's *relative* distribution in her national income distribution rather than by her absolute purchasing power.

The results of doing this are reported in Figure 13. We find similar overall levels of optimal marginal tax rates, although the implied marginal tax rate for middle-income households is even higher than in our main specification. This arises largely because compared to the main specification, the alternative specification considered here implies a more pronounced drop in elasticities in the middle of the income distribution. The overall shape of the underlying elasticity function is similar across all specifications.

Figure 12: The Optimal Income Tax



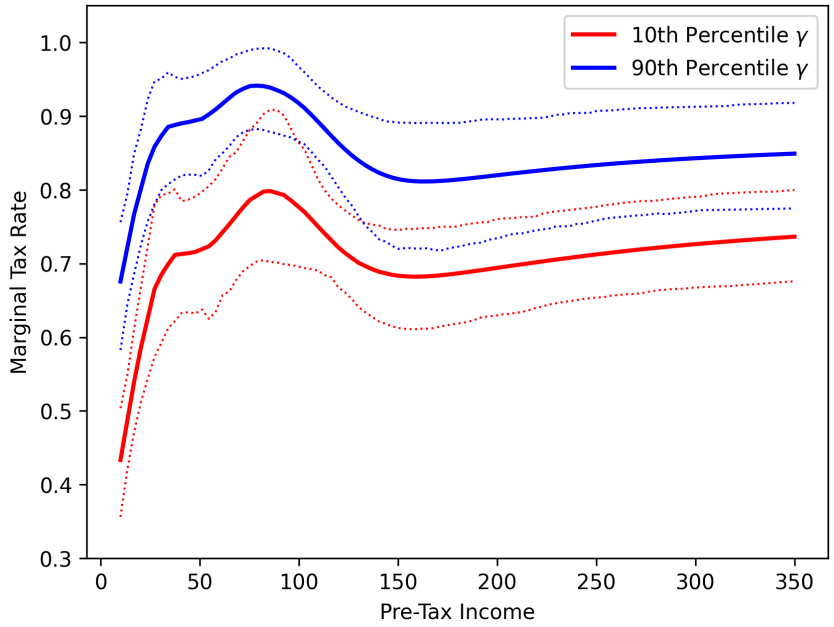
(a) Marginal Tax Rates



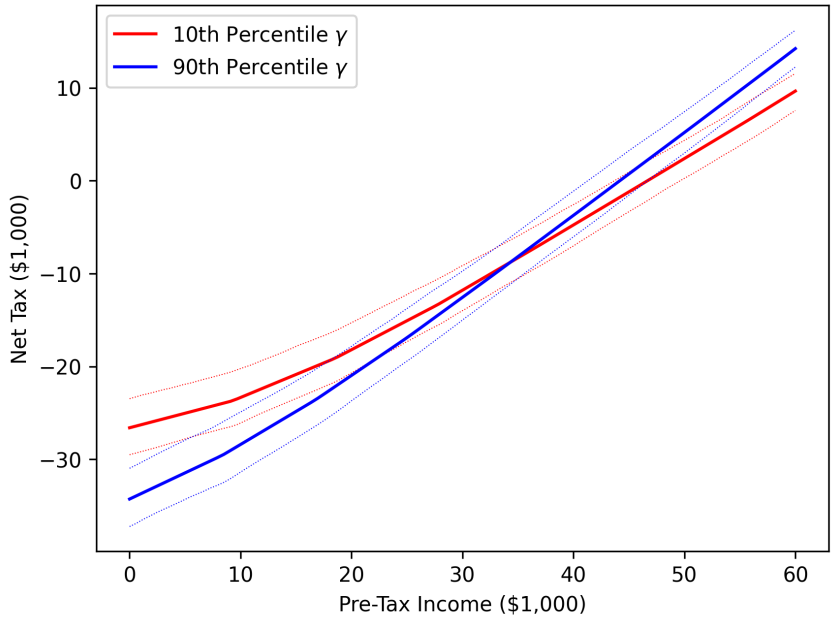
(b) Total Tax Burden

Notes: Panel A shows optimal marginal tax rates as a function of pre-tax income. In red, we illustrate the optimal tax schedule using the 10th percentile of the distribution of CARA parameters. These are the parameters preferred by people with the least redistributive preferences. In blue we show the 90th percentile of CARA parameters, reflecting the most redistributive preferences. The difference between the red and blue lines gives our measure of *normative* disagreement about redistribution. The dotted lines are confidence intervals of the optimal tax rate, resulting from *positive* uncertainty about the elasticity of taxable income. At the bottom of the income distribution, normative disagreement is large while positive uncertainty is small. At the top of the distribution, confidence intervals become wide, implying that positive uncertainty becomes more important. Finally, Panel B illustrates how the marginal tax rates from Panel A translate into the implied total tax burden in 1,000 USD at each income level. Negative values are transfers.

Figure 13: The Optimal Income Tax



(a) Marginal Tax Rates



(b) Total Tax Burden

Notes: Panel A shows optimal marginal tax rates as a function of pre-tax income. In red, we illustrate the optimal tax schedule using the 10th percentile of the distribution of CARA parameters. These are the parameters preferred by people with the least redistributive preferences. In blue we show the 90th percentile of CARA parameters, reflecting the most redistributive preferences. The difference between the red and blue lines gives our measure of *normative* disagreement about redistribution. The dotted lines are confidence intervals of the optimal tax rate, resulting from *positive* uncertainty about the elasticity of taxable income. At the bottom of the income distribution, normative disagreement is large while positive uncertainty is small. At the top of the distribution, confidence intervals become wide, implying that positive uncertainty becomes more important. Finally, Panel B illustrates how the marginal tax rates from Panel A translate into the implied total tax burden in 1,000 USD at each income level. Negative values are transfers.

C Details for Counterfactual Calculations

In this appendix, we provide additional details about how we compute the optimal income tax, given a draw of $\varepsilon(z)$, and given a value of redistributive preferences, γ. The overall procedure proceeds in two steps. In step 1, we translate a draw of $\varepsilon(z)$ into a value of the disutility function $v(z)$, and a distribution of types F_θ. In step 2, we use these two structural primitives to compute the optimal income tax schedule. We describe each of the two steps below.

C.1 Step 1: Extract Structural Primitives from Elasticity

Given a draw of $\varepsilon(z)$, recall that disutility and elasticity are related via the ordinary differential equation $v'(z)/(zv''(z)) = \varepsilon(z)$. Subject to the normalization that $v(0) = 0$ and $v'(0) = 1$, we have that we can construct v from ε via

$$v(z) = \int_0^z \exp\left(\int_0^{z'} \frac{dz''}{z''\varepsilon(z'')}\right) \mathrm{d}z'.$$

We thus obtain $v(z)$ by numerically computing the above integral.

To obtain a distribution of types F_θ, we use two sources of external information. Specifically, we need to know the current endogenous distribution of incomes $H(z)$ as well as ther current income tax schedule $T^{current}(z)$. Given these two pieces of information, we make use of two facts about the mapping between primitive types θ and endogenous incomes z:

1. The z choice corresponding to type θ satisfies first-order condition $\theta = (1-(T^{current})'(z))/v'(z)$
2. Given a fixed tax schedule $T^{current}$, z is an increasing function of θ.

These two facts imply the following relation between F_θ and $H(z)$:

$$H(z) = \Pr[\theta \geq (1 - (T^{current})'(z))/v'(z)] = 1 - F_\theta((1 - (T^{current})'(z))/v'(z)).$$

This functional relation implies that, once H and $T^{current}$ is observed, F_θ can also be recovered by solving the above equation.

In practice, we follow Hajargasht and Griffiths (2013) and approximate $H(z)$ as following a Pareto Log-Normal distribution. We estimate the parameters of this distribution to match the quantiles of the US income distribution from DQYDJ (2020). We approximate $T^{current}$ using the approximation to the US tax code from Heathcote et al. (2017).

C.2 Step 2: Structural Primitives to Optimal Tax

Before describing our approach to computing the optimal income schedule, first recall that the optimization problem that we wish to solve takes the form

$$\max_T \int \mathcal{W}(U(T;\theta,v);\gamma)\,\mathrm{d}F(\theta) \quad \text{s.t.} \quad \int T(z^*(T;\theta,v))\,\mathrm{d}F(\theta) \geq \underline{R}.$$

As a first step towards solving this problem, recall that by the revelation principle, it suffices to instead represent the optimal tax schedule via an incentive-compatible revelation mechanism, where the government asks each citizen to announce their type θ, and assigns to each type a required income $Z(\theta)$ and consumption $C(\theta)$.

Define the indirect utility function from the mechanism as $U(\theta) \equiv C(\theta) - \theta v(Z(\theta))$. Then the usual envelope theorem arguments show that incentive compatibility can equivalently be characterized by the requirements that

1. $Z(\theta)$ is decreasing in θ
2. $U'(\theta) = -v(Z(\theta))$

The second condition above pins down $U(\theta)$ up to a constant of integration. This constant of integration is in turn pinned down by the fact that any type who does not work receives a utility of 0. Thus, if $\bar{\theta}$ is the type just indifferent to work, we have $U(\bar{\theta}) = 0$

Thus, any tax schedule can equivalently be characterized by the monotone endogenous mapping $Z(\theta)$ from type to income that it induces. Moreover, a given monotone $Z(\theta)$ implies a unique implied optimal tax function. To see why, note that integrating U', we have

$$U(\theta) = -\int_{\bar{\theta}}^{\theta} v(Z(\theta))\,\mathrm{d}\theta, \tag{12}$$

and since $C(\theta) = U(\theta) + \theta v(Z(\theta))$, this pins down the mapping from type to consumption as well. The tax burden of a type θ individual is given by

$$T(\theta) = Z(\theta) - C(\theta), \tag{13}$$

so by a change-of-variables, the implied income tax as a function of income is pinned down implicitly by

$$T(z) = z - C(Z^{-1}(z)),$$

where Z^{-1} is the inverse of $Z(\theta)$ and is well-defined since Z is monotone.

To solve for the optimal nonlinear tax schedule, we therefore directly parameterize the endogenous mapping $Z(\theta)$ as a rich cubic B-spline:

$$Z(\theta) = \sum_{k=1}^{K} \omega_k \mathcal{B}_k(\theta),$$

where $\mathcal{B}_k$ are B-spline basis functions and ω_k are coefficients to be be optimized over. The monotonicity constraint on Z implies that $\omega_{k+1} < \omega_k$, $\forall k$. For a given guess of Z, Equation (12) gives us an implied value of $U(\theta)$, and using Equation (13), this also gives an implied tax burden function $T(\theta) = Z(\theta) - U(\theta) - \theta v(Z(\theta))$. The social planner's objective, given our B-spline parameterization can thus be written

$$\begin{aligned} \max_{\boldsymbol{\omega}} \int -\frac{1}{\gamma} \exp(-\gamma U(\theta)) \, \mathrm{d}F_\theta(\theta) \quad \text{s.t.} \quad & \int T(\theta) \, \mathrm{d}F(\theta) \geq \underline{R} \\ & Z(\theta) = \sum_{k=1}^{K} \omega_k \mathcal{B}_k(\theta) \\ & U(\theta) = \int_{\bar{\theta}}^{\theta} v(Z(\theta)) \, \mathrm{d}\theta \\ & T(\theta) = Z(\theta) - U(\theta) - \theta v(\theta) \\ & \omega_{k+1} \leq \omega_k \end{aligned} \tag{14}$$

This defines a nonlinear optimization problem with linear constraints, with choice variables given by the B-spline coefficients $\boldsymbol{\omega}$. We implement this optimization problem using the SLSQP implementation in optimization package of the SciPy library in Python.

D Comparing Our Results to Extreme Redistributive Preferences

The CARA specification is a simple parameterization which places voters on a spectrum in terms of how they navigate the equity-efficiency tradeoff in reasoning about taxes. To gain a broader context for thinking about the equity-efficiency tradeoff, in this section, we solve for optimal tax policy under two extreme cases. First, the case where $\gamma \to \infty$ corresponds to Rawlsian social preferences. Under this limit, taxpayers only care about maximizing the transfer given to the poorest member of society, i.e., the transfer given to someone who does not generate any

income on their own. In other words, rather than solving Equation (14), they instead solve

$$\begin{aligned} \min_{\boldsymbol{\omega}} T(\underline{\theta}) \quad \text{s.t.} \quad & \int T(\theta)\,\mathrm{d}F(\theta) \geq \underline{R} \\ & Z(\theta) = \sum_{k=1}^{K} \omega_k \mathcal{B}_k(\theta) \\ & U(\theta) = \int_{\bar{\theta}}^{\theta} v(Z(\theta))\,\mathrm{d}\theta \\ & T(\theta) = Z(\theta) - U(\theta) - \theta v(\theta) \\ & \omega_{k+1} \leq \omega_k \end{aligned} \tag{15}$$

A second extreme case occurs when $\gamma \to 0$. This corresponds to a social planner who simply wants to maximize total (unweighted) surplus. Such a social planner instead solves

$$\begin{aligned} \max_{\boldsymbol{\omega}} \int U(\theta)\,\mathrm{d}F_\theta(\theta) \quad \text{s.t.} \quad & \int T(\theta)\,\mathrm{d}F(\theta) \geq \underline{R} \\ & Z(\theta) = \sum_{k=1}^{K} \omega_k \mathcal{B}_k(\theta) \\ & U(\theta) = \int_{\bar{\theta}}^{\theta} v(Z(\theta))\,\mathrm{d}\theta \\ & T(\theta) = Z(\theta) - U(\theta) - \theta v(\theta) \\ & T(\theta) \leq Z(\theta) \qquad\qquad \omega_{k+1} \leq \omega_k. \end{aligned} \tag{16}$$

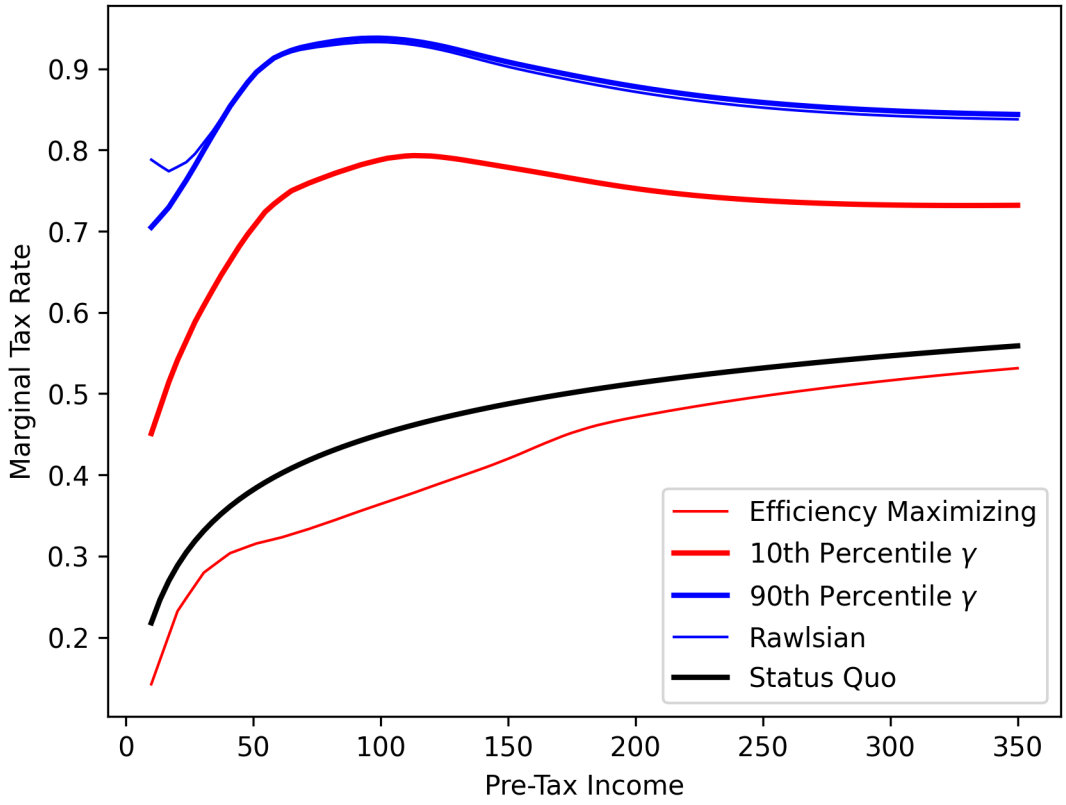
A subtlety in the above equation is that we have introduced a new constraint, that the total tax burden of any individual cannot exceed their income. This constraint is far from binding in all other optimization problems we have solved, but must be imposed for the efficiency maximizing social planner. This is because consumption utility itself is quasi-linear, so a fully unconstrained social planner would attempt to raise the entire tax burden via lump-sum taxes. Since such a tax scheme would force many agents in the economy into bankruptcy, it does not make much economic sense, so we explicitly require that the efficiency-maximizing social planner cannot raise more than is earned from each citizen.

We solve the optimization problems in Equations (15) and (16) for each draw of the elasticity function. In Figure 14 below, we plot the marginal tax rates from the 10^{th} and 90^{th} percentile redistribution preferences, in addition to the marginal tax rate from the Rawlsian and efficient social planners. We also plot the status quo marginal tax rate schedule.[14] We find that the

[14] In contrast to the body, here, we plot the Heathcote et al. (2017) approximation to the average US income tax schedule, which is inclusive of both federal and local taxes.

most redistributive citizens are very close to the Rawlsian social planner, while even the least redistributive citizens are closer to Rawlsian than to the efficient social planner. We also find that the efficiency-maximizing tax schedule is only slightly smaller than the status-quo tax schedule.

Figure 14: Rawlsian and Efficient Social Planners



Notes: This figure plots the median marginal tax rates under four specifications of the social planner: Rawsian, 90^{th} percentile γ, 10^{th} percentile γ, and efficiency maximizing.

E Comparing CRRA and CARA Redistributive Preferences

In our main specification, we parameterize redistributive preferences as taking on a constant absolute risk aversion (CARA) form. An alternative parameterization, which is sometimes found in the literature, is to assume that redistributive preferences are constant relative risk aversion (CRRA) instead. In this appendix, we discuss our choice in more detail.

Using our experimental data, we fit the CRRA analogue of our mixed logit model. Recall that in the CARA model, the decision-maker chooses redistribution to household 1 if and only if

$$T_1 \exp(-\gamma_i C_1)\varepsilon_{iq1} > T_2 \exp(-\gamma_i C_2)\varepsilon_{iq2}.$$

Under CRRA redistributive preferences, the choice to redistribute to household 1 is instead given by the condition

$$T_1 C_1^{-\gamma_i} \varepsilon_{iq1} > T_2 C_2^{-\gamma_i} \varepsilon_{iq2}.$$

We thus fit the CRRA model to our data.

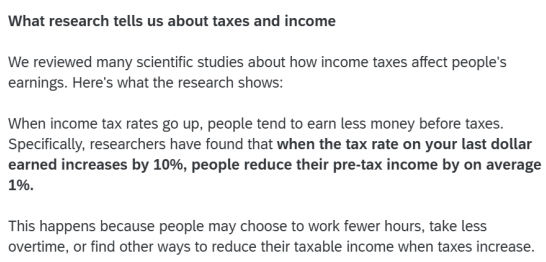
In practice, we find that while the log likelihood for the CARA model is -11,742, the log-likelihood of the CRRA model is -11,837. Thus, the CARA model fits the data slightly better than the CRRA model, although the difference is not statistically significant: a Vuong (1989) test of non-nested hypothesis yields a p-value of 0.70.

Nonetheless, we opted for a CARA specification on conceptual and computational grounds. The median CRRA exponent from our estimates is around 22, which implies that the marginal social welfare weight of someone at income level Z is 10^{22} times larger than the marginal social welfare weight of someone at income level $10Z$. This large ratio stretches the limts of computer precision, thus making our optimal taxation exercises difficult to implement in practice.

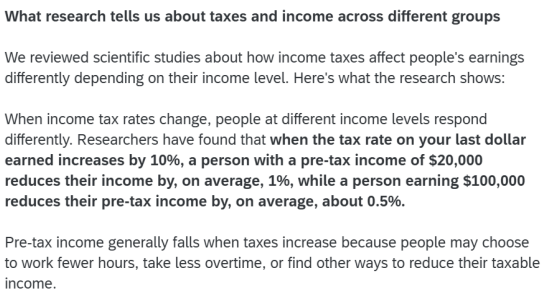
F Details on Information Treatments in Wave 2

Below we provide screenshots of each of the four information treatments used in the second experiment.

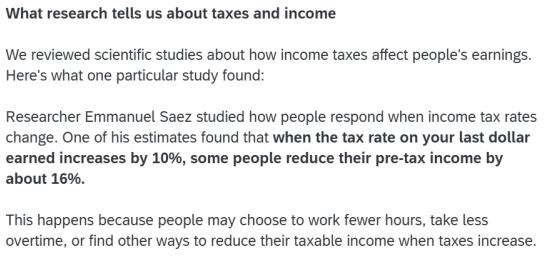
Figure 15: Information Treatments about the ETI



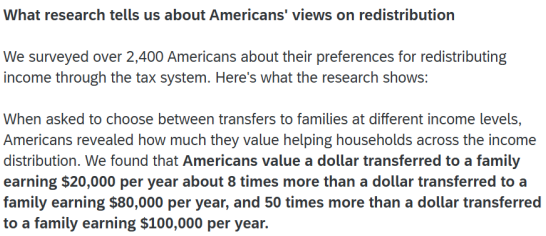
(a) Average ETI



(b) ETI for two Income Levels



(c) High ETI



(d) Social Norms

Notes: This Figure plots the three information treatments about the ETI, as well as the treatment about society's social norms regarding redistribution.

F.1 The Role of Beliefs about Efficiency and Equity for Welfare Weights

The demand for redistribution should depend not just on people's tastes but also on their beliefs about how redistribution changes behavior, which determines economic efficiency. In addition, a subject's demand may depend on how much she beliefs other people in society prefer redistribution. Based on the results of our first experiment, we ran a follow-up experiment with 2,456 subjects to study these questions. We ran the same study on the same platform but added a number of new features to the experiment.

First, we vary information about the efficiency costs of taxation by introducing respondents to scientific evidence on the elasticity of taxable income (ETI). Before subjects were making the transfer choices they were randomized either into a control group or one of three information treatments. Treatment T1 provides the average ETI estimate from our meta-analysis and is displayed in the box below.

What research tells us about taxes and income

We reviewed many scientific studies about how income taxes affect people's earnings. Here's what the research shows:

When income tax rates go up, people tend to earn less money before taxes. Specifically, researchers have found that **when the tax rate on your last dollar earned increases by 10%, people reduce their pre-tax income by on average 1%**.

This happens because people may choose to work fewer hours, take less overtime, or find other ways to reduce their taxable income when taxes increase.

Treatment T2 is designed similarly but uses an unusually high ETI. The relevant difference to the previous text is:

"[...] Researcher Emmanuel Saez studied how people respond when incomes tax rates change. One of his estimates found that **when the tax rate on your last dollar earned increases by 10%, some people reduce their pre-tax income by about 16%**. [...]"

The objective of this treatment is to induce a high belief about the ETI without deceiving participants.[15] Treatment T3 communicates heterogeneity in ETI across income levels, highlighting that richer households have a lower ETI than poorer households.

Independent of the assigned elasticity treatment, we also randomized subjects into a social norms treatment which would be seen subsequently: In this group, we informed subjects about the social welfare weights that participants in our first study revealed. To make these weights accessible to a non-technical audience, the information treatment simply illustrates what our weights implied about how people made trade-offs:

What research tells us about Americans' views on redistribution

We surveyed over 2,400 Americans about their preferences for redistributing income through the tax system. Here's what the research shows:

When asked to choose between transfers to families at different income levels, Americans revealed how much they value helping households across the income distribution. We found that **Americans value a dollar transferred to a family earning $20,000 per year about 8 times more than a dollar transferred to a family earning $80,000 per year, and 50 times more than a dollar transferred to a family earning $100,000 per year.**

[15] The approach of selectively reporting different values to induce different belief distributions is widely used in experiments and considered a deception-free way of calibrating beliefs (see, for instance, Bottan and Perez-Truglia, 2025).

F.1.1 Results

Figure 16 illustrates how income and transfer sensitivities change with the information treatments (using the same specification as previously for Figure 4). Panel A shows that information about behavioral responses to taxation reduces people's taste for redistribution. All ETI treatments reduce the (negative) income sensitivity coefficient, which falls by around 0.026–0.03 percentage points percentage points from a baseline effect of 0.15 percentage points per $1,000 in recipient income. These effects are each statistically significant from control at $p < 0.05$.

Panel B shows the corresponding effects on transfer amount sensitivity, though these are noisier and generally not statistically significant.

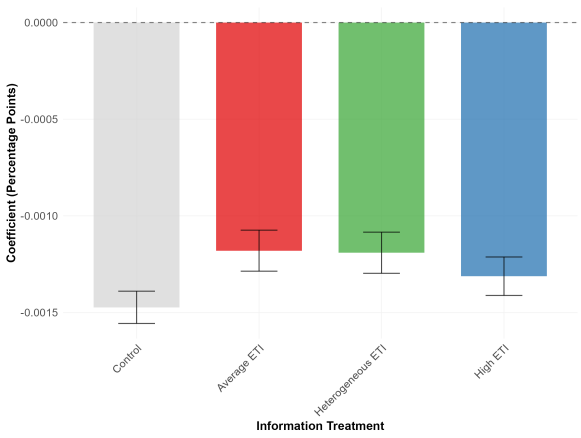
Panel A and B of Figure 17 presents the distribution of beliefs about the elasticity of taxable income (ETI) and social welfare weights across treatment groups. Panel A reveals that control group respondents substantially overestimate the ETI, with a modal belief of the pre-tax income response around 10% (i.e., an ETI of 1) and substantial right tail extending beyond 15%. The Average ETI treatment, T1, which informed respondents that income falls by 1% when tax rates increase by 10%, shifts the belief distribution leftward toward the treatment value (marked by the red dashed line). Similarly, T3 which provided income-specific elasticities of 1% for low earners and 0.5% for high earners, concentrates beliefs around these two values. In contrast, the High ETI treatment, T2, informing respondents of a 16% response implied by the 1.6 elasticity from Saez's study, exacerbates the overestimation of the ETI, shifting the distribution further rightward.

Despite shifting beliefs into different directions, the High ETI treatment has virtually the same effect on redistributive preferences as the other two. This rejects an interpretation in which behavior is driven purely by belief updating. Results rather point towards attention effects where the ETI information primes people to think about efficiency considerations, which increases reservations to redistribute.

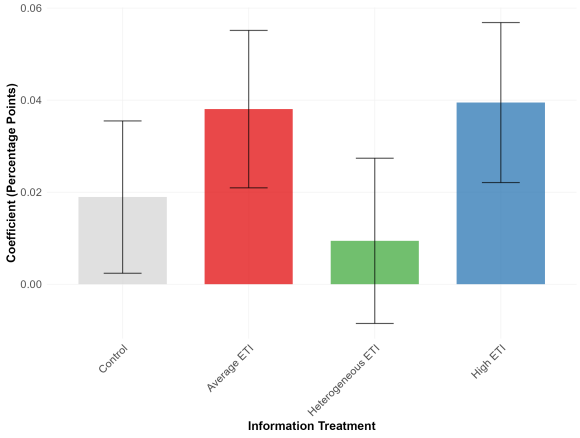
Panel C and D visualize the impact of informing people about the redistributive norms. The treatment increases the magnitude of income sensitivity from -0.00114 to -0.00145 percentage points (a 27% increase in redistributive preferences) while having a negative (but insignificant) effect on the transfer amount sensitivity. This asymmetric effect suggests that learning about societal values makes people more likely to redistribute to lower income families, as well. This is supported by the movement in belief distributions in Figure 18. In control, respondents' beliefs about how much more society values transfers to poor versus rich families are dispersed, with a modal value around 3-4. The norms information treatment shifts beliefs upwards toward the treatment value.[16]

[16] In unreported results, we find that this compression of beliefs only slightly translates into less uncertainty over the optimal tax schedule.

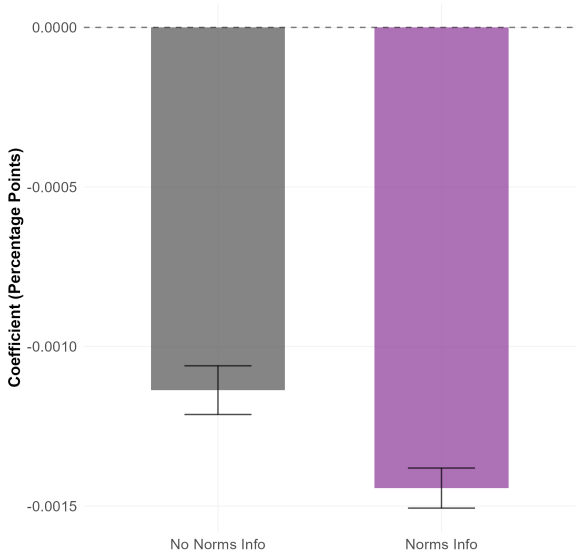
Figure 16: Main Treatment Effects on Redistributive Preferences



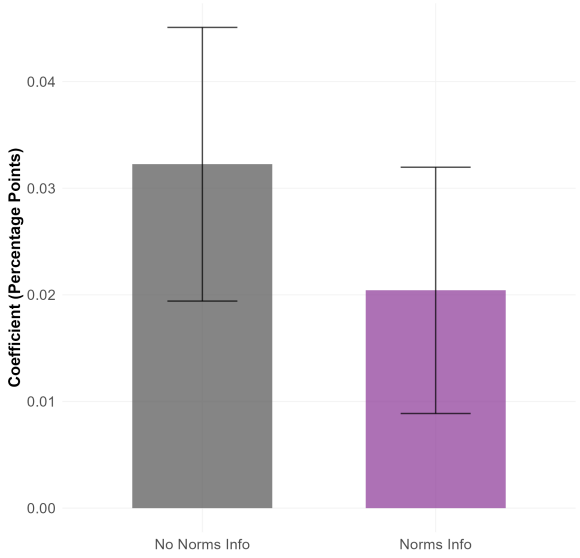
(a) Income Sensitivity To ETI Info



(b) Transfer Amount Sensitivity to ETI Info



(c) Income Sensitivity to Norms Info

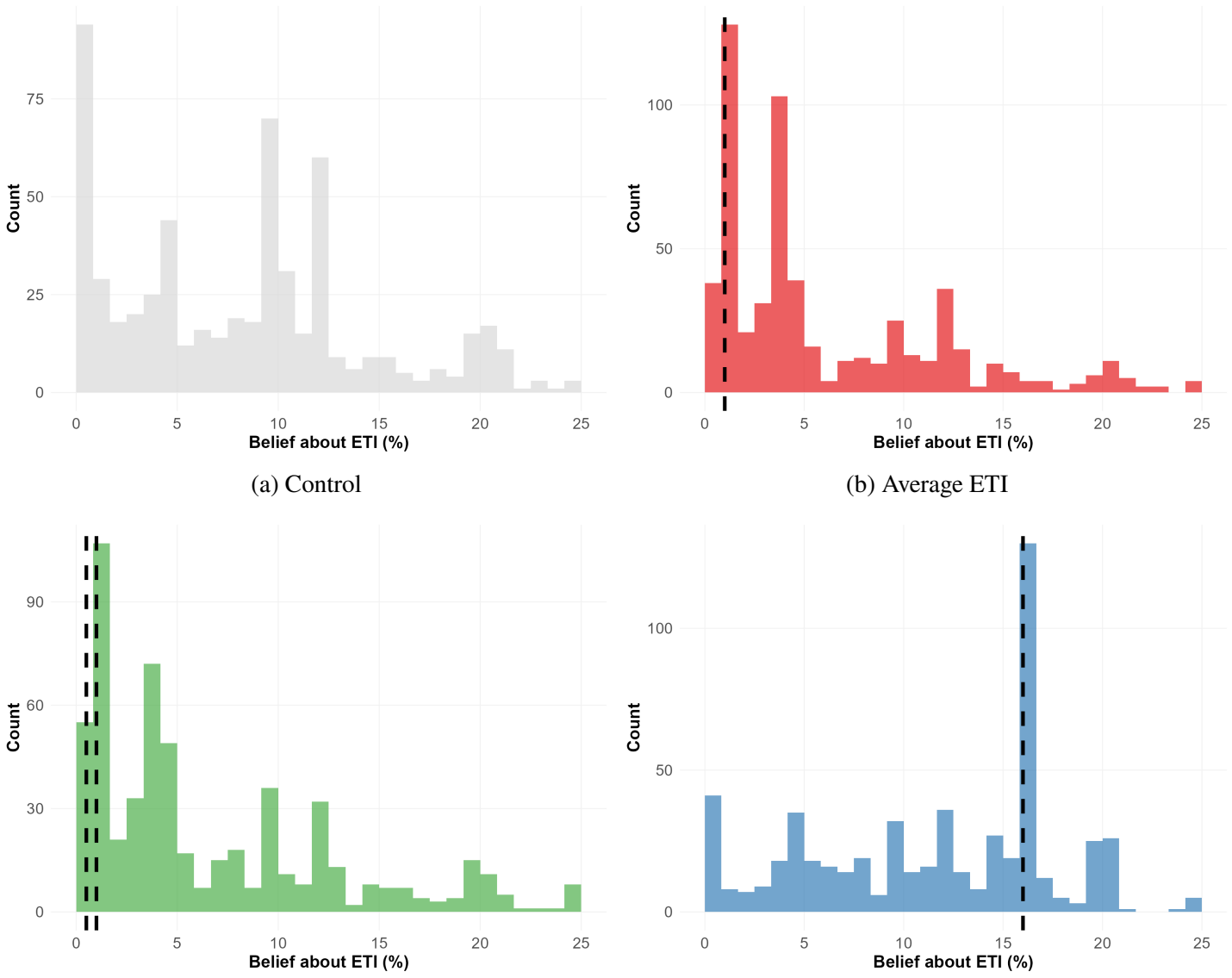


(d) Transfer Amount Sensitivity to Norms Info

Notes: This figure illustrates effects of the information treatments on redistributive preferences. Panel A and B show how information treatments affect redistribution preferences. Panel A displays the effect of a $1,000 increase in recipient income on transfer choice probability. Panel B shows the effect of a $1,000 increase in transfer amount. Panel C and D show how information about societal redistribution preferences affects individual choices. Respondents in the norms treatment learned that Americans value transfers to low-income families 8 times more than to high-income families. All coefficients are from regressions with respondent-clustered standard errors. Error bars represent 95% confidence intervals. N = 24,560 observations from 2,456 respondents.

We conclude that information about the ETI decreases support for redistribution to the poor, while social norms information increases it. Only the latter effect can be rationalized by movements in the belief distribution.

Figure 17: Beliefs about Elasticity of Taxable Income by Treatment



(a) Control

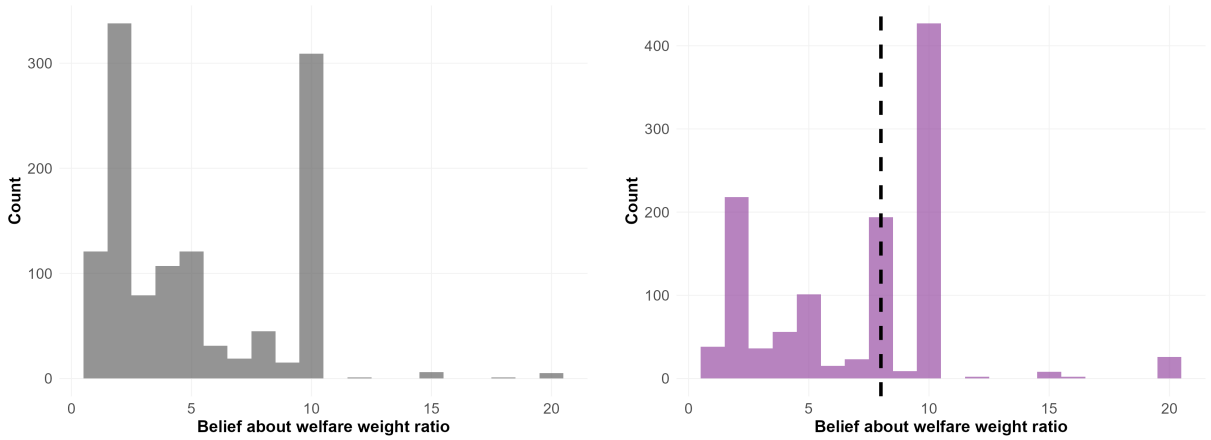
(b) Average ETI

(c) Heterogeneous ETI

(d) High ETI

Notes: Distribution of beliefs about the elasticity of taxable income (percent reduction in income from a 10% tax increase) across the four information treatments. Dashed lines indicate the values provided in each treatment: 1% for Average ETI, 0.5% and 1% for Heterogeneous ETI, and 16% for High ETI.

Figure 18: Beliefs about Social Welfare Weights by Treatment



(a) Control (b) Information about Norms

Notes: Distribution of beliefs about how much more society values $1 to poor versus rich families. The red dashed line in Panel B marks the treatment value of 8. N = 2,456 respondents.

G Uncertainty Estimation for γ Distribution

In this appendix, we show that sampling uncertainty is negligible relative to the true underlying heterogeneity in our estimates of the γ distribution. To do so, we perform a simple bootstrap procedure to estimate the uncertainty in estimating the γ distribution. We redraw our experimental data $B = 500$ times, and for each bootstrap draw, $b = 1, \ldots, B$, we re-compute a new distribution of γ, $\hat{f}_b(\gamma)$. The total variance in γ, which incorporates both sampling uncertainty and true heterogeneity is given by

$$\text{Var}(\gamma) \equiv \frac{1}{B}\sum_{b=1}^{B}\int \gamma^2 \hat{f}_b(\gamma)\,\mathrm{d}\gamma - \left(\frac{1}{B}\sum_{b=1}^{B}\int \gamma \hat{f}_b(\gamma)\,\mathrm{d}\gamma\right)^2.$$

On the other hand, the variation in γ that is attributable only to sampling variation is given by

$$\text{Var}(\mathbb{E}[\gamma|b]) \equiv \frac{1}{B}\sum_{b=1}^{B}\left(\int \gamma f_b(\gamma)\,\mathrm{d}\gamma\right)^2 - \left(\frac{1}{B}\sum_{b=1}^{B}\int \gamma f_b(\gamma)\,\mathrm{d}\gamma\right)^2.$$

In Table 3, we report $\text{Var}(\mathbb{E}[\gamma|b])$ and $\text{Var}(\gamma)$ separately. We find that $\text{Var}(\mathbb{E}[\gamma|b])$ is only 1.7% as large as $\text{Var}(\gamma)$, so heterogeneity in redistributive preferences are driven primarily by true underlying differences in opinion about γ, *not* sampling uncertainty.

Table 3: Uncertainty in γ

$\text{Var}(\mathbb{E}[\gamma\|b])$	$1.35e-9$
$\text{Var}(\gamma)$	7.65×10^{-8}

H Accounting for Optimization Frictions

A key assumption implicit throughout the paper is that our meta-analysis accurately reflect the "structural" elasticity, defined by the preferences of agents in the economy. A large behavioral economics literature has emerged arguing for the importance of optimization "frictions" in distorting behavior relative to a rational benchmark optimum. In light of this literature, it is important to understand the extent to which our results are robust to optimization frictions.

One very common type of optimization friction pointed out in the bunching literature, where many of our most precise point estimates come from, is that taxpayers may face *fixed costs of adjustment*. All common specifications of preferences from the literature, imply that if taxpayers are fully rational, the support of the income distribution should be disconnected, with "holes" occurring near kink and notch points. In practice, however, while mass near kink/notch points is greatly reduced, it is never observed to be 0 in practice. These concerns have been explicitly address in the literature on notches. Kleven (2016) points out that downwards notches in the tax code (all of the examples in our paper are of this form) generate a "dominated region" of income choices, which are suboptimal, regardless of what the true structural elasticity is. Estimators derived off of notches in the tax code can thus use the presence of mass in the dominated as an additional moment. In this case, biases due to fixed adjustment costs can be more directly controlled for by incorporating information about the share of individuals in the dominated region. We expect notch studies to therefore be especially robust to fixed cost optimization frictions. If fixed costs substantially bias kink studies, we should then expect a systematic difference between kink and notch studies. In Figure 19, we color-code kink and notch studies in our sample, run a nonparametric regression of elasticity on income in both samples. We find broad agreement on average elasticities regardless of which sample of studies we focus on. Overall, it seems that kinks tend to slightly overestimate the ETI but differences are not economically large.

The comparison between kinks and notches offers a fairly straightforward robustness test for our elasticity function and is mostly agnostic about the source of the optimization friction. This is useful because, as Chetty (2012) points out, the set of theoretically plausible sources of optimization frictions are difficult to enumerate, let alone control for.

Chetty (2012) provides an explicit partial identification methodology, which computes worst-case bounds on structural elasticities in the presence of optimization frictions. In his setting, a researcher compares behavior at two different tax policies, say τ_0, and τ_1. A perfectly frictionless

taxpayer would choose income z_0^* and z_1^* at these two tax policies, so the structural elasticity is approximately $\varepsilon \approx \frac{\log z_1^* - \log z_0^*}{\log(1-\tau_1)-\log(1-\tau_0))}$. However, in practice, optimization frictions cause actual choices to deviate from the frictionless benchmark, so the researcher instead observes $\log z_0 = \log z_0^* + \eta_0$, and $\log z_1 = \log z_1^* + \eta_1$. In that case, the observed elasticity is

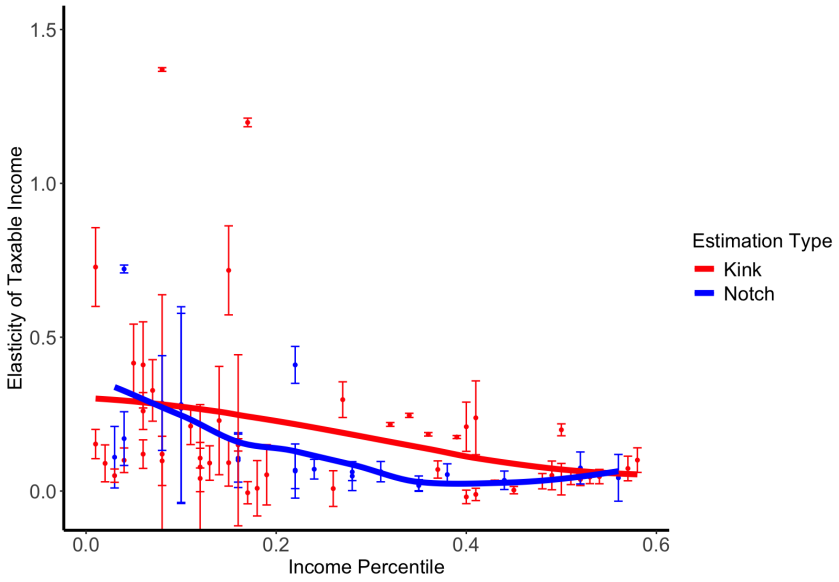
$$\hat{\varepsilon} = \frac{\log z_1 - \log z_0}{\log(1-\tau_1) - \log(1-\tau_0))} = \varepsilon + \frac{\Delta\eta}{\log(1-\tau_1) - \log(1-\tau_0))}, \tag{17}$$

where $\Delta\eta = \eta_1 - \eta_0$. The resulting bounds are very conservative because they allow for extremely adversarial (from the perspective of the econometrician) behavioral patterns arising from frictions. For example, the "upper bound" on the ETI is obtained from a data-generating process where 1) structural elasticities are very high, but 2) frictions cause taxpayers to generate too much income when taxes are high, while generating too little income when taxes are low. Applying these bounds to our data implies that for many of our point estimates, we can neither rule out elasticities near 0 nor sometimes unreasonably large elasticities above 6. This makes us hesitant to draw sharp quantitative conclusions from the bounds themselves. Instead, we view the kink–notch comparison as a credible robustness check: while it does not deliver formal partial identification, it avoids relying on extreme behavioral confusion of taxpayers and provides a more informative window into how optimization frictions may affect our elasticity estimates.

Our final approach to assessing the importance of optimization frictions is to examine whether elasticities estimated from large tax changes are less sensitive to frictions than those estimated from small tax changes. The logic follows directly from Equation (17): the same friction term $\Delta\eta$ has a larger effect on the estimated elasticity when the denominator, $\Delta\log(1-\tau) \equiv \log(1-\tau_1) - \log(1-\tau_0)$, is small. In other words, if optimization frictions matter, we would expect estimates based on small tax changes to be noisier and more variable than those based on large changes. Figure 20 puts this prediction to the data. We find no systematic difference in the variance of elasticity estimates across small versus large tax changes, suggesting that this form of optimization frictions does not play a first-order role in driving the heterogeneity of our elasticity function.

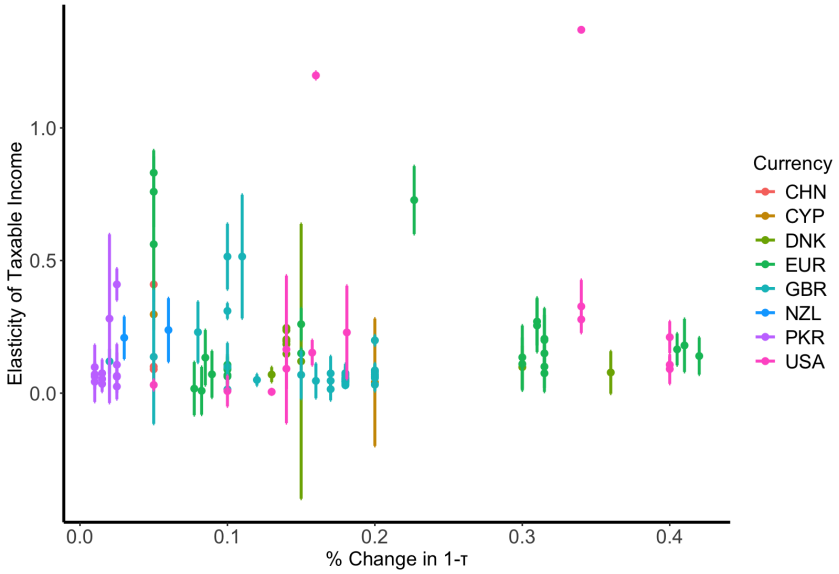
We conclude that our various tests suggest that optimization frictions, while undoubtedly present, do not fundamentally undermine our elasticity function. The heterogeneity we document appears to reflect genuine variation in behavioral responses across the income distribution rather than artifacts of measurement error or adjustment costs.

Figure 19: Kinks vs Notches



Notes: This Figure compares point estimates from studies using kinks in the tax code to studies studying notches. Each point represents the point estimate and income level of the discontinuity of a single estimate. The curves represent the results of nonparametric regressions using these points.

Figure 20: Estimates by $\%\Delta(1-\tau)$



Notes: This Figure plots estimates of elasticity of taxable income by the percent change in one minus the marginal tax rate.

I Experiment Screenshots

Figure 21: Experimental Instructions and Incentive Compatibility

In this study, you will be presented with multiple scenarios where you must choose **which financial transfer should be given to one of two families.**

Each scenario will vary in two key ways:

- **The size of the financial transfer** each family could receive.
- **The after-tax income of each family**, which represents their earnings after taxes have been deducted (excluding the transfer).

Since only **one of the two transfers** can be implemented, your decision determines which family will receive financial support.

How Your Choice Could Have Real Consequences

At the end of this study, we will randomly select two families from the U.S. population who are also part of the Prolific sample but **are not participants in this survey.** These families will be **matched** to income levels similar to those presented in the scenarios.
One of these families will receive the **exact transfer amount** chosen by a participant in this study, while the other family will receive **nothing.**

This means your decisions are not just hypothetical—your choices could directly impact how much money a real family will receive from us.

Figure 22: Comprehension Check

Please pay attention and answer the question below correctly.

Based on the previous description of this study, which of the following statements is true?

- ◯ We will randomly select two families from the U.S. population who are part of the Prolific platform, and give ONLY ONE of them the chosen transfer.
- ◯ The transfers in the study are purely hypothetical and will not be given to any real families.
- ◯ We will randomly select two families from the U.S. population who are part of the Prolific platform, and give BOTH of them a chosen transfer.
- ◯ We will select one participant in this survey and give him/her the transfer amount.