Taxpayer Bias in Perceived Income Distributions

Díez-Alonso, Daniel

University of Warwick

30 September 2020

Taxpayer Bias in Perceived Income Distributions *

PhD Thesis - Chapter 1

Daniel Díez Alonso
University of Warwick
Supervisors: M. Almunia, L. Gadenne and C. Perroni

January 2021

*Contact: D.Diez-Alonso@warwick.ac.uk. I thank Prof. Miguel Almunia, Prof. Carlo Perroni and Prof. Lucie Gadenne for their guidance and supervision; Prof. Eugenio Proto, Prof. Kiril Pogorelskiy and the team of academicians and students at the PSE Summer School in Experimental Methods for useful suggestions and discussions; my family and friends, for their unconditional support.

Abstract

Contrary to the predictions of classical models, poor people tend to display lower support for redistribution owing to a biased perception about the income distribution in the society. This paper explores the potential effects of income taxes on perceived income distributions and income rank. I present a theoretical model that maps the information from the tax schedule onto informative signals that taxpayers use to infer their perceived position in the income distribution. Using probabilistic regression analysis on the General Social Survey, I find evidence supporting that the changes in perceptions of income ranks observed in the USA in the last decades were influenced by changes in the federal income tax system. To identify causality in a controlled environment, I then test the main predictions of the theoretical model by randomizing tax systems in two online experiments conducted on Amazon Mechanical Turk with American workers. Results of a large pilot identify statistically significant differences between individuals facing a proportional tax system with a unique average tax rate and those facing a progressive tax system with increasing marginal tax rates. Compared to no tax information (control), facing the progressive tax system used in the experiment induces a 12% higher perceived average income level and a 25% lower perceived probability of being above the average income level, while the proportional system does not generate significant differences. These findings encourage further research to identify the exact elements in a tax schedule that generate a bias that can affect support for redistributive policies.

JEL CODES: D03, D31, D63, D90, H21, H24.

1 Introduction

Standard models of preferences for redistribution based on the median voter theorem predict that societies with average income above the median should implement redistributive policies since a majority of the population could benefit from them (Meltzer and Richard, 1981). Given the levels of inequality observed in most western democracies, however, empirical evidence shows lower support for redistribution than expected, creating what the literature has called the *inequality-redistribution puzzle* (Benabou, 1996; Kenworthy and McCall, 2007).

From a theoretical perspective, the fact that changes in income inequality in a society rarely translate into changes in preferences for redistribution remains a puzzle. Part of the literature rationalised this apparent inconsistency by introducing other variables influencing taxpayers' preferences, such as the *prospects of upward mobility* (POUM) hypothesis (Bénabou and Ok, 1998). Nevertheless, those models still rely on individuals effectively inferring the income distribution of their society and their position within it, which is most often not the case. Several studies[1] have consistently found that taxpayers actually have biased perceptions of the income distribution. Moreover, their levels of preferred redistribution seem to respond to those biased perceptions rather than their true position, which can be corrected by giving them accurate information on their actual position or the real level of inequality (Cruces et al., 2013; Kuziemko et al., 2015; Karadja et al., 2017; Hvidberg et al., 2020).

The United States case illustrates this problem very well. Income inequality in the USA has been steadily increasing over the past 50 years and, while citizens seem to be aware of such a trend, their perceived distance from the average family income has barely changed (Figure 1). Similarly, support for redistribution has not increased, especially among families with the lowest levels of income, who would benefit the most from redistributive policies (Figure 2). Meanwhile, income taxes in the USA have been significantly reduced over the same period, especially for the rich. Have changes in preferences led to such changes in policy, or may the policies themselves shape preferences?

As Gimpelson and Treisman (2018) consistently estimate for a wide range of countries using

[1] Kenworthy and McCall (2007); Cruces et al. (2013); Hauser and Norton (2017); Fernández-Albertos and Kuo (2018); Hvidberg et al. (2020).

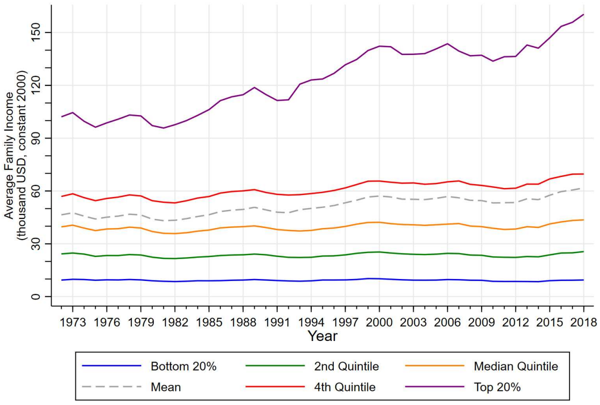
ISSP[2] data, people's support for redistribution seems to be driven by their perceived income distribution rather than the true one. Therefore, understanding what shapes those perceptions is a key political question. So far, the focus has been on analysing the role of reference groups (Cruces et al., 2013; Hvidberg et al., 2020), treating tax systems as the mere result of taxpayers' perceptions and preferences. However, what if tax schedules generated reference points that influenced income perception biases? This paper adds a new perspective to the discussion: if taxpayers infer distributional information from tax policies, the policies themselves may generate a bias that affects their public support.

Some behavioural economics papers have built on classical models to explain some of the most common biases, such as the representativeness heuristic when inferring population distributions from a reduced sample (Cruces et al., 2013). However, empirical evidence from survey data in the USA suggests that the predictions of those models do not explain the behaviour at the bottom tail of the income distribution. The representativeness heuristic in this context implies that individuals believe their self-selected reference group is more representative of the whole population than it really is, leading them to feel closer to the average than they actually are. In a model with such base-rate neglect, the probability of an individual perceiving herself as earning the average income should decrease as income approaches the boundaries (zero at the bottom and the highest possible income at the top) since the probability that other incomes in their reference group extend beyond those boundaries is zero. Nevertheless, I observe that the left tail of the income distribution does not behave this way. Specifically, it seems that all respondents with family incomes below a level that coincides with the income tax personal allowance (i.e. all levels of income exempt from paying federal income tax) perceive they are at the same distance from the average family income, regardless of their level of earnings. I argue that the tax-free allowance sets a reference value that influences individuals' perceived position.

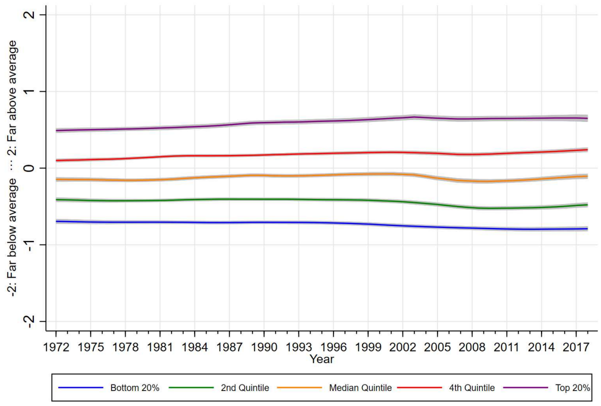
If this hypothesis is true, when thresholds of the tax bands or their marginal tax rates change, poorer households may perceive themselves closer to the average family income and therefore reduce their support for redistributive policies, allowing the social planner to cut taxes on higher-income groups. This is, in fact, what one observes when analysing the evolution of the Federal

[2]The International Social Survey Programme (ISSP) conducts annual surveys covering countries from North and South America, Africa, Europe, Asia and Oceania, with a module on social inequality.

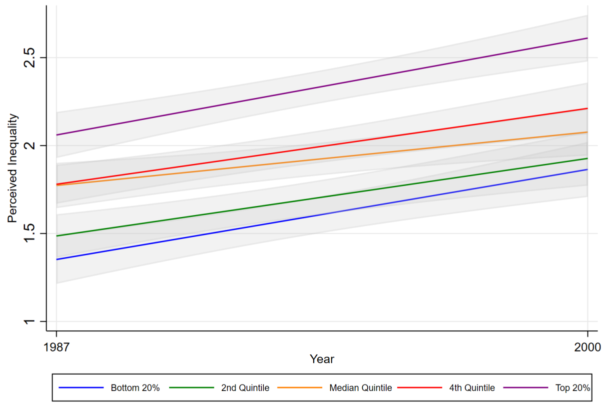
Figure 1: Evolution of household incomes, perceived position, and perceived inequality



(a) Family Income Distribution



(b) Perceived distance from the average family income

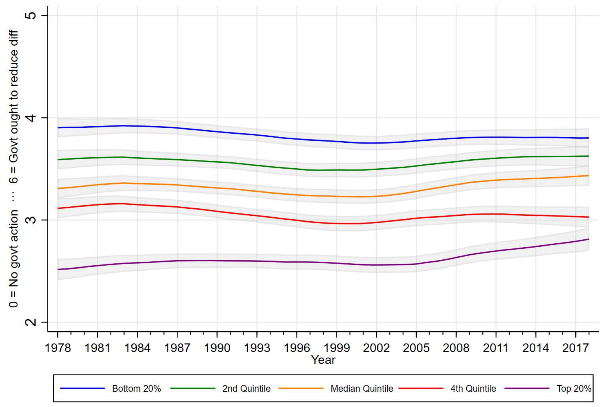


(c) Perceived pay inequality

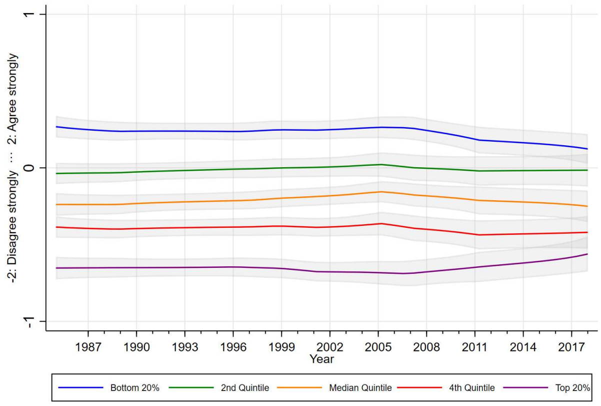
Source: Elaborated by the author using General Social Survey (GSS) and Census data.

Notes: Figure 1a shows the evolution of the average income per quintile and depicts the increase in inequality throughout the last 50 years. Figure 1b shows how families in each income quintile perceive their income compares to the average. Figure 1c shows the change in perceived pay inequality ($\ln \frac{perceivedexecutivepay}{perceivedskilledworkerpay}$) between 1987 and 2000, by income quintile.

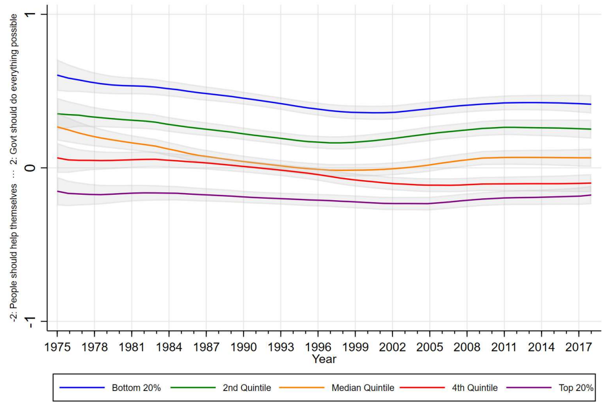
Figure 2: Evolution of preferences for redistribution, by income quintiles



(a) Should the government act to reduce differences between rich and poor?



(b) It is the government's responsibility to reduce income differences between rich and poor.



(c) Should the government improve the standard of living of poor people?

Source: Elaborated by the author using GSS and Census data.
Notes: The three graphs in this figure summarise the evolution of the answers to three questions included in the General Social Survey (GSS) through the last four to five decades. The answer to question (a) was a scale between zero (*No governmental action*) and six (*Government ought to reduce differences*). Question (b) was answered on a five-point scale, between minus two (*Disagree strongly*) and plus two (*Agree strongly*). The answer to question (c) was also a five-point scale, in this case recording positions between *People should help themselves* (minus two) and *Government should do everything possible* (plus two).

Income Tax Law in the USA for the last half of the 20th century: despite rising inequality, the government approved significant tax cuts for top incomes through several reforms of the federal income tax schedule. Simultaneously, poor households' perceived distance from the average family income and their level of support for redistributive policies barely changed. Two questions arise from those facts. First, why did low-income families fail to update their perceived distance from the average income, even when they perceived increasing pay inequalities in their society? And second, why did they not increase their support for redistributive policies?

In this study, I use data from the General Social Survey (GSS) to identify patterns on different measures related to perceived income distributions which the existing models with reference group bias cannot explain. I present a theoretical framework that builds on existing behavioural models and maps information from the income tax schedule onto informative signals that taxpayers use to infer income ranks, what I call the *tax burden heuristic*. I then test the model's predictions in an online experiment using Amazon Mechanical Turk, with participants based in the USA.

The final experiment of the paper finds statistically significant differences between individuals facing a progressive tax system with increasing marginal tax rates (informative) compared to those facing a proportional tax system with a unique flat rate (uninformative) and those in the control group who do not have any information about the tax system. In particular, taxpayers in the progressive tax group believe they are approximately 25% less likely to be above the average income in their reference group than those in the control and the proportional tax groups. They also estimate the average income to be around 12% higher. This finding is consistent with the hypothesis presented in this paper: setting different contribution levels above specific income thresholds affects how taxpayers infer their position in the income distribution.

This work contributes to the growing literature on (mis)perceptions of inequality. The new theory I present helps to explain the *inequality-redistribution puzzle* and provide a better understanding of the political economics of redistributive preferences, building on the work of Piketty (1995), Alesina and Giuliano (2011) and Cruces et al. (2013). Moreover, the conclusions of this study may also have implications for the literature of economics of happiness, which sustains that happiness is affected by relative income (rank) rather than absolute income levels (Boyce et al., 2010). Additionally, the notion that income tax thresholds constitute reference points that trigger reactions beyond labour supply and tax avoidance decisions is of relevance to the literature of public finance and should be taken into account when deriving optimal income taxes.

The rest of the paper is structured as follows: Section II develops the theoretical framework and presents the concept of *tax burden heuristic*; Section III provides supporting evidence from USA survey data, exploiting tax law changes that generated quasi-experimental variation; Section IV presents the online experiment aimed to test the predictions of the model; Section V concludes.

2 Theoretical Framework

One of the most consistent biases in perceived income rank is that people tend to feel closer to the average income than they really are. To explain that bias, Cruces et al. (2013) derived a model in which taxpayers suffer from base-rate neglect[3] and therefore take the income distribution of their reference group (to which they are self-selected by similarity) as if it was representative of the overall population. This leads people to believe they are closer to the average income than they truly are, as depicted in Figure 3. In the example of the figure, the blue solid line represents a hypothetically true cumulative distribution of incomes. The red dashed line would be the distribution of beliefs regarding income rank if income were bounded at both ends (minimum and maximum income levels), and respondents based their perception on their reference group, to which they are self-selected by similarity. Notice that although the bias increases as incomes depart from the average (marked by the grey vertical line at 0.2), it reduces again as incomes approach either boundary, since the probability of income levels beyond the boundaries is zero (in terms of the reference group, if you earn the minimum possible income, it cannot be that you know someone who earns less). A similar prediction would apply to a question regarding the position of the household relative to the income of the average family, since the probability of being above (below) the average tends to zero as incomes approach the lower (upper) boundary, for any possible distribution with positive densities along the feasible income range and zero density for income values beyond its boundaries. Hvidberg et al. (2020) argue that the bias at the tails may be larger due to mean-reversion since people at the bottom cannot underestimate their position and people at the top cannot overestimate it. However, that same argument would imply that the dispersion of answers should decrease as it approaches those boundaries[4], which is not what we observe in the data. Perceived position at the lower tail, the area that does not behave as predicted by a model with base-rate neglect, presents larger variation than at any other levels of income, perhaps signalling less accurate information.

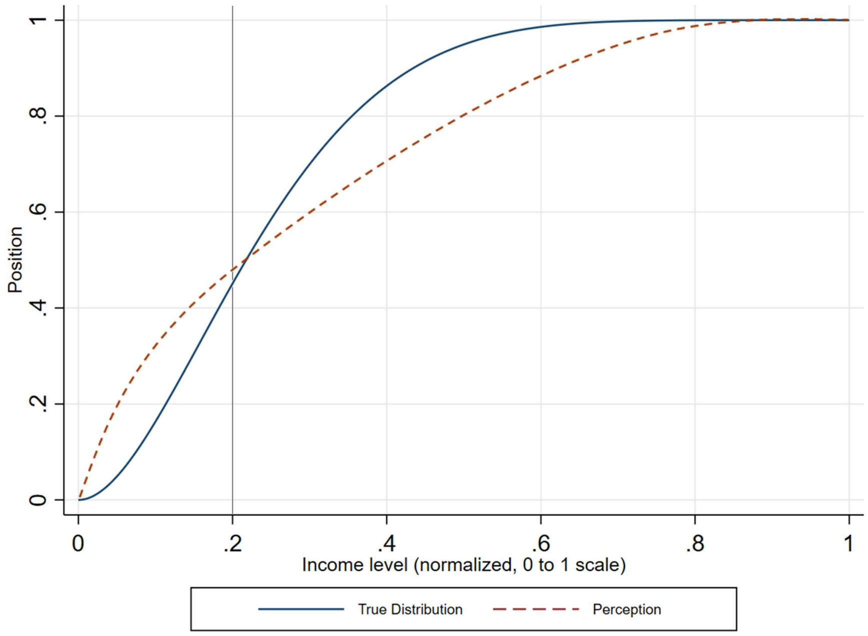
A question in the GSS specifically asks: *"Compared with American families in general, would you say your family income is...?" [Far below average - Below average - Average - Above average*

[3]The concept of base-rate neglect is such of sampling bias described by Kahneman and Tversky (1972).

[4]The dispersion could, however, increase for other reasons, hence challenging this argument. Nevertheless, the correlation between income and the perception bias revealed by the GSS data has a clear inflexion point way before the answers are close to the extremes of the scale, rather than flattening smoothly as answers get closer to their lower bound (see Figure 4).

- *Far Above average]*. In Figure 4a, I recoded the answers to show the proportion of people in each income level (normalised by the average family income of each year) perceiving themselves as below, around, or above the average for all survey years between 1984 and 2018. The grey vertical line represents the average income for the year of each survey. As expected, the proportion of people feeling below (above) the average is larger for lower (higher) incomes. Meanwhile, the proportion of people feeling around the average is the highest around the true average family income and decreases as income levels depart from it. It is relevant to notice that the proportion of respondents perceiving their income above the average remains positive and stable (around 10 per cent) at lower income levels, even if this is not consistent with the predictions of models based on similarity-selected reference groups. Similarly, the proportion of people feeling below the average

Figure 3: Perceived position under base-rate neglect and self-selected reference group



Notes: In this model, the central assumption is that the reference group generating the base-rate neglect only covers a subsample of the population distribution around the income level of the individual reporting her perceived position. Since incomes in that group are closer to the respondent's income than other incomes in the general population, the respondent underestimates her distance from the average and median incomes. However, for agents at the extremes of the distribution (i.e. with the highest or the lowest possible incomes), the probability of knowing a person with income beyond those boundaries is zero, hence reducing the bias. If *position* refers to the percentile, the lowest possible perceived percentile will depend on the frequency of the minimum income (not necessarily zero). However, if *position* refers to the probability to have an income above the average, as in one of the GSS questions, this should tend to zero as the respondent's income approaches the lower boundary of the income distribution, and it should tend to one as it approaches the other end. The graph reproduces the case for this second meaning of *position* relative to the average income, which is the central measure used in this paper and that of most relevance to redistributive preferences.

seems fairly stable (at around 60 per cent) for all people earning any income below a quarter of the average family income. Figure 4b represents the distribution of perceived position by true income rank. The blue dots and fitted solid line represent the mean value of answers per income group in each year, while the green dashed line serves as a reference value for a calculated measure of true distance from the average family income.

Given the top coding of the income variable in the GSS data, I focus the analysis on the anomaly observed at the left tail of the distribution. The average answer stops decreasing below a certain income level despite incomes being at the lowest end of the distribution, and uncertainty around perceived positions increases. This contradicts the prediction of models with base-rate neglect, which would yield more accurate perceptions at lower income levels, where the probability of being above the average family income tends to zero.

What is the driver of such increased bias at the lower tail? Is there a specific reference value from which people have a fixed perception of their position in the income distribution? I suggest an alternative source of information that may be shaping those perceptions: the tax burden.[5] Do taxpayers infer their income position from how much they pay in taxes rather than uniquely from the value of their income? Certainly, in any progressive tax system with increasing marginal tax rates, people earning the least will pay a lower proportion of their income in taxes, while those earning top incomes will be taxed at a higher rate (larger share). The tax burden is exactly the same for all income levels below the tax-free allowance (zero). If they were to infer their position based on how much they pay in taxes rather than how much they earn, all those in the tax-free band would perceive themselves at the same position, despite their differences in income.

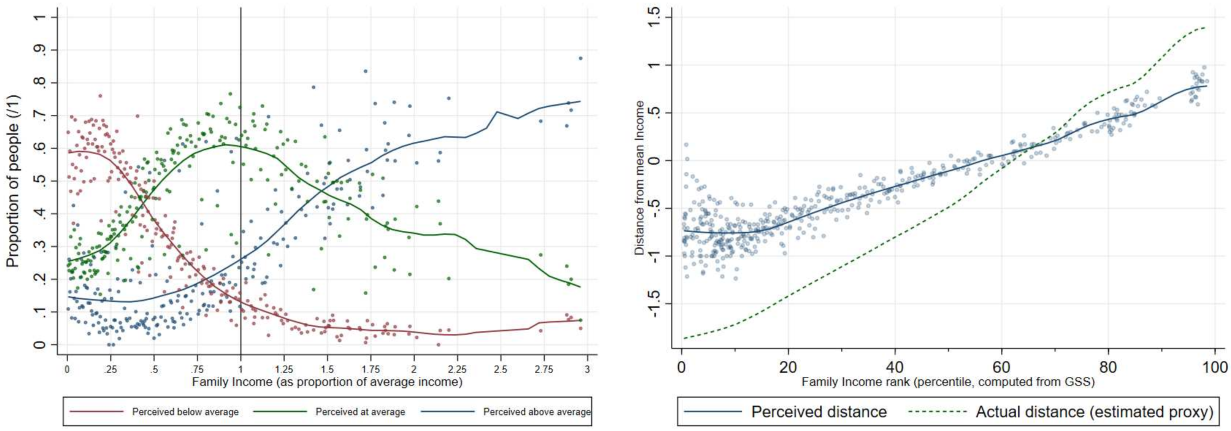
To formalise the hypothesis, I present a model that rationalises why taxpayers may infer their position in the income distribution from tax schedules and, specifically, their tax burden (average income tax rate). The model's main prediction is that a higher (lower) tax burden will be used as a signal of a higher (lower) level of relative income[6] and a higher (lower) rank in the income distribution.

Assume a country is ruled by a government that has to provide public services for its population,

[5]Throughout this paper, I refer to the proportion of income paid in taxes as the *tax burden*.

[6]The term *relative income* is used in this paper as income level relative to the average level of income in the population.

Figure 4: Perceived position relative to the average family income



(a) By relative income

(b) By true income rank

Source: Elaborated by the author using GSS and Census Data.
Notes: These two graphs plot the distribution of answers to the GSS Question 202: *Compared with American families in general, would you say your family income is...*, which has five possible answers: *Far below average; Below average; Average; Above average; Far above average.* The left figure (a) shows the proportion of respondents perceiving their family income below, around, or above the average family income by the true proportion of their family income with respect to the average. Families with the average income are those around value one in the horizontal axis, while those on the left (right) are families with incomes below (above) the average. The blue dots and fitted polynomial (blue line) in the right figure (b) represent the average score of all answers in a given year, with zero denoting respondents who perceive their family income is at the USA average and values one and two (positive and negative) for the different intensities of the above and below average answer options. The horizontal axis in the second graph denotes the true percentile of the respondent's family income. The green dotted line is only plotted for visual reference and is an arbitrary (albeit proportional) measure of distance from the average family income based on the relative size of the family income at each percentile of the distribution compared to the average income. Alternative measures could be computed, leading to different slopes, but they would all cross the true average income level (zero on the vertical axis) at the same point, slightly above the 60th percentile, and cross the blue fitted line afterwards.

and it needs to fund such public expenditure *G* through tax revenues. In the simplest case, assume the only tax in place is the income tax, so $\sum_i (\tau_i \cdot y_i)$ represents the total budget, collected from each individual's income y_i at an average tax rate τ_i. If individuals were subject to a non-progressive (purely proportional) system, with income taxed at a flat rate $\tilde{\tau}$, the flat rate would be determined by $\frac{G}{Y}$, where Y represents the total addition of individual incomes in the population, $\sum_i (y_i)$. The total tax revenue collected by the tax authority given a tax schedule S thus can be expressed as $G(Y \mid S) = \tilde{\tau} \cdot Y$.

However, in most countries, the income tax schedule is composed of a progressive system of thresholds and increasing marginal tax rates above each income threshold. Such a system is thus defined by a set of *K+1* thresholds ($z_0 = 0 < z_1 < \cdots < z_K$) and a corresponding set of marginal tax rates ($r_0 < r_1 < \cdots < r_K$) where any marginal tax rate r_k is applied to any earned income between z_k and z_{k+1}.[7] The tax burden of an individual, her proportion of income paid in taxes, is measured by the average tax rate (τ_i) as in Equation 1, and the total tax revenue can be expressed in this case as $G(Y \mid S) = \int_0^\infty y\,\tau(y)\,f_y\,dy = \sum_{n=1}^{N} y_n\,\tau(y_n)$, where N is the total number of people subject to the tax system and f_y the number of them earning income y.

$$\tau(y_i) = \tau_i = \frac{\sum_{k=0}^{K}\{r_k \cdot min\,(z_{k+1} - z_k, max\,(0, y_i - z_k))\}}{y_i} \quad \text{where } z_{K+1} = +\infty \tag{1}$$

Therefore, in any given income tax system, the range of possible tax burdens has a lower bound at 0 and an upper bound equal to the top marginal tax rate, since $\lim_{y_i \to \infty} \tau_i = r_K$. This also means that the reference rate $\tilde{\tau}$ that would apply if all taxpayers faced the same proportional flat rate must lie within $(0, r_K)$.

Any given tax schedule with progressive tax rates must be constructed so that those contributing above the hypothetical proportional share $\tilde{\tau}$ compensate the forgone revenue from those contributing below that proportional share. Calling $\tilde{y}$ the income level such that $\tau(\tilde{y}) = \tilde{\tau}$, Equation 2 represents such balance condition:

$$\int_0^{\tilde{y}} (\tilde{\tau} - \tau(y))\; y\, f_y\, dy = \int_{\tilde{y}}^{\infty} (\tau(y) - \tilde{\tau})\; y\, f_y\, dy \tag{2}$$

If agent i earns income y_i with a resulting tax burden $\tau(y_i) < \tilde{\tau}$, Equation 2 can be re-written as Equation 3, which establishes the relation in terms of transfers between incomes below and above that of agent i to ensure a target revenue G.

$$\int_0^{y_i} (\tilde{\tau} - \tau(y))\; y\, f_y\, dy + \int_{y_i}^{\tilde{y}} (\tilde{\tau} - \tau(y))\; y\, f_y\, dy = \int_{\tilde{y}}^{\infty} (\tau(y) - \tilde{\tau})\; y\, f_y\, dy$$

$$\int_0^{y_i} (\tilde{\tau} - \tau(y))\; y\, f_y\, dy = \int_{y_i}^{\infty} (\tau(y) - \tilde{\tau})\; y\, f_y\, dy \tag{3}$$

[7] A tax-free allowance would be represented by $r_0 = 0$, and the initial tax band would then kick in at rate r_1 for any unit of income above z_1. Similarly, all units of income above z_K will be taxed at the maximum marginal tax rate, r_K.

2.1 Theoretical predictions

Imagine a society with only three types of agents: those earning a high income (y^H), those with medium income (y^M), and those earning a low income (y^L). The government needs to raise a fixed amount G through income taxes. If it implements a proportional tax rate $\tilde{\tau}$ (no progressivity), individuals cannot relate their tax burden to their position in the income distribution (since all individuals pay the same tax rate, independent of their income).

Alternatively, the government could consider a very simple progressive tax system with two marginal tax rates, r_0 up to income level y^*, and then a higher r_1 for any income exceeding that level. This means the possible range of tax burdens for income earners in this society is defined between r_0 and r_1, with $r_0 < \tilde{\tau} < r_1$.

Recalling the balance condition of progressive tax schedules (Equation 2), there must be at least one group with a tax burden below the proportional share $\tilde{\tau}$ and at least another group with a tax burden above it. Since the minimum possible tax rate is r_0, which by definition must be smaller than $\tilde{\tau}$, this implies the following information can be inferred by taxpayer i when realising their tax burden, without knowing which income group they belong to, and even without knowing $\tilde{\tau}$:[8]

$$P(y_i \in L \mid \tau(y_i) = r_0) > 0\,;\; P(y_i \in L \mid \tau(y_i) = r_1) = 0 \tag{4}$$

$$P(y_i \in H \mid \tau(y_i) = r_0) = 0\,;\; P(y_i \in H \mid \tau(y_i) = r_1) > 0 \tag{5}$$

In addition, assuming $\tilde{\tau}$ is known to taxpayers, the tax system automatically becomes more informative:

$$P\left(y_i \in H \mid r_0 \leqslant \tau(y_i) < \tilde{\tau}\right) = 0 \Rightarrow P(y_i \in M \cup L \mid r_0 \leqslant \tau(y_i) < \tilde{\tau}) = 1 \tag{6}$$

$$P\left(y_i \in L \mid \tilde{\tau} < \tau(y_i) \leqslant r_1\right) = 0 \Rightarrow P(y_i \in M \cup H \mid \tilde{\tau} < \tau(y_i) \leqslant r_1) = 1 \tag{7}$$

And incorporating the balance condition (2):

$$P(y_i \in M \mid \tau(y_i) < \tilde{\tau}) \propto P\left((\tau(y^H) - \tilde{\tau})\, y^H\, f^H > (\tilde{\tau} - \tau(y_i))\, y_i\, (N - f^H)\right) \tag{8}$$

[8] In these equations, the expression $P(y_i \in L)$ denotes the probability that income y_i belongs to the group of low income earners.

$$P(y_i \in M \mid \tau(y_i) > \tilde{\tau}) \propto P\left((\tilde{\tau} - \tau(y^L))\, y^L\, f^L > (\tau(y_i) - \tilde{\tau})\, y_i\, (N - f^L)\right) \tag{9}$$

Notice that the probabilities in (6) and (7) both decrease as the distance between $\tau(y_i)$ and $\tilde{\tau}$ increases, everything else constant. Therefore, a lower tax burden below the proportional share implies a lower probability of belonging to income group M (in favour of group L), and a higher tax burden above the proportional share implies a lower probability of belonging to income group M (in favour of group H), for any prior belief on the distribution of incomes. Combining those conditions, this implies the conditional probability of belonging to income group L weakly increases when $\tau(y_i)$ decreases, while the conditional probability of belonging to group H weakly increases when $\tau(y_i)$ increases: $\frac{\partial P\left(y_i \in y^L | \tau(y_i)\right)}{\partial \tau(y_i)} \leq 0$ and $\frac{\partial P\left(y_i \in y^H | \tau(y_i)\right)}{\partial \tau(y_i)} \geq 0$, everything else constant.

Such calculations of probabilities are complex and thus have a high cognitive cost. However, the qualitative conclusion of this model can be approximated by a heuristic, which I call the *tax burden heuristic*. This heuristic consists in inferring a higher position in the income distribution from a higher *relative* tax burden (τ_i relative to the maximum rate in the tax system or the proportional share $\tilde{\tau}$, when known). Notice that a *higher position* in this model means a larger share of population income accumulated below the respondent's income, which can imply a higher rank or a higher income relative to the average (larger distance from the average income if above it, and shorter distance if below), or both.

3 Empirical evidence in the USA

To test the model's predictions, in this section, I show supportive evidence from the USA using household panel data from the General Social Survey (GSS) for years 1984-2018, exploiting the exogenous variation of tax burdens introduced by the Income Tax Law changes within this period. Although the GSS data is available for years as early as 1972, I restrict my analysis to the period from 1984, since Personal Income Tax thresholds were only indexed to inflation from that time. While this survey has its limitations regarding the accuracy of income measures (the public dataset groups households by income ranges rather than disclosing the exact answer from each household, which results in top coding), it is the only publicly available longitudinal dataset including income data, perceived position relative to the average family income, and preferences for redistribution. It also includes the more ambiguous question on "position in the social ladder" from the ISSP module.

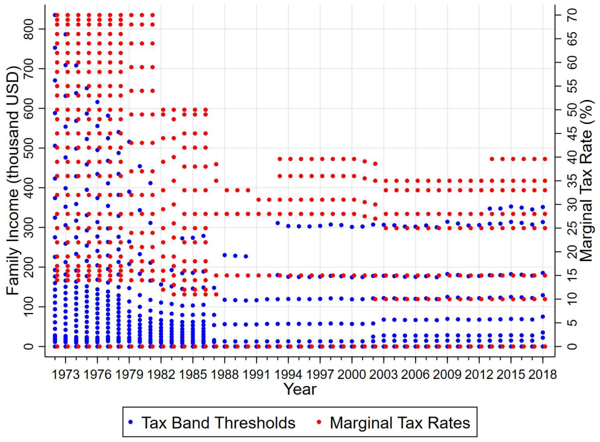
3.1 Historical context and main reforms

The Federal Personal Income Tax Law in the United States has experienced massive changes during the last 50 years. As depicted in Figure 5a, tax thresholds were only indexed to inflation after 1984. This resulted in progressive changes in the real value of those thresholds during the first ten years represented in the graph, although these were most probably not as salient for taxpayers as those resulting from changes in the tax law. All income values are counted in real terms, as constant US dollars with base year 2000. Each blue dot represents the threshold for each additional tax band, and the red dots depict their corresponding marginal tax rates. One can observe a very relevant reduction in the number of tax bands (blue dots) for the first third of the series, followed by a slow progressive increase in subsequent years. Due to the scale of the graph, one cannot notice the changes on the bottom band, the one with zero marginal tax rate (tax-free allowance), but this is represented in the graph on the right, Figure 5b. The dashed black line shows the evolution of the tax-free allowance (right vertical axis), and the same chart depicts the evolution of tax burdens for the average family income level within each income quintile (left vertical axis). Changes in the tax-free allowance mechanically have a more significant impact on the burden of lower incomes. However, one can observe how, in most periods where the tax-free allowance was increased, it was actually the burden of the highest income quintile that was most reduced, implying that those reforms also included significant tax cuts for higher incomes.

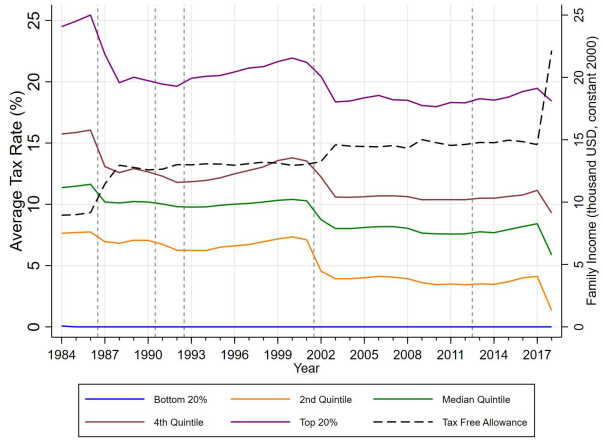
The first relevant change in the income tax during 1972-2018 was introduced by the Tax Reduction and Simplification Act of 1977 and the Revenue Act of 1978, which consisted in removing ten of the intermediate tax thresholds with the main purpose of simplifying the tax schedule. Tax rates were kept constant for the remaining bands, and thus the subtraction of intermediate thresholds effectively resulted in a reduction of taxes on most income levels, especially the highest ones.

The second relevant reform arrived with the Tax Equity and Fiscal Responsibility Act of 1982, which removed the three top tax bands, resulting in a huge drop of the top threshold and corresponding marginal tax rate. That threshold was raised again progressively in the following two years, with the addition of two new bands. The Economic Recovery Tax Act of 1981 also set that

Figure 5: Evolution of Federal Income Tax for married couples in the USA



(a) Summary of band thresholds and tax rates



(b) Evolution of tax burdens, by income quintiles

Source: Elaborated by the author using Census Data.

Notes: In the left graph (a), the blue dots represent the number of bands set by the Federal Income Tax Law in each year. Their position on the (left) vertical axis represents the threshold for each band. The vertical distribution of these blue dots gives an idea of the piece-linearity of the tax schedule. Meanwhile, the red dots represent the marginal tax rate assigned to each corresponding income band. The chart on the right (b) summarises how tax burdens, measured as the proportion of gross income paid in taxes, changed over time for the average income of each income quintile. The dashed black line (right vertical axis) was added to complement the information included in Figure 5a, since it would correspond to the lowest blue dot in every year, indistinguishable from the other bands in most years. The vertical dashed grey lines highlight the years with major changes to the Federal Income Tax Law.

all tax thresholds would be adjusted for inflation after 1984, except for the first year after a new law directly affected the value of a tax band.

The most relevant change in the history of income taxes in the USA was implemented in 1987 after the US Congress passed the Tax Reform Act of 1986. The number of tax bands was reduced from 15 to only 6, and tax rates dropped significantly.

A few years later, the Omnibus Budget Reconciliation Act of 1990 started raising the top marginal tax rate, but it was not until 1993 that two additional tax bands were added at the top of the scale by the Omnibus Budget Reconciliation Act of 1993. This reform was followed by the most stable period regarding the Income Tax Law until the Economic Growth and Tax Relief Reconciliation Act of 2001 introduced new tax reductions.

The last big changes of the Income Tax Law took place in 2013, with the introduction of an additional band at the top of the scale, and in 2018, with a substantial increase of the tax-free allowance.

3.2 Description of the survey data

The General Social Survey (GSS)has been running in the USA in a panel format since 1972, either annually or biennially, and is publicly available on the website of the NORC (University of Chicago).[9] It is one of the few large panel surveys that has consistently recorded some measure of income rank perception, together with the usual socio-demographic characteristics of respondents. I use this data to find correlations that support the model presented in the previous section to motivate an online experiment that allows testing the model's predictions with identified causal interpretation. I present the results for the period 1984 to 2018, including 22 rounds and comprising 4510 households, of which more than half were interviewed in at least 11 different rounds.

The main dependent variable of interest is the answer to Question 202: *"Compared with American families in general, would you say your family income is...? [Far below average; Below average; Average; Above average; Far above average]"*. In addition, in the years where questions from the International Social Survey Programme (ISSP) were incorporated into the GSS, a similar (but less specific) question is also available. This question has been broadly used for cross-country

[9] `http://www.gss.norc.org/` [last accessed on 14/09/2018].

comparisons in the literature of perceived income rank and preferences for redistribution[10]: *"In our society there are groups which tend to be towards the top and those that are towards the bottom. Here we have a scale that runs from top to bottom. Where would you put yourself on this scale? [1=Top ········· 10=Bottom]"*. I run the analysis on both questions, which I refer to as Q1 and Q2, respectively, through the rest of this section.

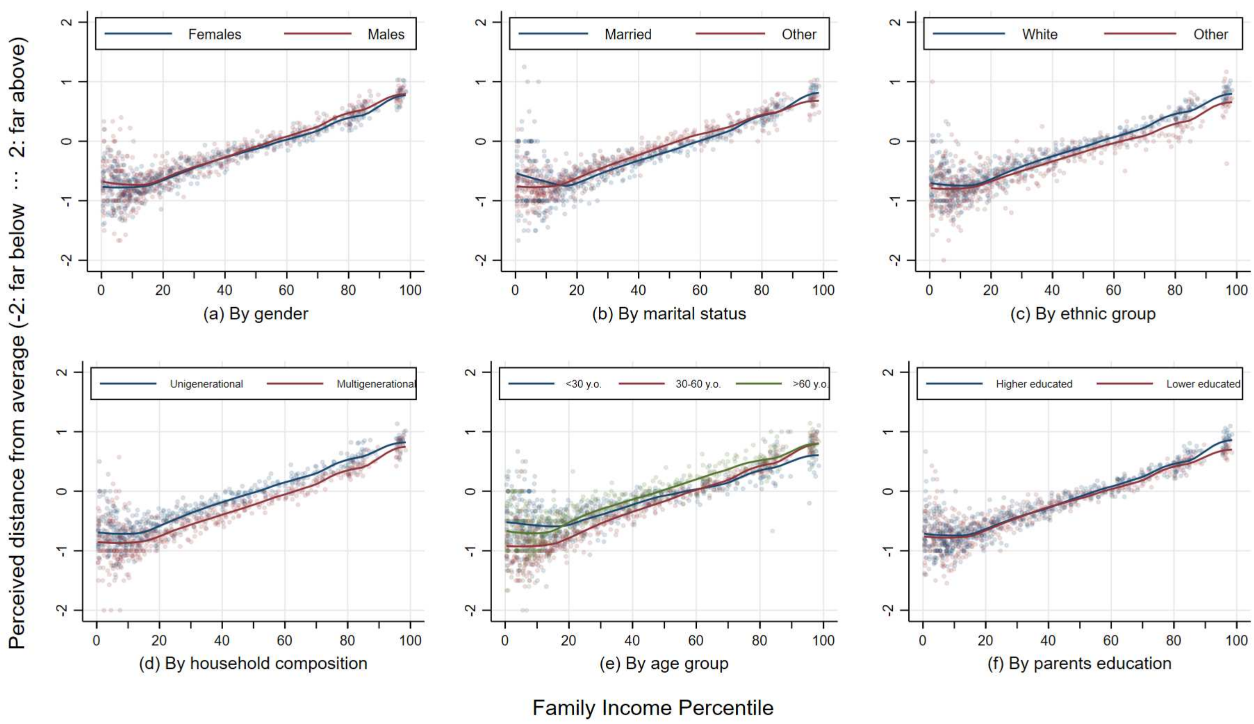
As mentioned in the introduction, and as consistently identified in the literature, people below the mean tend to overestimate their position while the reverse happens to people above the mean, and both groups underestimate their distance from the average or median income (Cruces et al., 2013; Hvidberg et al., 2020). More surprisingly, while perceived position (distance from the mean) seems to be correlated with income for most of the range, there is a threshold at the lower (left) tail of the distribution from which perceived position does not decrease with lower income levels. Specifically, the correlation between income and perceived position within the bottom quintile of the income distribution disappears, with the average perceived position being constant for households below the 20th income percentile. This anomaly, incompatible with predictions based on reference groups and base-rate neglect, is consistent across different socio-demographic dimensions (Figure 6).

The ISSP modules also include a question on preferences for redistribution, which I call Q3: *Some people think that the government in Washington ought to reduce the income differences between the rich and the poor, perhaps by raising the taxes of wealthy families or by giving income assistance to the poor. Others think that the government should not concern itself with reducing this income difference between the rich and the poor. Think of a score of 1 as meaning that the government ought to reduce the income differences between rich and poor, and a score of 7 meaning that the government should not concern itself with reducing income differences. What score between 1 and 7 comes closest to the way you feel?*. In Figure 7, I show how answers to this question highly correlate with the inverse of the tax burden.

The survey does not include any question on how much taxes the respondent is paying, but I use the yearly family income variable to approximate the corresponding income tax in a given year. Although this measure is not exact due to the lack of information on tax deductions, it is expected

[10] Alesina and Giuliano (2011) and Evans and Kelley (2004), among others.

Figure 6: Distribution of perceived position, by socio-demographic characteristics



Source: Elaborated by the author using data from the GSS, years 1984 to 2018.
Notes: For each socio-demographic variable, a dot represents the average answer of households within the same percentile in a given year. The solid lines are fitted polynomials (Epanechnikov kernels) superposed for visual clarity.

to be highly correlated with the real unobserved measure. In the main analysis, I normalise it by the maximum marginal tax rate in that year (potential maximum tax burden as income tends to infinity). This way, I can capture the idea of the relative tax burden, as described in the theoretical model.

The correlation between the answers to each of those three questions and the income level can be observed in Figure 7. The income measure on the horizontal axis has been rescaled for every year so that value zero corresponds to the tax-free allowance threshold (red vertical line). One can see that the distribution of answers to each of the three questions (blue) has an inflexion point exactly around the tax-free allowance, which is only observed in the distribution of the tax burden rather than on the actual distributions of percentile and distance from the average income.

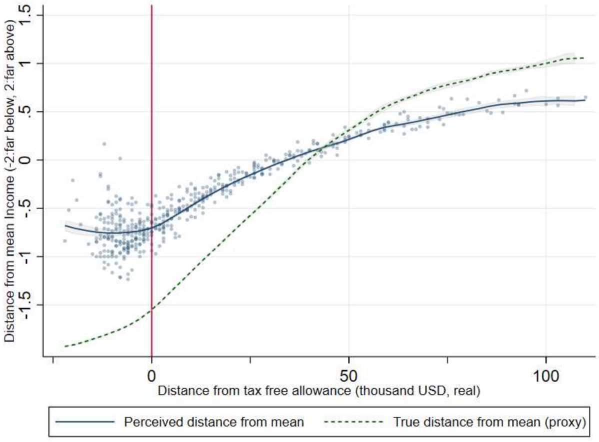
3.3 Empirical analysis

Exploiting the exogenous shocks to tax burdens introduced by the changes in the income tax law, I explore the explanatory power of such a source of information to taxpayers. I regress the perceived position (Q1) on the relative tax burden (average tax rate on the respondent's household income, $\tau(y_{it})$, relative to the maximum possible tax rate r_{K_t}) and control for family income (y_{it}), a recent change in financial situation ($FinChange_{it}$), socio-demographic variables, and year (t) fixed effects. Given that the dependent variable is an ordered scale, I estimate the coefficients using an ordered probability model (OProbit), both as cross-sections and as a panel with individual fixed effects (α_i). The sample used in the analysis excludes households above the 90th income percentile, affected by the top coding of the income variable.

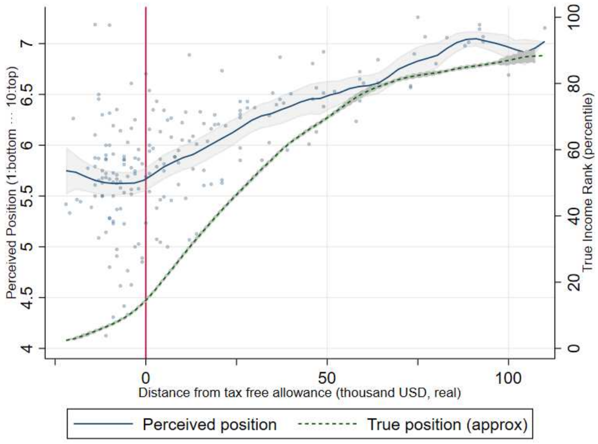
$$PerceivedPosition_{it} = \beta_1 \frac{\tau(y_{it})}{r_{K_t}} + \beta_2 \ln(y_{it}) + \beta_3 FinChange_{it} + \Gamma' \text{Controls}_{it} + \alpha_i + \epsilon_{it} \quad (10)$$

Results are shown in Table 1, columns 1 to 4. The first 3 columns treat the observations as cross-sections, clustering errors by household ID. The fourth column corresponds to the panel regression. The relative tax burden has a highly significant explanatory power for perceived position, and it is robust across specifications. Moreover, its size is not negligible: the coefficient for a percentage

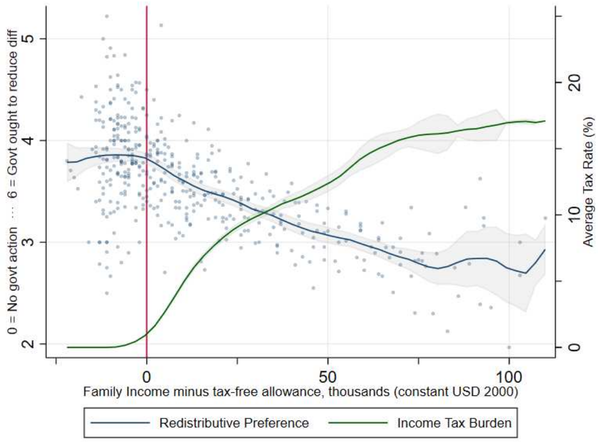
Figure 7: Distribution of answers to GSS questions, by distance to personal tax allowance



(a) Q1: Perceived distance from average family income



(b) Q2: Perceived position in society



(c) Q3:Preference for redistributive policies

Source: Created by the author using GSS and Census data for years 1984-2018.
Notes: The blue dots and fitted polynomial in each of the three graphs represent the average answer to three questions of the GSS detailed in previous paragraphs. In these figures, the income of each respondent has been deducted their corresponding personal allowance according to the Income Tax Law in place the year of each survey round. Therefore, the horizontal axis represents the distance between the respondent's household income and the tax-free allowance, adjusted for inflation to 2020 US dollars. The green dashed line in graph (a) is the same reference measure of Figure 4b based on the actual distance of the respondent's household income from the average family income. In graph (b), the green dashed line represents the percentile of the respondent's household income (right vertical axis). In graph (c), instead of a measure of income position, the green line represents the income tax burden of the respondent's household (the proportion of income paid in taxes) on the right vertical axis.

point change in relative tax burden is close to the effect of an increase of family income by one per cent. The other four columns (5 to 8) correspond to the same regressions using Q2 as the dependent variable. In this case, the statistical significance of household income practically disappears, while the relative tax burden remains with a very stable coefficient value and very high statistical significance (above 99% confidence level). However, the explanatory power of the regressions for Q2 (measured by the Pseudo R-squared) is significantly lower, reflecting that the ambiguity of the question makes it less related to household income (people may interpret very differently what the social ladder is based on).

Table 1: Effect of the relative tax burden on the perceived position

	Q1: Distance from average income				Q2: Position in the social ladder			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax Burden (as % of potential maximum)	0.041***	0.027***	0.037***	0.040***	0.016***	0.013***	0.015***	0.019***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Family Income (log, constant USD 2000)		0.252***	0.050***	0.047***		0.050***	-0.036*	-0.014
		(0.013)	(0.015)	(0.015)		(0.018)	(0.022)	(0.021)
Change in financial situation			0.364***	0.336***			0.132***	0.105***
			(0.008)	(0.007)			(0.014)	(0.013)
Observations	49,379	49,379	49,042	49,217	14,515	14,515	12,333	12,360
Year FE	NO	NO	YES	YES	NO	NO	YES	YES
Demographic Controls	NO	NO	YES	NO	NO	NO	YES	NO
Household FE	NO	NO	NO	YES	NO	NO	NO	YES
Pseudo R-squared	0.0961	0.102	0.138	0.315	0.0122	0.0123	0.0275	0.306
Number of Households				3,904				2,841

Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Demographic controls include age, gender, marital status, and educational level.

3.4 Limitations

Although the results of this section are robust to different econometric specifications and measures of the tax burden, they should be taken with caution. This section provides some suggestive evidence of a relation between tax schedules and perceived distributions, but the data and context used for the analysis suffer severe limitations. First, the income data from the GSS is encoded within income ranges, introducing variation in income that is purely a result of changes in the definition of such ranges with different versions of the survey used in different years. Second, that same limitation of the income variable equally affects the calculation of income taxes, which is also likely to introduce spurious variation in my measure of tax burden. Third, the lack of information on tax deductions and benefits reduces the accuracy of the measure of tax burden. Finally, my hypothesis suggests a reciprocal effect of taxpayers electing policies that simultaneously affect their perception and may subsequently change their policy preferences. Since elected politicians introduced the tax changes, those are likely to be endogenous, and it is difficult to identify causality. Moreover, changes in the income tax law were often just a part of broader fiscal policies affecting several taxes, and even within the tax law, they involved changes in the number of bands, thresholds, and tax rates, all at the same time. To accurately understand what elements of the tax schedule are informative to taxpayers, we would need a change that only affected either the tax-free allowance, the other tax band thresholds, or the marginal tax rates, one at a time.

4 Experimental Approach

To overcome the limitations presented in the previous section, I designed an online experiment to analyse the influence of progressive tax schedules on perceived income distributions. I set up a synthetic environment that tried to mimic a real-life situation: participants had to perform some tasks from which they earned a level of income subject to a tax. After learning their earned income and the tax system in place, participants reported measures of perceived statistics of the income distribution. Participants faced a tax schedule that was randomised across treatment groups. This experiment focused explicitly on the role of tax schedules creating progressivity (differences) in the tax burden, and compared a proportional tax system (same flat tax rate on all incomes) with a very simple progressive tax system with two (increasing) marginal tax rates. Therefore, this experiment is limited to testing the main underlying assumption of the theoretical model: that tax systems directly impact perceived income distributions. Further experiments will be necessary to test the specific theoretical predictions comparing different levels of relative tax burden.

An online experiment provides the opportunity to introduce very specific changes to the tax system applied, one at a time, holding all other variables constant. It also allows controlling the set of information made available to respondents at every moment, thus ensuring a situation as close as possible to the one assumed in the theoretical model. I programmed the experiment using oTree, an open-source software for experiments and surveys created by Chen et al. (2016) based on Python, and implemented it on Amazon Mechanical Turk with location restricted to USA workers only. Respondents were matched to records of other 99 previous participants, creating an overall group of 100 individuals on which they had to infer different income distribution statistics. Monetary amounts in the experiment are expressed in Experimental Currency Units (ECU) to represent realistic yearly income values closely. The equivalence with real-world currency is 50,000 ECU to 1 USD. All tasks and questions are incentivised, with participants' final payment based on their performance on the tasks and the accuracy of their answers. A random lottery selects which questions to pay, eliciting truthful beliefs by eliminating hedging opportunities. In addition to the amount earned through the experiment, all participants were paid a participation fee of 0.40 USD upon completion.

All components and concepts in the experiment were tested with volunteers of different ages and backgrounds to ensure a correct level of comprehensibility. The experiment received internal ethical approval from the Department of Economics at the University of Warwick and was registered at the American Economic Association's registry for randomized controlled trials.[11] Further technical details, including the set of instructions, explanations of the reward structure chosen to ensure the right incentives, screenshots of different sections and more, can be found in Appendix A.

4.1 Structure of the experiment

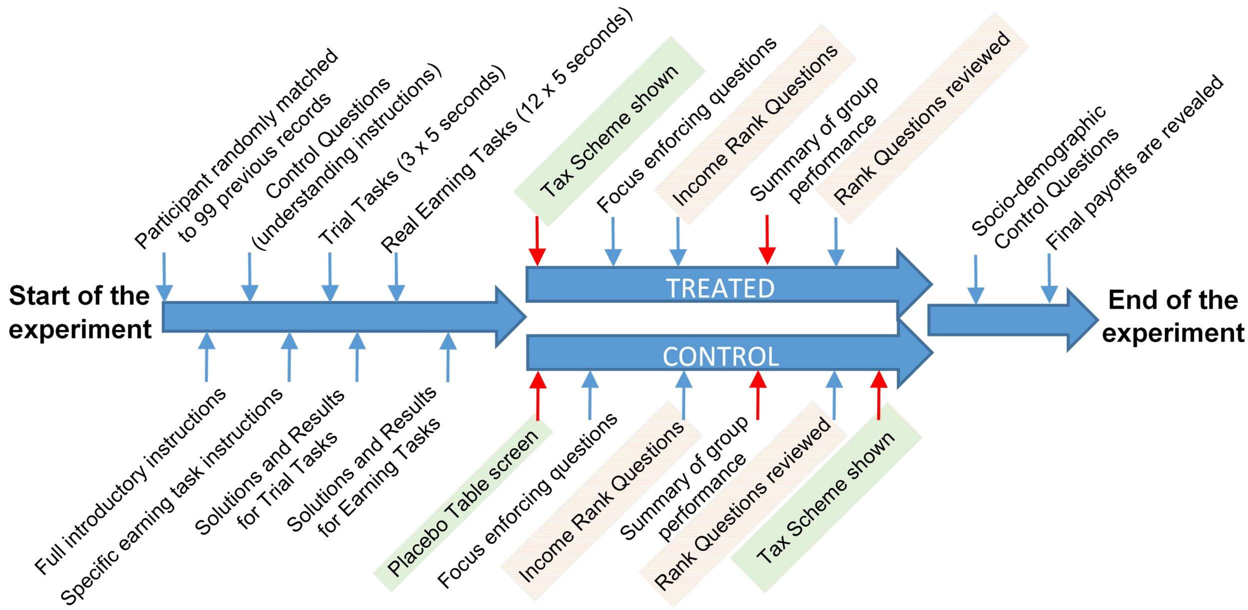
The experiment was implemented in May 2020 and had an average duration of approximately 12 minutes. As summarised by Figure 8, the experiment has a first part where participants had to solve twelve tasks to earn their gross income, and they were later given different pieces of information before being asked the target questions on statistics of the perceived distribution. There were two treated groups: one facing a proportional flat tax system and another facing a progressive tax system with two bands. The control group saw a placebo screen displaying a table with similar numbers but framed as a *captcha*[12] to prove they were not a bot and were only revealed the tax system at the end of the experiment. After the first measures of perceived statistics were collected, all participants saw the results of a subset of other players in their group (reference group), and they were given the chance to modify their answers. This provides a measure of elasticity of respondents' beliefs to a set of information that is objectively informative (even if partial), adding validity to the findings of this study. This information was disclosed after treated groups had already learnt the tax system

[11] The ethical approval reference is *ECONPGR 05/19*, and the AEA RCT ID code is *AEARCTR-0003365*.

[12] Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA).

in place and had provided an initial measure of their perceived statistics of the income distribution.

Figure 8: Timeline of the experiment



Notes: Randomization is highlighted by the red arrows. Allocation to groups was also random, and the green background boxes highlight the screens that were different for the treatment and control groups. Orange background boxes highlight the main target questions of the experiment.

The questions asked in this experiment were:

1. **Self-perceived rank (Q1):** *How do you think you rank among the 100 players in your group?*

2. **Perceived average income (Q2):** *What do you think was the average income in your group (you included)?*

3. **Perceived position relative to the mean (Q3):** *With what probability do you think your income may be above the group average?*

4.2 Identification strategy and randomization

As Figure 8 shows, the experiment was designed to compare between individuals in the control group and each of the two treatment groups. Identification was achieved through random allocation of participants to these three groups:

- **Treatment 1 (T_1):** participants saw the proportional tax schedule before being asked the target questions. The proportional tax system consisted of a unique average tax rate for all levels of income.

- **Treatment 2 (T_2):** participants saw the progressive tax schedule before being asked the target questions. The progressive tax system consisted of a tax-free allowance and a marginal tax rate above that amount.

- **Control:** participants saw a placebo screen instead of the tax schedule, which was presented as a *captcha* to verify they were not bots. They were aware their income would be subject to taxes at the end of the experiment, but they were not told the exact tax system until the end.

To ensure that treated participants paid attention to the tax schedule and control individuals exerted similar effort on the placebo screen, all participants were required to input a few pieces of information. They had to type their gross income earned from the 12 tasks, their proportion of income paid in taxes (control individuals were asked to calculate a ratio of the two values shown in the placebo screen), and the maximum level of income that paid zero taxes (control individuals were asked to input the highest of the values displayed in a table).

Participants were asked to report their beliefs on the statistics Q1, Q2, and Q3 after seeing those tax schedules (or the placebo screen if in the control group). Right after introducing these measures of beliefs, participants saw a little graph showing the income of other nine players in their reference group and were given the opportunity to amend their answers. In this experiment, participants were randomly matched with one of two groups of previous records: high performing or low performing, the latter displaying lower median and average incomes (see Figure 11).

Therefore, the measures of perceived statistics on the distribution taken at first (X_{i1}) were based on the participant's own income (y_i, measured in thousand ECU), the range of possible incomes ($M \in [0, 120]$), prior personal knowledge and biases (ϵ_i, unobservable) and, in the case of treated individuals, the information from the tax schedule (T_1 or T_2). As mentioned before, after collecting these measures, participants were shown partial information on their reference group (g_H or g_L) and were allowed to modify their answers, hence reporting a final measure $X_{i2}(y_i, M, T_i, g_i, \epsilon_i)$. Variation between individuals occurs in income (endogenous), tax system (exogenous), and reference group (exogenous).

Figure 10 depicts the distribution of tax burdens across income levels (left graph) and the difference between gross and net incomes under each tax system (right graph). Those earning $60,000$ ECU were taxed the same under both systems. Remember, participants in the control group did not count with any of this information before reporting their perceived statistics and were only shown the tax system in place and their tax due at the end of the experiment, after all target questions had been answered.

Based on the theoretical framework proposed in Section II of this chapter, taxpayers may derive information about their position from progressive tax schedules, but they cannot do so from proportional tax systems. This is because a proportional flat tax rate implies that everybody pays the same share of income in taxes (and hence there is no difference in tax burden across income

Figure 9: Randomized tax schedules shown to participants

Income (ECU)	ATR*
From 0$	15%

(a) Flat/Proportional Tax (Treatment 1)

Income (ECU)	MTR*
0$ to 40,000$	0%
From 40,001$	45%

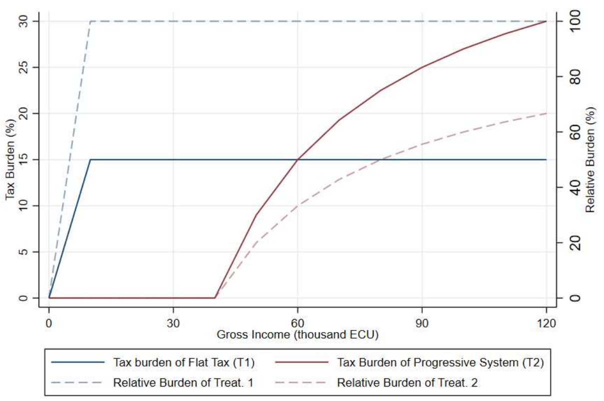
(b) Progressive Tax (Treatment 2)

40,000
10,000

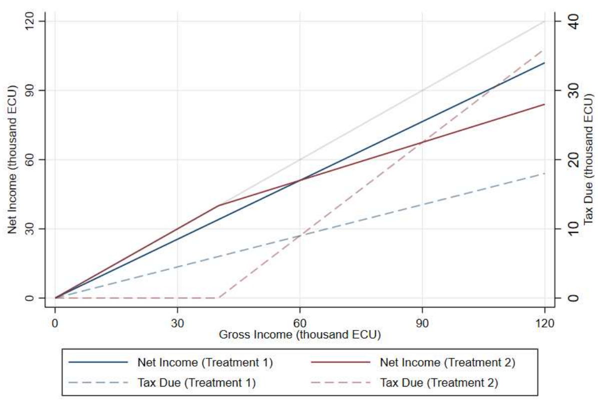
(c) Placebo Captcha (Control)

Notes: ATR stands for Average Tax Rate, applied on all gross income. MTR stands for Marginal Tax Rate, applied only on the fraction of income that falls within each band.

Figure 10: Income taxes, by tax schedule



(a) Distribution of tax burdens



(b) Gross vs Net incomes

Notes: Graph (a) shows the tax burden that individuals earning different levels of income faced depending on their tax schedule, as their average tax rate (solid lines, on the left vertical axis), or relative to the top marginal rate (dashed lines, on the right vertical axis). The solid lines in graph (b) illustrate the differences in tax burden by plotting the net income corresponding to each level of gross income under the two possible tax schedules. The dashed lines show the amount of taxes paid by each level of gross income.

levels). Therefore, we would expect to find significant differences between Control and Treatment 2 (progressive tax schedule) but not between Control and Treatment 1 (proportional tax system). However, if significant differences were identified between Treatment 1 and Control, without significant differences between both treatments, my hypothesis about the *tax burden heuristic* would be rejected while confirming a role of taxes on perceived income distributions through a mechanism to be further explored. On the other hand, if no statistical differences were identified between the Control group and either of the Treatment groups, that would indicate that income taxes are unlikely to be the main driver of the large bias observed at the bottom of the income distribution in the GSS data.

It is important to notice that if participants suffered from representativeness bias[13], the starting bias of those above and below the mean would have the opposite sign, potentially affecting any posterior updating of beliefs. Therefore, I created a dummy (H_i) to differentiate effects on high-income levels with respect to low-income levels. I defined as *top half* incomes from $60,000$ ECU ($H = 1$ if $y_i \geq 60$).

4.3 Empirical analysis

The sample of participants in the experiment consisted of $1,535$ people, distributed evenly between the three groups, as seen in Table 2. The percentage in parenthesis is the proportion of participants in each cell being matched to the high-performance group. To reduce noise, I exclude the tails of the distribution, thus focusing on the income range $[10, 100]$. This reduces the sample to $1,368$ valid observations.

For the first measures of perceived statistics of the distribution (X_{i1}), participants only knew the set of tasks, their gross earned income (y_i) and, treated individuals, the tax system (T_1 or T_2). In the regression analysis, income effects are allowed to be non-linear by using a set of dummies (C), and the control group is used as the base category, adding dummies for each treatment group. I also add the dummy (H_i) for the top half of the income range included in the analysis ($y_i \geq 60$) interacted with the treatment dummies. For individuals in Treatment 2 (progressive tax schedule), the dummy for the top half of the income distribution also captures the effect of the relative tax burden (higher

[13]The consequence of the representativeness bias in this context would be that agents believe they are more representative of the population than they are, hence underestimating their distance from the average income.

Table 2: Distribution of participants across groups, by income level

	Income (thousand ECU)													
	0	**10**	**20**	**30**	**40**	**50**	**60**	**70**	**80**	**90**	**100**	**110**	**120**	**Total**
Treatment 1	33	40	63	62	71	63	53	52	33	17	14	4	3	508
	(51.5%)	(47.5%)	(57.1%)	(51.6%)	(50.7%)	(46%)	(60.4%)	(51.9%)	(51.5%)	(35.3%)	(57.1%)	(50%)	(66.7%)	(51.8%)
Treatment 2	59	58	36	61	66	50	58	45	55	22	12	5	1	528
	(42.4%)	(50%)	(44.4%)	(54.1%)	(47%)	(46%)	(50%)	(55.6%)	(58.2%)	(63.6%)	(50%)	(0%)	(100%)	(50%)
Control	51	49	39	47	66	60	72	46	31	16	11	10	1	499
	(56.9%)	(51%)	(48.7%)	(46.8%)	(54.5%)	(53.3%)	(52.8%)	(39.1%)	(41.9%)	(56.3%)	(36.4%)	(70%)	(0%)	(50.5%)
Total	143	147	138	170	203	173	183	143	119	55	37	19	5	1535
	(49.7%)	(49.7%)	(51.4%)	(51.2%)	(50.7%)	(48.6%)	(54.1%)	(49%)	(52.1%)	(52.7%)	(48.6%)	(47.4%)	(60%)	(50.7%)

Notes: In each cell, the number on the top is the number of individuals with that income level in a given treatment group. The percentage in parenthesis is the proportion of those who were matched to the high-performing reference group.

on the top half). Notice that the income dummies already absorb any differences between the lower and top half of the income range in the control group. The set of socio-demographic controls (W_i) added to the regression are age, education level, previous experience filing taxes, and the quintile of the USA income distribution they believe to belong to in real life. Moreover, the measures of relative position (self-perceived percentile and perceived probability of being above the mean) are likely to be affected by the perceived average income. Therefore, I run an additional specification for those two regressions, controlling for perceived average income ($\tilde{y}_i$). This specification allows testing whether the tax has a direct impact on the perceived position, as suggested by the *tax burden heuristic*, or all the effect goes through an update on the whole perceived distribution (reflected by the change in perceived mean). Continuous variables are used in natural logarithms so that the coefficients can be interpreted as percentage changes. The estimation equation is:

$$\ln X_{i1} = \theta_0 + \Theta_1' \mathbb{1}\{T_i\} + \Theta_2' \mathbb{1}\{T_i \times H_i\} + [\theta_3 \ln \tilde{y}_i] + \Theta_4' \mathrm{C}_i + \Theta_5' \mathrm{W}_i + \varepsilon_i \tag{11}$$

Table 3 shows the results of the ordinary least squares (OLS) regressions for the first measures of perceived statistics of the income distribution. None of the three measures (Q1, Q2, Q3) shows any statistically significant difference between either of the treated groups and the control group. Nevertheless, the sign of the coefficients is consistent across specifications, and it is the opposite for each treatment (T_1, T_2). Individuals facing the flat tax system seem to perceive themselves at a slightly higher percentile than those in the control group if their income is in the bottom half ($y_i < 60$), and slightly lower if their income is in the top half ($y_i \geq 60$). On the other hand, individuals facing the progressive tax system seem to believe to be at a slightly lower percentile

Table 3: Impact of tax schedules on perceived rank and average income (first measure)

	Q2: Mean Income		Q1: Percentile			Q3: $Pr(y_i > \tilde{y})$		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Flat Tax System (T_1)	-0.070	-0.078	0.014	0.024	0.004	0.062	0.078	0.044
	(0.051)	(0.050)	(0.068)	(0.065)	(0.064)	(0.117)	(0.111)	(0.108)
Progressive Tax System (T_2)	0.026	0.021	-0.081	-0.103	-0.098	-0.020	-0.053	-0.044
	(0.049)	(0.048)	(0.073)	(0.070)	(0.069)	(0.124)	(0.116)	(0.114)
Flat Tax × Top Half ($T_1 \times H_i$)	0.064	0.080	-0.080	-0.097	-0.076	-0.100	-0.136	-0.101
	(0.061)	(0.061)	(0.086)	(0.086)	(0.084)	(0.137)	(0.135)	(0.131)
Prog. Tax × Top Half ($T_2 \times H_i$)	-0.007	-0.005	0.052	0.094	0.093	-0.075	0.011	0.009
	(0.058)	(0.059)	(0.089)	(0.088)	(0.086)	(0.147)	(0.142)	(0.140)
Perceived mean income ($\ln \tilde{y}_i$)					-0.259***			-0.437***
					(0.051)			(0.093)
Observations	1,368	1,363	1,368	1,363	1,363	1,368	1,363	1,363
R-squared	0.297	0.306	0.122	0.171	0.198	0.206	0.275	0.301
Income Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Socio-Dem. Controls	NO	YES	NO	YES	YES	NO	YES	YES

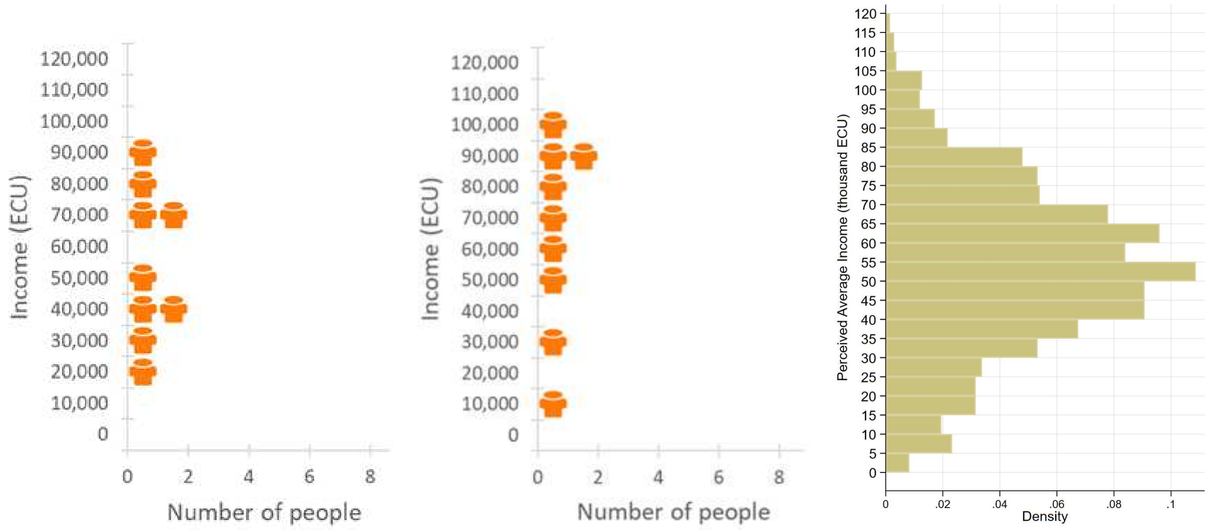
Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.
Notes: Socio-demographic controls include age, gender, employment status, and educational level.
The first two columns report coefficients for the regressions of the perceived average income (Q2). The dependent variable in columns 3 to 5 is the perceived income percentile (Q1). The final three columns correspond to the regressions of the perceived probability of being above the average income (Q3). The coefficients of the income dummies and control variables have been removed from this output table (vectors Θ'_4 and Θ'_5 in Equation 9).

than those in the control group only if their income is below $60,000$ ECU.

Since participants were allowed to modify their answers after seeing the summary of incomes of other members in their group, I ran the same regression with those final *amended* measures, incorporating the new set of information. Given that all three groups (T_1, T_2 and *Control*) saw the partial summary statistics of the income distribution of their reference group at the same point in the experiment, it is impossible to differentiate how much of the change between both measures results merely from realising they had introduced a wrong amount at first, and how much was a result of the disclosed results of other participants. For that reason, the most relevant regression will be one using the final measure only, after participants were allowed to correct their initial answers. Nevertheless, I include the regression on the *correction* (updating) process for robustness.

Participants were shown a summary of the performance of other nine players in their reference group, which could be a low-performing or a high-performing group. The two first graphs from the left in Figure 11 correspond to those pieces of information participants received. The low-performing reference (left) has the income level of nine players, with a median value $50,000$ and mean $54,444$ ECU. The summary of the high-performing group (centre), on the other hand, has a median value $70,000$ and mean $64,444$ ECU. The last graph in the same figure shows the distribution of the reported perceived average income in the first set of questions (before seeing the group information). Answers follow a normal distribution centred around $52,500$ ECU. Fifty per cent of respondents reported a prior belief for the average income below the mean of the lower reference group, while the number of participants with a prior belief below the mean income of the higher reference group was substantially larger, up to seventy per cent.

Figure 11: Reference groups and prior perceived average income



(a) Lower Group ($R_i = 0$) (b) Higher Group ($R_i = 1$) (c) Distribution of perceptions

Notes: The partial distributions in (a) and (b) are the pieces of information about other participants' performance that players were shown after submitting their initial set of answers. Chart (c), on the other hand, summarises the distribution of initial answers about the perceived average income, with mean value at 37.6 thousand ECU.

Participants were given a chance to modify their initial answers after seeing this new piece of information. If participants were to update their beliefs merely based on the new piece of information, and given the distribution of initial answers (with an average estimated mean income below $40,000$ ECU), we would expect a statistically significant positive change in perceived mean (Q2) for respondents matched to either group, though larger for those with the higher reference group ($R_i = 1$), with no differences between control and treated groups. The sign of the coefficients should be the exact opposite for the measures of perceived position (Q1 and Q3).

Nevertheless, I already mentioned that giving participants the opportunity to correct their answers may cause changes that respond to other factors than the new piece of information. Most of them are unobservable, but I can control for the information they saw previously (treatment group).

Therefore, I regress the changes to the reported measures, denoted by $\triangle X_i = \ln(\frac{X_{i2}}{X_{i1}})$. As defined in the estimation equation 10, I control for mean reversion by adding the initial measure X_{i1} in natural logarithms, add treatment dummies for changes uncorrelated with the reference group, and interact the treatment and reference group dummies with the dummy for the top half values of the income range (H_i) to allow for differential effects on high and low incomes. To identify the direct impact of tax systems on perceived rank, I control for changes in perceived average income too, $\triangle\tilde{y}_i = \ln(\frac{\tilde{y}_{i2}}{\tilde{y}_{i1}})$. As in Equation 3, C_i is the set of income dummies, and W_i the set of controls. Results of this set of OLS regressions are shown in Table 4.

$$\triangle X_i = \gamma_0 + \gamma_1 \ln X_{i1} + \Gamma_2' \mathbb{1}\{T_i\} + \Gamma_3' \mathbb{1}\{T_i \times H_i\} \\ + \gamma_4 R_i + \gamma_5 \{R_i \times H_i\} + [\gamma_6 \triangle \tilde{y}_i] + \Gamma_7' \mathrm{C}_i + \Gamma_8' \mathrm{W}_i + \varepsilon_i \quad (12)$$

Table 4: Impact of reference groups on perceived rank and average income (change)

	Q2: $\triangle$ Mean Income		Q1: $\triangle$Percentile			Q3: $\triangle Pr(y_i > \tilde{y})$		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prior Belief	-0.258***	-0.266***	-0.207***	-0.217***	-0.202***	-0.197***	-0.219***	-0.204***
	(0.045)	(0.045)	(0.024)	(0.024)	(0.024)	(0.026)	(0.027)	(0.027)
Flat Tax System (T_1)	0.096***	0.093***	-0.052	-0.052	-0.016	-0.169**	-0.164**	-0.100
	(0.026)	(0.026)	(0.042)	(0.042)	(0.040)	(0.069)	(0.068)	(0.071)
Progressive Tax System (T_2)	0.103***	0.102***	-0.012	-0.010	0.022	-0.243***	-0.246***	-0.190**
	(0.027)	(0.026)	(0.044)	(0.043)	(0.042)	(0.078)	(0.078)	(0.079)
Flat Tax $\times$ Top Half ($T_1 \times H_i$)	-0.107***	-0.108***	0.070	0.072	0.031	0.142	0.139	0.066
	(0.031)	(0.032)	(0.053)	(0.053)	(0.051)	(0.087)	(0.087)	(0.089)
Prog. Tax $\times$ Top Half ($T_2 \times H_i$)	-0.136***	-0.141***	0.045	0.048	0.001	0.157	0.171*	0.091
	(0.031)	(0.032)	(0.054)	(0.053)	(0.052)	(0.101)	(0.102)	(0.103)
Higher Reference Group (R_i)	0.055**	0.054**	-0.111***	-0.117***	-0.100***	-0.096	-0.104	-0.074
	(0.022)	(0.022)	(0.037)	(0.037)	(0.036)	(0.064)	(0.063)	(0.061)
Higher Ref. Group $\times$ Top Half ($R_i \times H_i$)	0.103***	0.098***	-0.049*	-0.040	-0.007	-0.185***	-0.179***	-0.118**
	(0.014)	(0.014)	(0.028)	(0.028)	(0.029)	(0.049)	(0.050)	(0.050)
Change of perceived mean income ($\triangle\tilde{y}_i$)					-0.322***			-0.574***
					(0.068)			(0.150)
Observations	1,368	1,363	1,368	1,363	1,363	1,368	1,363	1,363
R-squared	0.259	0.274	0.156	0.164	0.201	0.126	0.142	0.181
Income Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Socio-Dem. Controls	NO	YES	NO	YES	YES	NO	YES	YES

Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: The regressions in this table have the change in the reported statistics as dependent variables. That is, the log-difference between the initial value reported and the final value submitted when individuals were offered the opportunity to amend their answers after seeing the scores of a subset of other 9 participants in their group. The first two columns report coefficients for the regressions on the change in perceived average income (Q2). The dependent variable in columns 3 to 5 is the change in perceived income percentile (Q1). The final three columns correspond to the regressions of the change in perceived probability of being above the average income (Q3). Socio-demographic controls include age, gender, employment status, and educational level.

Individuals in the lower half of incomes of both treated groups increased their perceived average income (columns 1 and 2) by around 10% more than control individuals. Among individuals in the higher half of incomes, those facing the progressive system (T_2) reduced their perceived mean by a statistically significant 3% compared to the control group, while there weren't significant differences between those facing the proportional system (T_1) and the control group. Moreover, individuals seeing the summary for the higher-performing reference group significantly increased their perceived average income (Q2) by another five per cent, and the magnitude of such increase was 10 percentage points larger for those in the top half of the income range (see interaction term $R_i \times H_i$). Seeing summary information from the higher-performing reference group also had a negative and statistically significant impact on the perceived percentile (Q1) and the probability of being above the mean (Q3). Changes in perceived percentile, however, do not significantly differ between control and treated individuals. Significant differences between Control and Treated groups arise in the other measure of position (Q3). Individuals facing the flat tax system (T_1) reduced their perceived probability of being above the mean, although its significance level drops after controlling for the actual update on the perceived average income level. Individuals facing the progressive tax system (T_2), on the other hand, also significantly reduced their perceived probability to be above the mean, and such change remains statistically significant after controlling for the actual change in perceived average income level (column 8). This means that the change in perceived position of those in Treatment 2 went beyond what can be explained through the change in perceived average income.

The last set of answers is likely to report the respondent's beliefs more accurately, and having a richer set of information is likely to reduce variance. Therefore, the last regression is the most relevant to confirming tax schedules' influence on perceived income distributions and perceived position. Equation 11 aims to identify the causal impact of tax systems and reference groups on perceived average income, perceived percentile, and perceived position relative to the mean:

$$X_{i2} = \lambda_0 + \Lambda_1' \mathbb{1}\{T_i\} + \Lambda_2' \mathbb{1}\{T_i \times H_i\} + \lambda_3 \mathbb{1}\{R_i\} + \lambda_4 \mathbb{1}\{R_i \times H_i\} + [\lambda_5 \ln \tilde{y}_i] + \Lambda_6' \mathrm{C}_i + \Lambda_7' \mathrm{W}_i + \epsilon_i \quad (13)$$

Effectively, Table 5 confirms a significant impact of the progressive tax system. First, columns 1 and 2 (Q2) confirm that the progressive tax system had a significant impact on the average income perceived by lower-income individuals (most of which are within the tax-free allowance), while the flat tax system did not generate significant differences in Q2 compared to the control group. Individuals with incomes below $60,000$ ECU in Treatment 2 (progressive tax) perceived an average income level in their reference group 12% larger than those in the Control group and 8% larger than those in the Treatment 1 group (flat tax). Columns 6 to 8 also confirm that the progressive system had a statistically significant impact on the perceived position with respect to the average income, which is the most relevant information according to classical models of preferences for redistribution, where individuals below the average income have incentives to support redistributive policies since they would benefit from them. Compared to control individuals, those facing the progressive tax system believed they were nearly 25% less likely to be above the mean if their income was in the bottom half, although the impact is only half that magnitude on individuals with higher incomes ($y_i \geq 60,000$ ECU), controlling for socio-demographic characteristics. The impact of the progressive tax system on the perceived probability of being above the mean remains above 90% confidence level after controlling for the perceived average income level, which could indicate taxpayers assessed their position relative to the average income by using the *tax burden heuristic*. Columns 3 to 5, however, do not show statistical significance for either treatment effect, and therefore the regression fails to reject that the tax system had no impact in shaping the perceived income rank (percentile) of taxpayers.

Table 5: Impact of tax schedules on perceived rank and average income (final measure)

	Q2: Mean		Q1: Percentile			Q3: $Pr(y_i > \tilde{y})$		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Flat Tax System (T_1)	0.045	0.036	-0.041	-0.033	-0.023	-0.119	-0.103	-0.083
	(0.048)	(0.046)	(0.069)	(0.067)	(0.065)	(0.121)	(0.114)	(0.111)
Progressive Tax System (T_2)	0.123***	0.117***	-0.077	-0.091	-0.058	-0.259**	-0.287**	-0.223*
	(0.046)	(0.044)	(0.075)	(0.071)	(0.071)	(0.128)	(0.119)	(0.117)
Flat Tax × Top Half ($T_1 \times H_i$)	-0.058	-0.048	0.007	-0.005	-0.018	0.064	0.034	0.008
	(0.057)	(0.057)	(0.086)	(0.086)	(0.084)	(0.144)	(0.140)	(0.136)
Prog. Tax × Top Half ($T_2 \times H_i$)	-0.139**	-0.143***	0.087	0.121	0.081	0.100	0.181	0.104
	(0.055)	(0.054)	(0.090)	(0.088)	(0.088)	(0.153)	(0.149)	(0.146)
Higher Reference Group (R_i)	0.058	0.058	-0.122**	-0.123**	-0.107*	-0.080	-0.088	-0.057
	(0.037)	(0.036)	(0.059)	(0.057)	(0.056)	(0.101)	(0.095)	(0.094)
Top Half × Higher Ref. Group ($H_i \times R_i$)	0.085***	0.080***	-0.058	-0.035	-0.012	-0.223***	-0.195***	-0.152**
	(0.024)	(0.025)	(0.042)	(0.044)	(0.044)	(0.066)	(0.071)	(0.070)
Perceived mean income ($\ln \tilde{y}_i$)					-0.282***			-0.544***
					(0.060)			(0.096)
Observations	1,368	1,363	1,368	1,363	1,363	1,368	1,363	1,363
R-squared	0.239	0.251	0.208	0.250	0.274	0.271	0.335	0.362
Income Dummies	YES	YES	YES	YES	YES	YES	YES	YES
Socio-Dem. Controls	NO	YES	NO	YES	YES	NO	YES	YES

Robust standard errors in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

Notes: The first two columns report coefficients for the regressions on the perceived average income (Q2). The dependent variable in columns 3 to 5 is the respondents' perceived income percentile (Q1). The final three columns correspond to the regressions of the perceived probability of being above the average income (Q3). Socio-demographic controls include age, gender, employment status, and educational level. The coefficients of the control variables and the income dummies are not included in this table.

5 Conclusions

This research project aims to fill a relevant existing gap in the literature of redistributive preferences. Its initial results, despite its limitations, seem to demonstrate that tax systems play a role in shaping perceived income distributions and, as a result, may influence support for redistributive policies by biasing taxpayer perceptions.

The initial exploratory analysis using the longitudinal survey GSS in the United States and exploiting quasi-experimental variation from tax law changes shows a systematic inflexion in perceived position for households with incomes below the tax-free allowance. This paper presents a rational framework to explain why taxpayers may infer signals from tax systems using the most salient information: their own tax burden and the maximum possible tax rate in the system. A similar reasoning could be used to develop a model with different reference points. For example, one where taxpayers at the bottom of the distribution used the tax-free allowance as a reference, while taxpayers at the top of the distribution used the top threshold as a reference.

Despite requiring large sample sizes to ensure enough statistical power to identify small effects, online experiments provide a synthetic environment where the different key elements of tax schedules can be tested one at a time. My experiment, using a large number of participants to compare two very different tax systems, revealed statistically significant differences in perceived income distributions of individuals facing a proportional flat tax system compared to those facing a progressive tax system. On average, individuals facing the progressive tax system believe they are between 10% and 25% less likely to be above the average income than those in the control and flat tax system groups. Moreover, the effect remains significant at the 10% confidence level and with a fairly stable value even after controlling for differences in the perceived average income value. At the same time, the perceived average income is significantly higher (+12%) for low-income individuals facing the progressive tax system than for those in the control and flat tax groups. These findings seem to indicate that tax systems influence perceptions of relative position as well as overall distributions of income. On the other hand, however, differences in perceived percentile were not significant, hence challenging the robustness of the findings.

Despite its limitations, the conclusions of this paper are of relevance to the literature in redistributive preferences, contributing to solving the *inequality-redistribution puzzle*. My results suggest that taxpayers use tax information to create their beliefs on income distributions and their position within those. Further research is needed to gain a better understanding of the mechanism behind the impact of progressive tax systems on perceived income distributions, as well as the role of income tax credits and different tax rebates, which were not considered in this study. If posterior studies confirmed the existence of the proposed tax burden heuristic, it would highlight the risks of misinformed voters. If tax systems themselves contribute to biasing perceptions of citizens, it is of utmost importance to provide accurate information to taxpayers on their true position in the distribution so they can correctly assess what policies would be of most benefit to them. Otherwise, by reducing the highest marginal tax rates as it happened during the second half of the 20th century, taxpayers at the lower end of the income distribution saw their tax burden get closer to the maximum tax rate, potentially inducing a biased overestimation of their position relative to the average family income. This, at the same time, could translate into lower support for redistributive policies, hence increasing support for the tax reduction on high incomes that caused the bias in the first place, and leading to a continued reduction of progressivity in income tax schedules.

References

A. Alesina and P. Giuliano. Preferences for redistribution. In *Handbook of social economics*, volume 1, pages 93–131. Elsevier, 2011.

R. Benabou. Inequality and growth. *NBER macroeconomics annual*, 11:11–74, 1996.

R. Bénabou and E. A. Ok. Social mobility and the demand for redistribution: the poum hypothesis. *NBER Working Paper*, (w6795), 1998.

C. J. Boyce, G. D. Brown, and S. C. Moore. Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological science*, 21(4):471–475, 2010.

D. L. Chen, M. Schonger, and C. Wickens. {oTree} — {An} open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016.

G. Cruces, R. Perez-Truglia, and M. Tetaz. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *Journal of Public Economics*, 98: 100–112, 2013.

M. D. Evans and J. Kelley. Subjective social location: Data from 21 nations. *International Journal of Public Opinion Research*, 16(1):3–38, 2004.

J. Fernández-Albertos and A. Kuo. Income perception, information, and progressive taxation: Evidence from a survey experiment. *Political Science Research and Methods*, 6(1):83–110, 2018.

V. Gimpelson and D. Treisman. Misperceiving inequality. *Economics & Politics*, 30(1):27–54, 2018.

O. P. Hauser and M. I. Norton. (mis) perceptions of inequality. *Current opinion in psychology*, 18: 21–25, 2017.

K. B. Hvidberg, C. T. Kreiner, and S. Stantcheva. Social position and fairness views. *NBER Working Paper*, (w28099), 2020.

D. Kahneman and A. Tversky. Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3):430–454, 1972.

M. Karadja, J. Mollerstrom, and D. Seim. Richer (and holier) than thou? the effect of relative income improvements on demand for redistribution. *Review of Economics and Statistics*, 99(2): 201–212, 2017.

L. Kenworthy and L. McCall. Inequality, public opinion and redistribution. *Socio-Economic Review*, 6(1):35–68, 2007.

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

A. H. Meltzer and S. F. Richard. A rational theory of the size of government. *Journal of political Economy*, 89(5):914–927, 1981.

T. Piketty. Social mobility and redistributive politics. *The Quarterly journal of economics*, 110(3): 551–584, 1995.

Appendix

A Experiment Design

This appendix provides further details of the experiment and its reward structure, as well as some sample screenshots.

Table 6: Randomisation checks

	Treatment 1	Treatment 2	Control
Males	52%	52%	49%
Age, years	36.4	36.74	37.36
	(12.14)	(11.76)	(12.25)
Higher Education	72%	72%	75%
Employed	77%	77%	80%
Ever filed taxes	89%	90%	92%
Earnings in real life	$ 67,512	$ 51,627	$ 52,575
	(372021)	(58204)	(54033)
Taxes in real life	$ 6,703	$ 6,288	$ 6,633
	(14205)	(12450)	(10159)
Quintile in real life	3.6	3.55	3.55
	(1.06)	(1.09)	(1.09)
Focus measure	0.89	0.90	0.89
	(0.16)	(0.16)	(0.16)
Correct tasks (/12)	4.68	4.86	4.78
	(2.39)	(2.52)	(2.36)
Matched to High Ref. Group	52%	51%	49%

Notes: This table provides summary statistics of socio-demographic and performance variables to check whether the random allocation of participants to treatment groups yielded balanced samples. Only the subsample used in the analysis was used for the calculations: participants with a focus measure above 25% and solving correctly between one and ten real effort tasks. Higher education refers to university level (Bachelor or above). For continuous variables, the standard deviation is reported in parenthesis.

A.1 Earning scheme and incentives

Participants earned money from three different sources:

1. **Participation:** all participants that completed the experiment were awarded a participation fee of 0.40 USD.

2. **Net earned income**: amount earned in the initial task ($10,000$ ECU for each correct answer), minus taxes. The equivalent to real currency is $50,000$ ECU to 1 USD.

Participants in the two treatment groups were told the tax scheme they were facing when their earned income was revealed, while participants on the control group were told their earnings would be subject to a tax, but the system in place was only revealed at the end of the experiment.

3. **Bonus income**: additional tax-free amount ($25,000$ ECU) earned for correctly answering the questions regarding income distribution statistics.
   Since all three questions are related, the answer to one question could condition another one. To prevent this, a lottery decides which one of the three questions is chosen for the bonus payment.

The random matching to 99 other players was achieved through a random match of the player to one of 2 pre-existing groups of 99 previous records. Matching to 99 pre-loaded previous records is an alternative to a 100 player simultaneous game, which would introduce additional issues (waiting time, potential drop-outs, etc.). Each of those groups was formed by bootstrapping from a sample of pre-existing records from the trial phase. At a point in the experiment (see Figure 8), participants were shown a selected subsample of 9 of those players in their matched group as a source of partial information to update their beliefs.

A.2 Choice of tax schemes

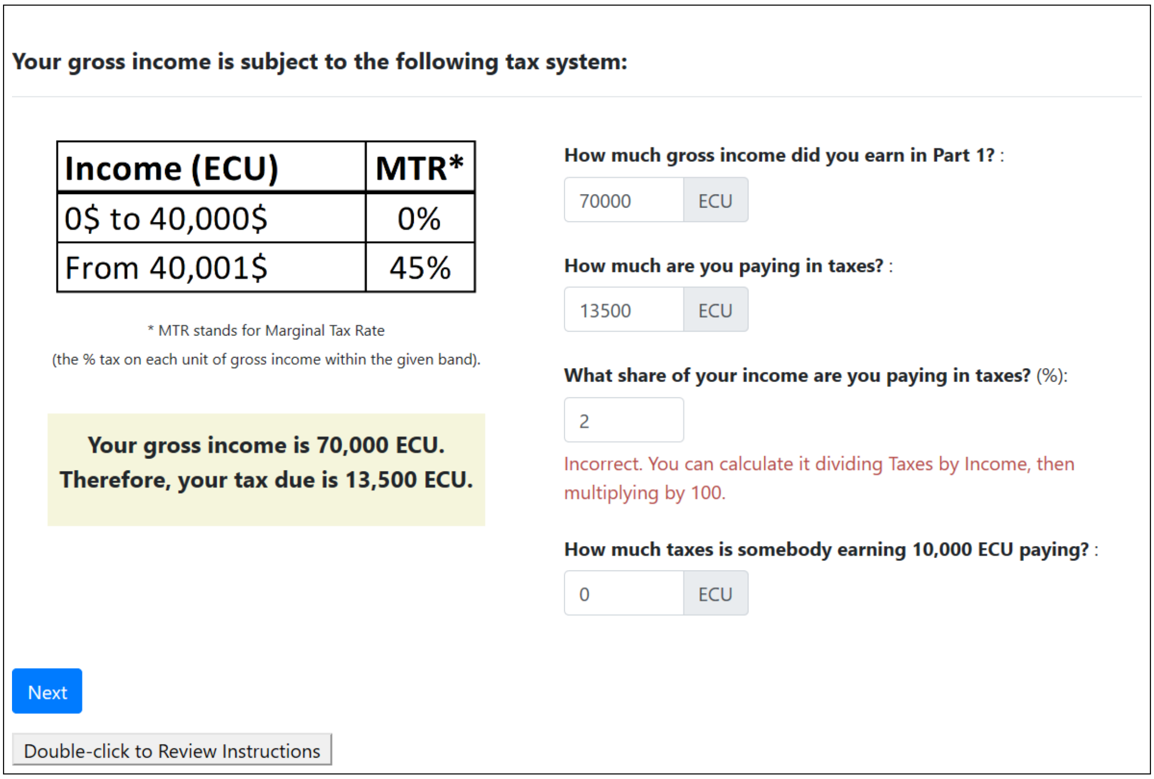
The number of brackets, income thresholds and marginal tax rates for the progressive tax schedule were chosen taking the UK Income Tax as a main reference.[14] To ensure that participants looked carefully at the tax schedule, they were asked a few simple details of the tax system in place.

A.3 Real effort tasks

The task was partially inspired by **?** and is similar in spirit to **?**, remaining simpler than **?**. It consists of solving four mathematical operations and eight scrambled words (four names of countries and four names of animals), and it is preceded by a trial task involving one question of each type so

[14] Details are published on-line by the HMRC: `https://www.gov.uk/government/publications/rates-and-allowances-income-tax/income-tax-rates-and-allowances-current-and-past`.

Figure 12: Income Tax Information for Treatment 2 (progressive system)



that players become acquainted with the dynamics. Each word/operation has a time limit of five seconds to be solved, and every correct answer yields $10,000$ECU.

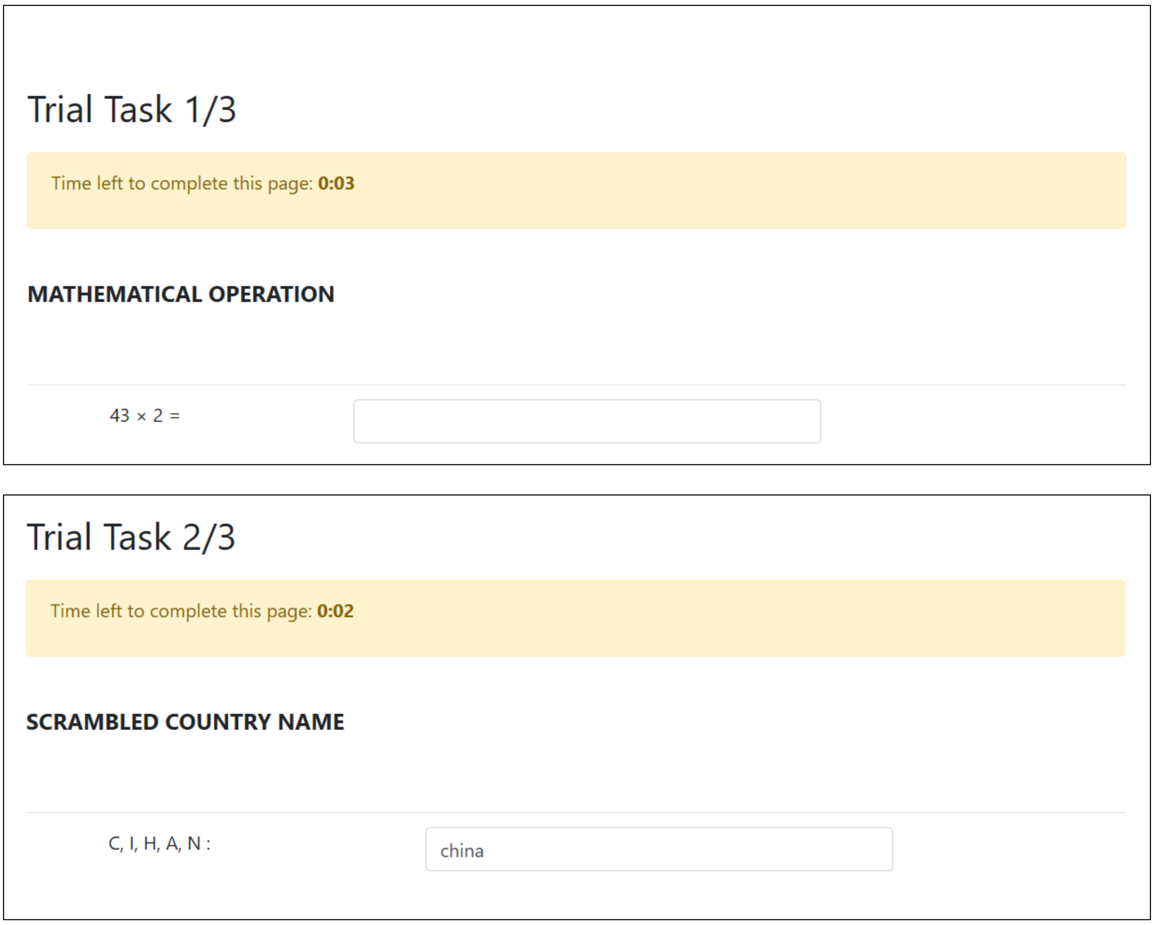
For the purpose of the experiment, the earning task needs to follow three main requirements: incomplete information on the income distribution, realism, and low dispersion. The need for incomplete information and realism is obvious, while low dispersion helps to minimize income effects (making respondents more comparable); it also helps with planning the budget of the experiment.

Incomplete information was ensured since all players knew the reward structure, probability of multiplier, time limitations and the lower and upper income bounds, but they did not know other players' performance. Although an upper bound does usually not exist in reality, it is unavoidable in an experiment.

A.4 Instructions

Since this experiment was not distributed in a lab, the instructions had to be fully displayed on the screen, and participants were not provided with a printed copy. The set of initial instructions was available for review on every screen through a button at the bottom of the page.

Figure 13: Examples of real effort tasks



Screenshots of the instructions at each step are provided below, with self-explanatory headings:

Figure 14: Initial page of the Experiment

Introduction

You have been randomly matched with other 99 previous participants (forming a group of 100 people).

Part 1:

- You need to **solve** 12 mathematical operations and scrambled words within the given time.
- **Every task** solved correctly will earn you **10,000 ECU**.
- At the end of this part, your total **gross income** will be subject to a **Tax System**.
- The Tax System was designed to collect a target amount to run additional sessions of this experiment.

Part 2:

- You will be asked three additional **bonus questions**.
- Only **one of them** (picked at random) will be paid, adding **25,000 ECU**.
- Income earned in Part 2 is not subject to the Tax System (you will be paid the full earned amount).

Other remarks:

- Amounts in the experiment are expressed in **Experimental Currency Units (ECU)**, with value **50,000 ECU = 1 USD**.
- Remember to write down the code given at the end of the experiment in order to claim your earned income on Amazon Mechanical Turk.

You will be shown detailed instructions at every step of the experiment.

Next

Figure 15: Instructions for Part 1 (real effort tasks)

Part 1

You are required to solve some **short mathematical operations and scrambled names** of countries and animals.

- For every task, you will be told if it is a mathematical operation, scrambled animal name or scrambled country name.
- You will have **5 seconds** to answer each task by typing the number or ordered letters in the textbox.
- Your answer will be **automatically submitted after timeout** (5 seconds) and a new task will be shown.
- You can **submit your answer before timeout** by pressing the *Enter* key on your keyboard.
- **Every correct answer has a value of 10,000 ECU.**

You will start with a trial of 3 tasks and at the end you will be shown your trial results.

The **trial section does not award any income** and is for practice only.

After the trial, you will begin your actual 12 tasks.

You may want to use paper and pen, or a calculator.

Next

Double-click to Review Instructions

Figure 16: Instructions for Part 2 (elicitation of beliefs)

Part 2 - Bonus Questions

Remember you were randomly matched to 99 other players at the beginning of the experiment, forming a group of 100.

You will now be asked **3 questions** related to the earnings of that group of 100 people in Part 1.

A random draw will define which question will be paid at the end of the experiment.

If **Questions 1 or 2** are selected by the random draw, you will be paid the bonus if your answer is correct.

If **Question 3** is the one selected, the probability you select as an answer will define whether your payment depends on the statement in Question 3 being true, or on a random lottery*.

This payment mechanism is designed so that answering honestly every question gives you higher probability of earning the bonus.

The bonus amount is **25,000 ECU**.

*Learn more about the payment method for Question 3.

Next

Figure 17: Additional information about the payment structure

*Learn more about the payment method for Question 3.

How is your payment determined if Question 3 is randomly selected for payment?

- The computer will generate a lottery box with 100 balls, with a random proportion of purple balls (from 1 to 100).
  Note: all proportions of purple balls are equally likely (have the same probability of happening).
- Your **reported probability** defines if your bonus will depend on the statement in Question 3 being indeed true, or on a random draw from the lottery box. How?
  - If the box has a higher number of purple balls than your reported probability:
    A ball will be randomly drawn. If it is **purple, you win**.
  - If the box has a lower or equal number of purple balls than your reported probability:
    The statement in Question 3 is checked. If it was indeed **true, you win**.
- Your bonus depends on the lottery only when the probability of wining it is higher than your reported probability of Question 3 being true. Therefore, the method applied will always favour your probability to earn the bonus.

Example:

- Imagine you **reported a probability of 20 percent** that the statement in Question 3 is true.
- The computer randomly generated a box with 100 balls, 40 of which are purple.
- This means the probability to randomly draw a purple ball from the box is **40 percent**.
  Since 40 (purple balls) is higher than 20 (reported probability):
  - **Your payment depends on the box**, and it does no longer matter whether the statement in Q3 is true or false.
  - A ball will be randomly drawn from that box: **if it is purple, you win** the bonus. Otherwise, you don't.

- Imagine instead that you had **reported a probability of 60 percent** that the statement in Question 3 is true.
- This time, the computer randomly generated a box with 100 balls, 50 of which are purple.
  Since 60 (reported probability) is higher than 50 (purple balls):
  - **Your payment depends on the statement** in Question 3, and the lottery box no longer matters in this case.
  - The statement in Question 3 is checked: **if it is true, you win** the bonus. Otherwise, you don't.

Next