

**Online Appendix for *Who Pays for the Welfare State? Preferences for  
Redistribution and Perceptions of the Costs Thereof***

## Contents

<b>1</b>	<b>Data</b>	<b>1</b>
1.1	Data collection and ethical issues . . . . .	1
1.2	Survey design and randomization . . . . .	2
1.3	Descriptive statistics . . . . .	2
<b>2</b>	<b>Methods</b>	<b>4</b>
<b>3</b>	<b>Difference in means tests between experimental groups</b>	<b>5</b>
3.1	First experiment . . . . .	5
3.2	Second experiment . . . . .	5
<b>4</b>	<b>National income distribution in Brazil</b>	<b>6</b>
<b>5</b>	<b>Income level estimation of the top 10%</b>	<b>6</b>
<b>6</b>	<b>Regression results</b>	<b>7</b>

## 1 Data

### 1.1 Data collection and ethical issues

The online survey was conducted by the market research institution Opinion Box. Before answering the questionnaire, participants were informed of possible rewards. The prizes ranged among credits for cell phone calls, donations to institutions, and points exchangeable for products and services in specific commercial establishments. The compensation calculation took into account several factors, such as the questionnaire length and the interviewed audience.

All respondents agreed with the informed and voluntary consent where they were apprised of the research study in which they would take part and the name and e-mail of the responsible researchers. This consent was obtained through an online form right before the survey started. Participants were also assured that they could quit at any moment and that all data collected would be kept anonymous. In case of any doubt or complaint, they were encouraged to contact the Research Ethics Committee of the researchers' home institution.

The survey asked respondents to what extent they agreed or disagreed with specific statements regarding eventual public policies. So, there was no intervention in the political processes, and the questionnaire did not employ deception or any form that could harm participants and other parties affected by the research.

Opinion Box company works in line with national laws and regulations governing their activities. The data were thus collected in accordance with the "General Data Protection Law" and with the standards of conduct of ESOMAR (a not-for-profit organization that provides ethical and professional guidance to market, opinion and social research and data analytics) and ABEP (Brazilian Research Companies Association).

## **1.2 Survey design and randomization**

The non-probabilistic survey divided respondents into similar subgroups in terms of their observable characteristics. This means that the variables income, age, sex, race and education are balanced between these subgroups. In order to demonstrate this balance, the mean tests provided in this appendix (Welch's T-test for two samples) confirm that there is no significant difference between the subgroups, either in the division made by the first experiment or in that made by the second.

The randomized treatment assignment in the first experiment was independent of the treatment assignment of the second experiment. That is, the chance of being assigned to a specific experimental group in the first experiment does not affect the group assigned in the second. Although all participants answered the questions in the same order, the two experiments were not positioned consecutively in the questionnaire.

## **1.3 Descriptive statistics**

The survey was designed based on sample quotas to proportionally represent all income groups, age, and regions of Brazil. It has 2209 respondents. The margin of error is 2.1 percentage points and there are no sample weights. As it is an online survey, there is some degree of sample selection bias: the level of education of those interviewed was, on average, above national parameters. This may compromise the external validity of the results.

Moreover, the experiment was fielded in February and March 2021, a particular period in terms of the income variable in Brazil. On the one hand, the impact of the coronavirus outbreak caused the biggest drop in the level of job occupation of people aged 14 or over ever recorded by the Brazilian Institute of Geography and Statistics (IBGE). According to national household survey data (PNAD Contínua and PNAD-Covid), this rate was, for the first time, below 50% in mid-2020, and remained at this level until March 2021. On the other hand, the standard of living of Brazilians during 2020 was also affected by emergency governmental aid, which directly benefited almost a third of the population. Although it is possible to argue that the application of the questionnaires in this context imposes some caveats to the data analysis, there is no reason to suppose that this compromises the external validity of the results found.

The following tables exhibit basic characteristics of the respondents.

household income	%
Up to BRL 1,100	15,05
From BRL 1,101 to BRL 2,200	24,87
From BRL 2,201 to BRL 3,300	17,91
From BRL 3,301 to BRL 5,500	18,75
From BRL 5,501 to BRL 11,000	16,22
From BRL 11,001 to BRL 16,500	4,68
More than BRL 16,500	2,52

sex (%)	
male	female
48,44	51,56

age	%
18 to 24 years old	26,62
25 to 29 years old	13,99
30 to 39 years old	25,76
40 to 49 years old	17,97
50 to 59 years old	10,28
60 years old or more	5,39

race	%
black	11,77
brown	38,39
white	47,89
indigenous	0,81
yellow	0,91
other	0,23

education level	%
Illiterate / Elementary school incomplete	1,22
Elementary school complete / Primary school incomplete	2,72
Primary school complete	2,08
Incomplete high school	4,84
Complete high school	33,18
Incomplete tertiary	20,33
Complete tertiary	23,31
Postgraduate studies	12,31

## 2 Methods

The article uses household earnings *per capita* as a proxy for individual income. Among the income measures collected in opinion polls, this is the one that best reflects the respondents' standard of living. In the experiments, the treatments with provision of information describe individual income values as tax exemption criteria. The two measures (*per capita* and individual income) are distinct, but do not compromise the analysis of the relationship between income and treatment effect.

The survey questionnaire categorizes respondents into income brackets. The household income variable was obtained from the midpoint of each interval, with the exception of the upper category, which received the value determined by a robust estimator adapted to the Pareto distribution (Von Hippel, Scarpino and Holas, 2016). Dividing by the number of household members, the final amount was the *per capita* income level.

In figures 4 and 5, the binary variable that stands for the attitudes toward the policies of each statement assumes a value of 0 if the respondent totally or partially disagrees with the statement, and a value of 1 if he or she totally agrees, agrees in part or does not agree or disagree. This criterion was adopted so that this variable could better capture the variation in policy rejection.

### Controls (used in figure 5)

#### Sex

0	female
1	male

#### Age

1	18 to 24 years old
2	25 to 29 years old
3	30 to 39 years old
4	40 to 49 years old
5	50 to 59 years old
6	60 years old or more

#### Education level

1	elementary school
2	primary school
3	high school
4	tertiary

#### Race

0	white or yellow
1	black, brown or indigenous

### 3 Difference in means tests between experimental groups

High p-values indicate that there is no significant difference between groups.

#### 3.1 First experiment

variable	control	treatment	difference	p-value
per capita income	1645.09	1508.32	136.77	0.11
% male	0.48	0.49	-0.01	0.80
% 30 years old or more	0.32	0.35	-0.03	0.17
% black, brown or indigenous	0.51	0.51	0.00	0.86
% up to complete high school	0.45	0.43	0.03	0.23

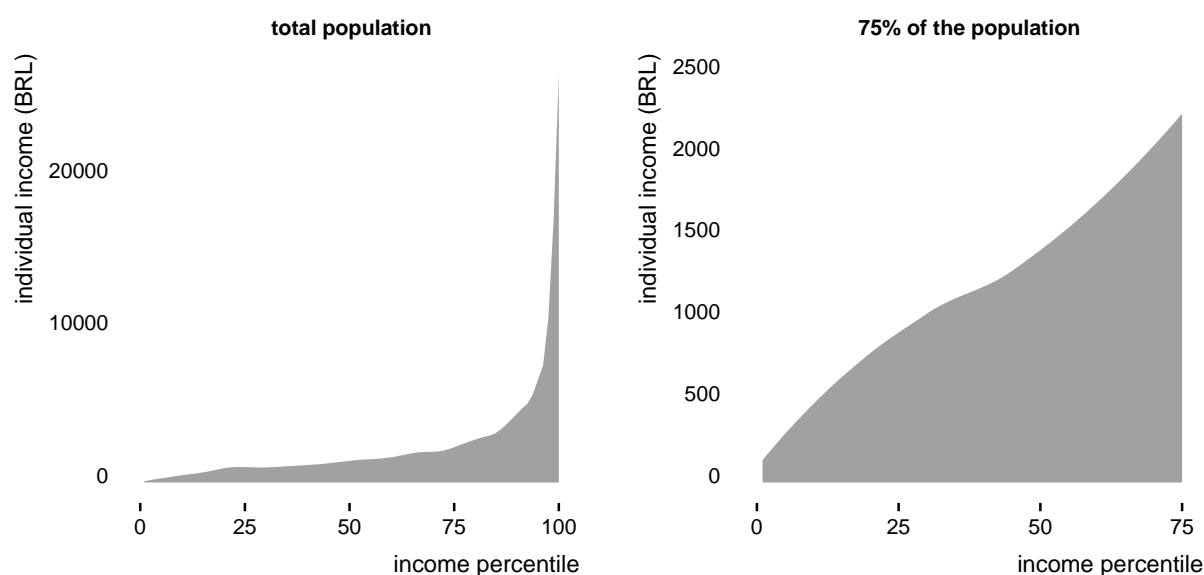
#### 3.2 Second experiment

variable	control mean	treat1 mean	diff	p-value
per capita income	1655.58	1541.64	113.94	0.29
% male	0.48	0.47	0.01	0.80
% 30 years old or more	0.34	0.34	0.01	0.79
% black, brown or indigenous	0.50	0.51	-0.02	0.55
% up to complete high school	0.46	0.43	0.03	0.26

variable	control mean	treat2 mean	diff	p-value
per capita income	1655.58	1533.49	122.09	0.28
% male	0.48	0.51	-0.03	0.28
% 30 years old or more	0.34	0.33	0.01	0.69
% black, brown or indigenous	0.50	0.52	-0.02	0.34
% up to complete high school	0.46	0.44	0.02	0.47

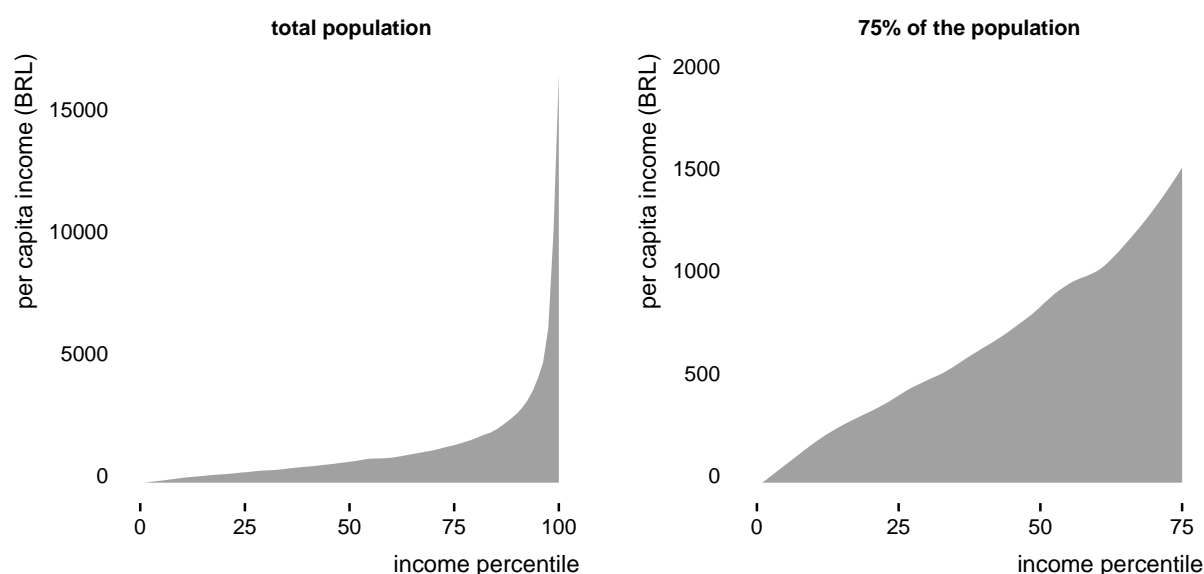
## 4 National income distribution in Brazil

Figure A - 1: Labour income level by percentile of the national distribution



Note: data from national household survey data (PNAD Contínua 2019). The income variable stands for the monthly individual labour income.

Figure A - 2: Household per capita income level by percentile of the national distribution



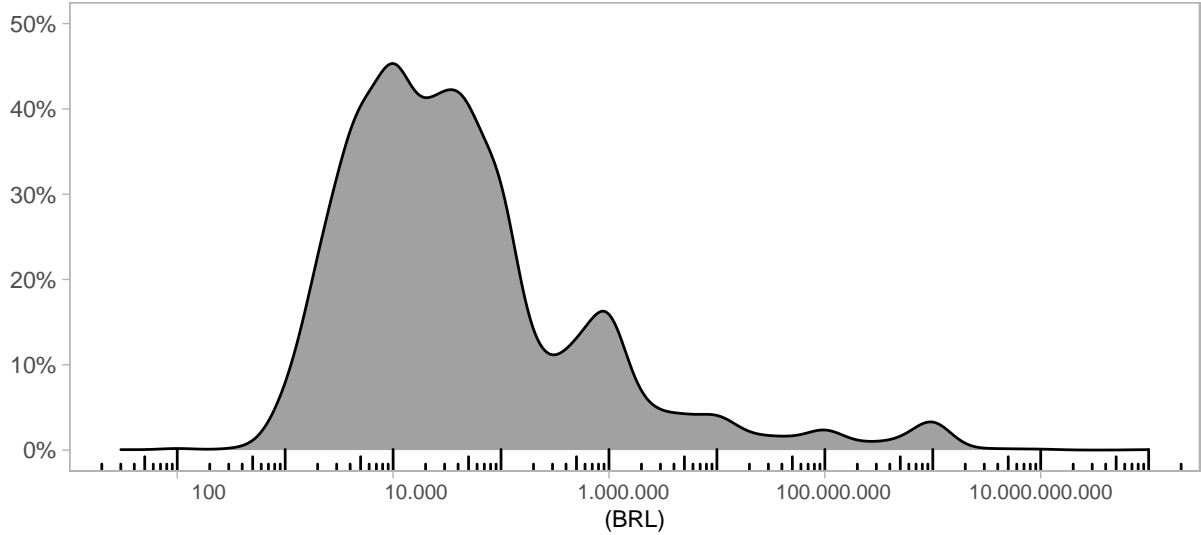
Note: data from national household survey data (PNAD Contínua 2019). The income variable stands for the monthly household *per capita* income from all sources.

## 5 Income level estimation of the top 10%

The Oxfam Brasil/Datafolha nationally representative survey asks respondents “In your opinion, how much do you think a person needs to earn per month to be part of the richest 10% in Brazil?”. The following

figure plots the distribution of responses from the surveys fielded in 2017 (N=2,025), 2019 (N=2,086), and 2020 (N=2,079). As a benchmark, according to administrative records, the actual value of the 90th percentile of the national distribution of individual labour income is close to BRL 4,000.

Figure A - 3: Income population assigns to the rich in Brazil



Note: data from Oxfam Brasil/Datafolha (2017; 2019; 2020).

## 6 Regression results

Table A - 1: Support for tax on the rich

	Binary Logit		Ordered Logit	
	(1)	(2)	(3)	(4)
income percentile	-0.004 (0.003)	-0.005 (0.003)	-0.001 (0.002)	-0.002 (0.003)
sex		0.015 (0.166)		0.091 (0.131)
age		-0.070 (0.056)		-0.027 (0.044)
race		-0.058 (0.171)		-0.052 (0.134)
education level		0.088 (0.114)		0.117 (0.091)
N	2209	2209	2209	2209

Regressions from the second experiment.

Note: data from Opinion Box (2021). Pooled results from five imputed datasets.

+  $p < 0.1$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table A - 2: Support for tax on BRL 2,400+

	Binary Logit		Ordered Logit	
	(1)	(2)	(3)	(4)
income percentile	-0.015 (0.003)***	-0.012 (0.003)***	-0.012 (0.002)***	-0.009 (0.003)**
sex		0.264 (0.155)+		0.197 (0.135)
age		-0.035 (0.054)		-0.053 (0.048)
race		0.427 (0.158)**		0.363 (0.140)**
education level		-0.090 (0.110)		-0.040 (0.096)
N	2209	2209	2209	2209

Regressions from the second experiment.

Note: data from Opinion Box (2021). Pooled results from five imputed datasets.

+  $p < 0.1$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table A - 3: Support for tax on BRL 10,000+

	Binary Logit		Ordered Logit	
	(1)	(2)	(3)	(4)
income percentile	-0.007 (0.003)*	-0.007 (0.003)*	-0.005 (0.002)*	-0.005 (0.003)+
sex		0.121 (0.161)		0.125 (0.134)
age		-0.068 (0.056)		-0.034 (0.047)
race		0.124 (0.168)		0.039 (0.139)
education level		0.068 (0.117)		0.001 (0.099)
N	2209	2209	2209	2209

Regressions from the second experiment.

Note: data from Opinion Box (2021). Pooled results from five imputed datasets.

+  $p < 0.1$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## References

Opinion Box. (2021). Pesquisa Mais Representatividade (dataset). Belo Horizonte.

Oxfam Brasil/Datafolha. (2017). Nós e as desigualdades: percepções sobre desigualdades no Brasil (dataset). São Paulo.

Oxfam Brasil/Datafolha. (2019). Nós e as desigualdades: percepções sobre desigualdades no Brasil (dataset). São Paulo.

Oxfam Brasil/Datafolha. (2020). Nós e as desigualdades: percepções sobre desigualdades no Brasil (dataset). São Paulo.

Von Hippel, P. T., Scarpino, S. V. Holas, I. (2016). Robust estimation of inequality from binned incomes. *Sociological Methodology* 46 (1): 212-251.