When Is It Fair to Tax the Rich? The Importance of Pro-Social Behavior*

Online Appendix

Kris-Stella Trump

Contents

Appendix A1: Detailed description of the coding scheme in Study 1 **2**

Appendix A2: Description of pilot survey experiment **13**

Research design 13

Results 16

Appendix A3: Details for Studies 2 and 3 **22**

Appendix A4: Additional pre-registered analysis of the 'hard work' variable **26**

Appendix A5: Equivalence tests **28**

Appendix A6: Heterogeneous effects by respondent income **30**

*Author affiliation: Johns Hopkins University. Contact: kstrump@jhu.edu

Appendix A7: Comparison of marginal means across US and Danish results **32**

Appendix A8: Differences in marginal means by party **33**

References **34**

Appendix A1: Detailed description of the coding scheme in Study 1

The study explored what Americans think of when they imagine the (un)deserving rich. The respondents were asked one of two sets of open-ended questions. Half of the respondents were asked two questions about rich people (presented in randomized order; randomized element in square brackets):

- "Some people make a lot more money than others. Please imagine a person who is rich and has a good income. Now, imagine that this person [clearly has more than what they deserve]/[deserves what they have, fair and square]. Please describe what comes to mind when you think about this person."

The remaining half were asked a broader question about what considerations come to mind when they think about asking the rich to "pay their fair share":

- "Sometimes people talk about asking the rich to"pay their fair share" in taxes. In talking to people, we've found that people think about a range of different things when they consider this. What comes to mind when you think about whether the rich "pay their fair share" in taxes?"

In addition, respondents were also asked standard survey questions regarding opportunities for upward mobility ("How much opportunity is there in America today for the average person to get ahead?" and "Some people say that people get ahead by their own hard work; others say that lucky breaks or help from other people are more important. Which do you think is most important?"). They answered one question about the level of household income at which they would consider a household rich. They also filled in the racial resentment scale, were asked to indicate how much they support various redistributive

policies ("The government should do more to redistribute from the rich to the poor, even if it means raising taxes."; "High incomes should be taxed more than is currently the case."; "We should decrease government spending on social welfare programs."), and were asked a standard battery of demographic variables including partisanship and household income.

The open-ended responses were organized by major recurring themes in two stages of coding. In the first stage, the author read the open-ended responses and identified recurring themes. The themes were developed in two separate sets: one set of themes to be applied to the "deserving rich" and "undeserving rich" responses, and a separate set of themes for the "pay one's fair share in taxes" responses. Each response could be coded as including several themes. Tables A1 and A2 show the resulting themes and the language that counted toward that theme (the one exception to the procedure described above is the 'system' variable which was added at the second stage of analysis; see below).

Some themes were coded separately based on whether respondents expected that the target *possessed* the trait in question, or whether they mentioned the target did *not* possess the trait in question. This was a particularly useful approach to the "hard work" category: while the undeserving rich prompt received several responses saying that an undeserving rich person presumably "did not work hard", it also received responses arguing that even an undeserving rich person must have "worked hard" for what they had. Coding these two types of mentions separately therefore gives a more accurate view of respondents' reactions than a unified "hard work" category would. Additionally, coding separately for the presence and absence of a trait helps clarify whether traits are most prominent in their presence or absence. For example, expectations that an undeserving rich person treated their workers unfairly were more common than expectations that a deserving rich person treated their workers fairly.

The variable for "giving back" is also worth highlighting, as this category was coded only for whether the respondent mentioned some form of giving back or charity. This choice

was made because nearly all mentions of giving back were phrased conditionally (for the deserving and undeserving prompt alike). Conditional phrasing included "they should still have to share with people who have less", "they should still donate money since they have an excess that they will not use", or "they should want to help others". Overall, responses that mentioned charitable giving in a conditional way tended to imply that a charitable rich person would be more deserving; this occurred to a similar extent for both prompts.

Table A1: Coding scheme for open-ended reactions to (un)deserving rich people.

Variable	Content	Includes mentions of
hardwork	Worked hard	Worked their way up / has a job / Successful / grew business from ground up / sense of purpose and direction / had good invention / hard earned money
nothardwork	Did not work hard	Lazy
giveback	Should give back to society	Give to charity / Donate / Contribute to society / Help others / Invented something that helps others / supports the community
inherit	Inherited wealth	Rich parents / Trust fund baby / Family wealth / Family connections / Got a job through parents / Grew up in privilege
modestbgr	From modest background	Started modestly / From poor background / Low income background / Came from modest means / Started from the bottom / Built their own wealth / Self-made
d_premise	No one deserves great wealth	Question whether it is possible to deserve wealth / Assertions that it is not possible to deserve great wealth

Variable	Content	Includes mentions of
u_premise	If one has wealth it is deserved	Question whether it is possible to have something but not deserve it / Assertions that it is not possible to not deserve wealth one already has / He has it therefore he deserves it / what does it mean to not deserve what you have / not my place to say
loopholes	Use loopholes	Loopholes / Offshores / Tax avoidance
noloopholes	Do not use loopholes	Pay their taxes / Do not have offshores / Do not look for loopholes
highertaxes	Should pay higher taxes	Should be taxed more than are / Should be taxed more than others / Should pay more taxes than their secretary / Pay more taxes compared to poor people / Income must be redistributed / There should be an income cap
keepmoney	Deserve to keep money	Use how they see fit / Do what they want with it / Should not be punished / No one can be told what to do with their own money / Entitled to their money / Does not include repetition of "they deserve it" for this gets too close to q wording
fairtreat	Treat others fairly	Treat workers fairly / Earn it without harming others / Earned it honestly / Gives fair wages / Gives good benefits
unfairtreat	Treat others unfairly	Pay low wages / Job harms others / Step all over people / Exploitation / Takes advantage of others
corrupt	Corrupt	Fraud / Illegal means / Scam / Corrupt / Stealing

Variable	Content	Includes mentions of
negative	Negative personality traits	Greedy / Selfish / Spoiled / Dishonest person / Unethical / Cheat / Bad person / Don't notice their privilege / Money-grubbing / Full of themselves / Snobby / Pretentious / Out of touch
positive	Positive personality traits	Generous / A good person / Compassionate / Caring / Understand the less fortunate / Understand working people / Humble / Has a understanding of the plight of the many
conspicuous	Conspicuous consumption	Flashy / Unnecessary things / Things that are overpriced / Lavish spending
modest	Modest habits	Humble / Modest about their wealth / Does not have more than they need / Does not have excess income / wealthy but not elite
lucky	Lucky	Got lucky / in the right place at the right time
morethanneed	Have more than they need	More than they could spend in a lifetime / Doesn't need all that money / Not able to use / Abundance left sitting around / Inability to take riches with us after death / More than they know what to do with
system	Systematic causes	Some people become rich because the system allows it / rich become richer

Table A2: Coding scheme for open-ended reactions to what it means for the rich to "pay their fair share"

Variable	Content	Includes mentions of
generic	Generic support for taxing the rich	Rich should pay their share / rich should pay more than poor / rich pay less in tax than the poor / rich currently don't pay fair share / includes comments that cannot be classified either as supportive of proportional or progressive taxes but do speak to importance of paying one's share
taxeshelp	Taxes help the community / others	Taxes help the poor / taxes provide services
loopholes	Concern about use of loopholes	Accountants / offshores / loopholes / tax avoidance / tax breaks / tax deductions / capital gains tax lower than income tax
corporate	Concern about corporate taxes	Too low corporate taxes / large companies paying too little tax
highertaxes	Explicit support for progressive / higher taxes on the rich	calls to increase taxes on rich / affirmation of progressive taxes / they do not pay enough in taxes / heavily tax the rich
flattax	Support for flat tax	They should pay same share / everyone should pay same rate
keepmoney	Deserve to keep money / pay enough in taxes	They deserve to keep the money / concerns about incentives / already pay enough / should not be punished for having more / taking significantly more is wrong / rich already pay most of the taxes

Variable	Content	Includes mentions of
morethanneed	Have more than they need	Already have enough / have too much / more than they need / can afford to pay more / their lifestyle wouldn't change
incentives	Concern about incentives to work/invest	shouldn't tax them to the point of disadvantage
system	Systematic causes	Some people become rich because the system allows it / rich become richer

In the second stage, two independent coders used the coding scheme to categorize the open-ended responses. This resulted in each response having three codes, one from each coder. As part of conversations with the second set of coders, one additional theme was added: 'system', which indicates responses that attributed agency/causality primarily to the social system rather than to individual actors.

The resulting coding scheme was evaluated for cross-coder consistency; see Tables A3-A5. These tables report both the rate at which a theme was present and the agreement between coders. The first column in each table reports whether a theme was present, where theme presence is defined as the share of responses where all three coders agreed that the response referred to the theme in question. The second column reports the share of responses for which coders unanimously agreed that a theme was not present. The third column reports the share of responses for which the coders disagreed; these were all coded as absence of the theme in the final dataset.

The tables allow the reader to evaluate different ways to evaluate cross-coder consistency. For the purposes of this paper, I define acceptable cross-coder consistency as unanimity among three coders (either for the presence or the absence of a theme) for at least 80%

of evaluated responses. In other words, the rate at which disagreement occurs among any pair of the three coders cannot exceed 20% of evaluated statements. By this criterion, most themes were coded with acceptable cross-coder consistency. However, some of the less commonly occurring themes had high ratios of coder disagreements to coder agreements of theme presence. To guard against idiosyncrasies in coder evaluations inflating the coded rate of theme occurrence, all instances of disagreement were coded as absences of the theme in the final dataset.

In general, the themes were somewhat more reliable when categorizing responses to the (un)deserving rich than when categorizing responses to the "fair share" prompt. One category proved too unreliable to use in further analyses: this was the category "generic" (which indicated vaguely worded support for higher taxes that nonetheless did not clearly indicate support for progressive principles or higher taxes on the rich), which only achieved a cross-coder agreement (as defined above) of 49%. Qualitative consultation with the independent coders confirmed that this was a category that easily overlapped with others, as well as being difficult to apply in cases where the respondents closely mimicked the language of the prompt. As a result, the category was deemed unreliable and is not used in further analyses. All other categories had cross-coder agreement of 80% or higher and are used in the final analysis.

Each response could be coded as corresponding to several themes, so the frequencies below are not an indicator of the share of respondents whose responses matched one or more of the themes. However, most respondents (78%) who were asked about the (un)deserving rich gave a response that matched a theme for at least one of those two prompts, with 43% giving a response that matched a theme both for the deserving and the undeserving prompt. For those assigned the "fair share" prompt, 41% gave a response that matched a theme.

Table A3: Frequency of occurrence and coder consistency of themes in the "undeserving rich" open-ended prompt.

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Hard work	0.04	0.9	0.06
No hard work	0.05	0.91	0.04
Give back	0.18	0.76	0.06
Inheritance	0.12	0.85	0.04
Modest background	0	1	0
Question the premise	0.02	0.83	0.14
Use loopholes	0.01	0.9	0.09
Do not use loopholes	0	0.98	0.02
Should pay higher taxes	0.09	0.87	0.04
Should be able to keep money	0.01	0.93	0.06
Treat workers fairly	0	0.99	0.01
Treat workers unfairly	0.06	0.9	0.03
Corrupt	0	0.96	0.03
Negative personality traits	0.11	0.76	0.13
Positive personality traits	0	0.97	0.03
Conspicuous consumption	0.06	0.9	0.04

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Modest habits	0	1	0
Got lucky	0.05	0.93	0.02
Have more than they need	0.05	0.85	0.1
System made them rich	0.01	0.94	0.05

Table A4: Frequency of occurrence and coder consistency themes in the "deserving rich" open-ended prompt.

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Hard work	0.37	0.5	0.13
No hard work	0	1	0
Give back	0.16	0.77	0.08
Inheritance	0	0.95	0.04
Modest background	0	0.96	0.04
Question the premise	0.01	0.91	0.09
Use loopholes	0	0.98	0.02
Do not use loopholes	0.01	0.98	0.01
Should pay higher taxes	0.03	0.94	0.03
Should be able to keep money	0.03	0.9	0.08
Treat workers fairly	0.02	0.96	0.02

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Treat workers unfairly	0	0.98	0.01
Corrupt	0	1	0
Negative personality traits	0.02	0.95	0.03
Positive personality traits	0.03	0.84	0.13
Conspicuous consumption	0	0.97	0.03
Modest habits	0	0.96	0.04
Got lucky	0.01	0.98	0.01
Have more than they need	0	0.96	0.04
System made them rich	0	0.98	0.02

Table A5: Frequency of occurrence and coder consistency themes in the "paying one's fair share in taxes" open-ended prompt.

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Generic	0.16	0.33	0.51
Taxes help	0.04	0.84	0.12
Loopholes	0.18	0.75	0.07

Variable	Theme present (coders agree)	Theme not present (coders agree)	Theme not present (coders disagree)
Corporate	0.03	0.93	0.04
Higher taxes	0.08	0.72	0.2
Flat tax	0.04	0.91	0.06
Keep money	0.03	0.91	0.06
More than need	0.06	0.83	0.12
Incentives	0.02	0.95	0.03
System	0	0.91	0.09

Appendix A2: Description of pilot survey experiment

Research design

This study was exploratory, and no hypotheses were pre-registered. The study was a conjoint survey experiment, which took the form of vignettes that described hypothetical rich individuals. The rich individuals varied in: household income, source of wealth, work ethic, how they treat workers, use of tax loopholes, charitable giving, and conspicuous consumption.[1] Each treatment variable was randomly assigned within each vignette. Random assignment was also independent across vignettes within respondents, and no profile constraints were applied. An example vignette is included below, with randomized elements shown in bold. Table A6 also shows the full wording of each condition.

> "Imagine a person who has a household income of **[$150,000 per year, which puts them in the top earning 20%] / [$250,000 per year, which puts**

[1] Additionally, each vignette was randomly assigned to be described with male or female pronouns. Gender was not a variable of theoretical interest a priori, and results showed no statistically significant impact of target gender on support for progressive taxation. The variable is included in statistical analyses but not reported on further.

them in the top earning 10%] / [**$500,000 per year, which puts them in the top earning 1%**] of American households. Most of her income comes from [**a good job gotten through family connections**] / [**a large inheritance**] / [**having built a company from the ground up**].

In the workplace, she is [**known as a hard worker who always puts in the effort**] / [**not known as a hard worker, but rather has a tendency to slack off**]. At the same time, she has a reputation for [**treating workers fairly and paying them a good wage**] / [**treating workers unfairly and paying them very low wages**].

She employs an accountant who does her tax returns. She tells the accountant to [**to make use of all available tax loopholes, and to just pay the minimum taxes legally possible**] / [**avoid making use of tax loopholes, and to just pay the tax rates as they are**].

Outside of work, she [**does not give money to charities**] / [**regularly gives money to different charities**]. At the same time she has [**a habit of buying expensive and flashy things**] / [**modest habits and does not often buy expensive and flashy things**]."

Table A6: Full wording of conjoint experiment conditions in pilot experiment

Variable	Condition	Full text
Household income	$150,000	$150,000 per year, which puts them in the top earning 20%

Variable	Condition	Full text
	$250,000	$250,000 per year, which puts them in the top earning 5%
	$500,000	$500,000 per year, which puts them in the top earning 1%
Source of wealth	Family connections	a good job gotten through family connections
	Inheritance	a large inheritance
	Built a company	having built a company from the ground up
Work ethic	Works hard	known as a hard worker who always puts in the effort
	Does not work hard	not known as a hard worker, but rather has a tendency to slack off
Treats workers	Fairly	treating workers fairly and paying them a good wage
	Unfairly	treating workers unfairly and paying them very low wages
Loopholes	Uses loopholes	to make use of all available tax loopholes, and to just pay the minimum taxes legally possible
	Does not use loopholes	to avoid making use of tax loopholes, and to just pay the tax rates as they are
Charity	Does not donate	does not give money to charities

Variable	Condition	Full text
	Donates	regularly gives money to different charities
Consumption	Conspicuous	the habit of buying expensive and flashy things
	Modest	modest habits and does not often buy expensive and flashy things

The outcome variable in this study was the question "What, in your opinion, would be fair when it comes to income taxes for people like him/her?". Respondents indicated their preference on a 5-point Likert scale labeled (from 1 to 5) "Reduced a lot", "Reduced a little", "Stay the same", "Increased a little", and "Increased a lot".

Results

The study was run through the Prolific Academic survey recruitment website in July 2020 with a convenience sample of 599 adult participants who were residents of the United States.[2] Each respondent saw three vignettes, resulting in 1797 vignette ratings. The respondents were 54% male, 72% white, 8% African-American, and 60% college-educated. The median household income was in the $50,000-$60,000 category; a rough estimate of the mean household income places it around $75,000. 11% of respondents (64 individuals) reported a household income of over $150,000. In this study, 53% of respondents identify or lean Democrat, 35% identify or lean Republican, and 12% are non-leaning independents.[3]

[2]The study was approved by the University of Memphis IRB. It included a consent form that was approved as part of this IRB application, and which informed the participants they were participating in research. The consent form is available as part of the survey protocol in the replication materials. The study involved no deception. The participants were paid $1.00 for participating in the study, which equaled an average payment rate of $11.69 per hour; this rate is above Prolific Academic's required minimum hourly rate and above the federal minimum wage in the United States.

[3]To improve representativeness in terms of partisanship over the sample in Study 1, this study was set up to recruit an even number of Democrats, Independents, and Republicans (as self-identified in their Prolific

The results of the experiment show that support for taxing the rich is generally high across conditions. Across all vignette combinations, mean support for taxing the rich (the grand mean) is 3.66 (standard deviation 0.97), meaning that on average, respondents support taxes being "increased a little". Also across all vignettes, only 9% of responses are in favor of decreasing taxes (6% of responses for reducing taxes a little and 2% of responses in favor of reducing taxes a lot). 36% of responses recommend leaving taxes as they are.

The conjoint experiment was analyzed using the 'cregg' package in R (Leeper and Barnfield 2020). The results are visualized in Figure A1, with average marginal component effects (AMCE) in the left panel, and marginal means in the right panel. The gray line in the right panel shows the grand mean of 3.66.

There is a significant impact of the target's household income on the respondents' willingness to increase the target's income taxes. Looking at marginal means, support for increasing taxes for a household that earns \$150,000 is (baseline), which rises to (AMCE , $p < 0.01$) for a household income of \$250,000 and (AMCE , $p < 0.01$) for a household income of \$500,000. The difference between the \$250,000 and \$500,000 incomes is itself also statistically significant.

Out of the non-income variables, the only treatments with substantial and significant effects are how the target treats workers (AMCE 0.36, $p < 0.01$) and their use of tax loopholes (AMCE 0.36, $p < 0.01$); each of these variables causes about a third of a point change on the 5-point scale of the outcome variable. Surprisingly, the estimates for the target's work ethic (AMCE 0.08, $p = 0.07$) as well as the source of their wealth (inheritance AMCE 0.01, $p = 0.88$; family connections AMCE 0.06, $p = 0.25$), while in the expected direction, were substantially smaller and not statistically significant at the 0.05 level. Likewise, the behaviors of charitable giving (AMCE 0.08, $p = 0.09$) and conspicuous consumption (AMCE

Academic profiles). Due to a combination of a) not all participants reporting the same party identification in this survey as they had provided in their participant profiles on Prolific Academic and b) the recruited independents leaning more Democratic than Republican, the final survey was not an even partisan split, but it was nonetheless an improvement over Study 1.

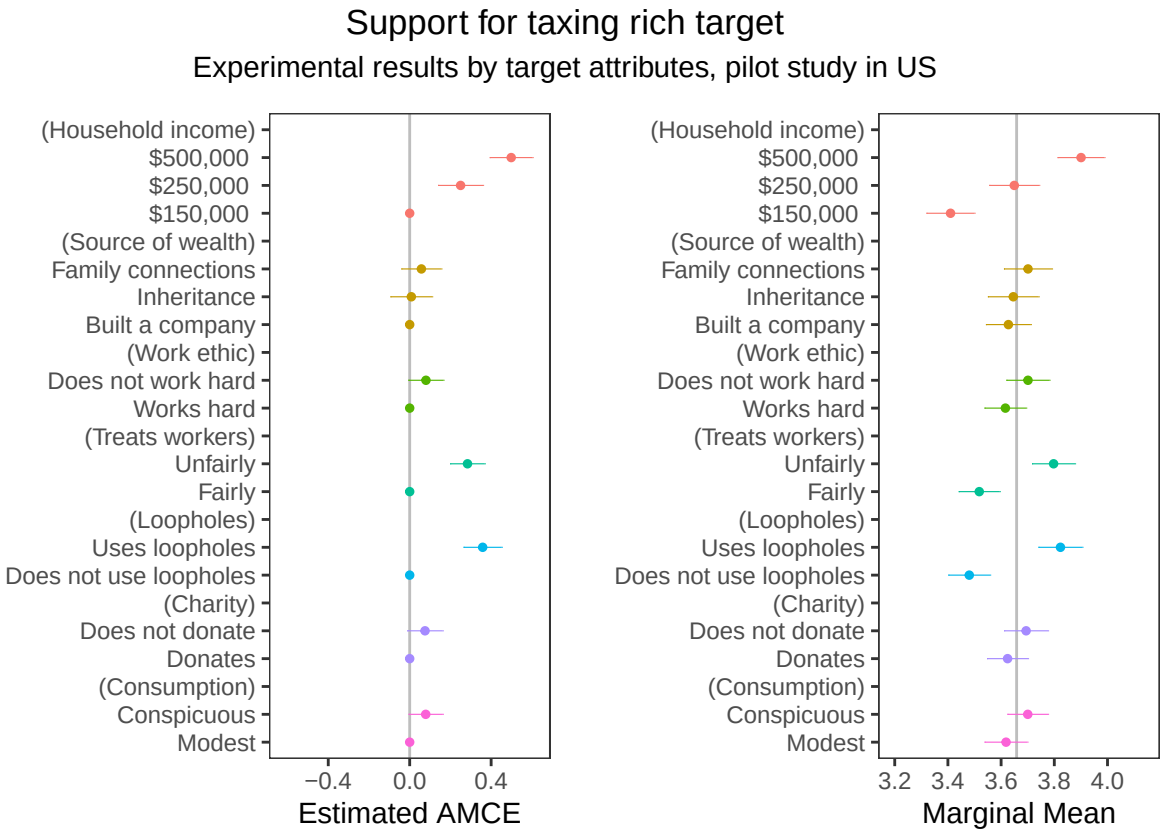

Figure A1: Results from pilot conjoint experiment. Showing average marginal component effects (AMCE) on the left and marginal means on the right. Higher values of the outcome variable indicate stronger support for increasing income taxes for the target individual. The grey line in the marginal means plot is the grand mean.

0.08, p = 0.07) had estimates in the expected direction, but the estimated effect sizes were substantively small and statistically insignificant.

To explore whether there were partisan differences in responses to the components of the vignette, Figure A2 visualizes the results by respondent partisanship (leaning independents are included in partisan groups; non-leaning independents are excluded from analysis). The two left panels show AMCEs for Republicans and Democrats separately; the right-most panel visualizes the *difference* in AMCEs between Republicans and Democrats. Note that Figure A2 shows estimated causal effects, not differences in support for taxation. Baseline support for taxation shows predictable and substantial partisan differences that persist across treatment conditions, with Republicans across the board being less likely to endorse higher income taxes for the rich. Figure A3 visually illustrates these differences in means.

Looking at differences in causal effects, we see that the two groups react similarly to most treatment conditions. Republicans react less strongly to the $500,000 household income condition, though for both groups household income remains one of the strongest predictors of willingness to increase taxes. There is also a small difference in how Democrats and Republicans react to inheritance as a source of income; while the effect of inheritance (compared to building a company from the ground up) is insignificant for both groups, the difference between Republicans and Democrats is significant, with Democrats slightly more willing to tax income acquired from "a large inheritance".

There were no partisan differences in reactions to the remaining variables: Democrats and Republicans react similarly and strongly to indications that a person treats their workers unfairly, and that they make use of tax loopholes to the largest extent legally possible. They are also similarly unresponsive to indicators that a person is not a hard worker, does not donate to charity, or engages in conspicuous consumption.

Finally, the results were analyzed for the potential pitfalls of conjoint experiments discussed in Scott F. Abramson, Kocak, and Magazinnik (2022). Out of the concerns they raise,

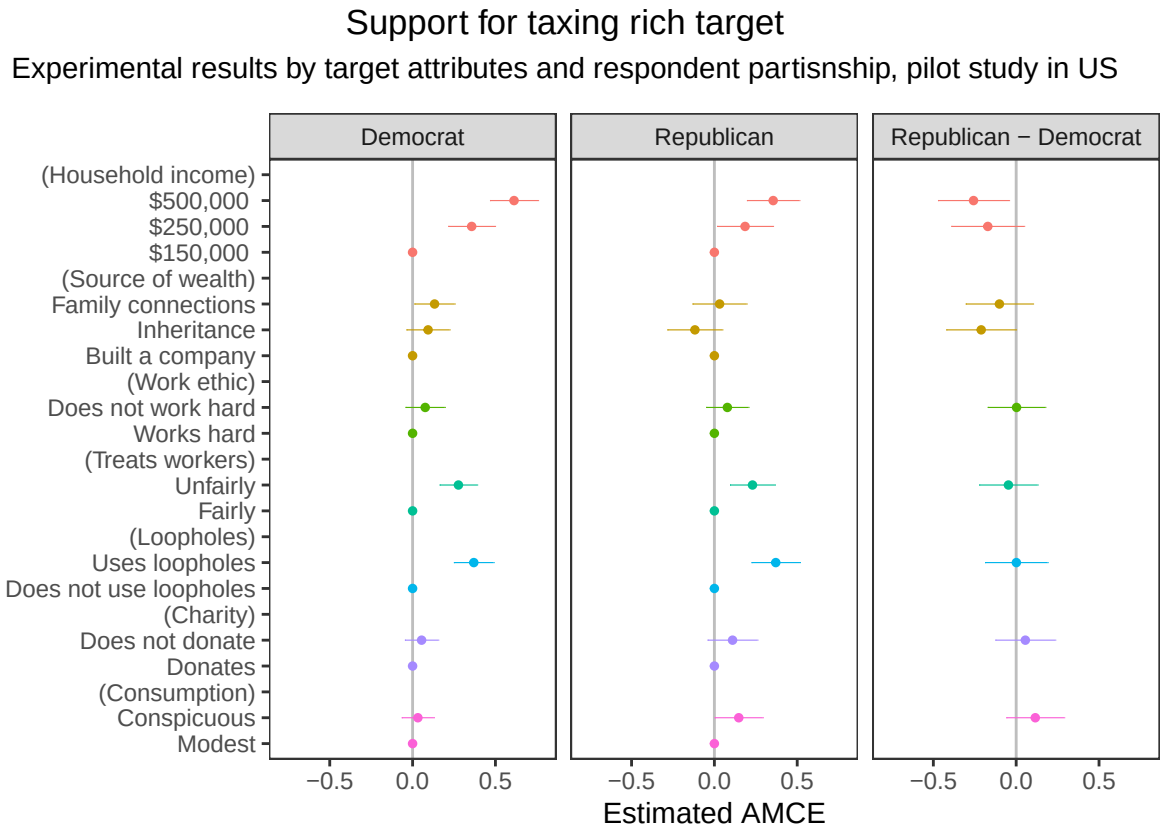

Figure A2: Results from pilot conjoint experiment. Showing Average Marginal Component Effects (AMCE) by partisanship. The right-most panel shows differences in causal effects between Republicans and Democrats.

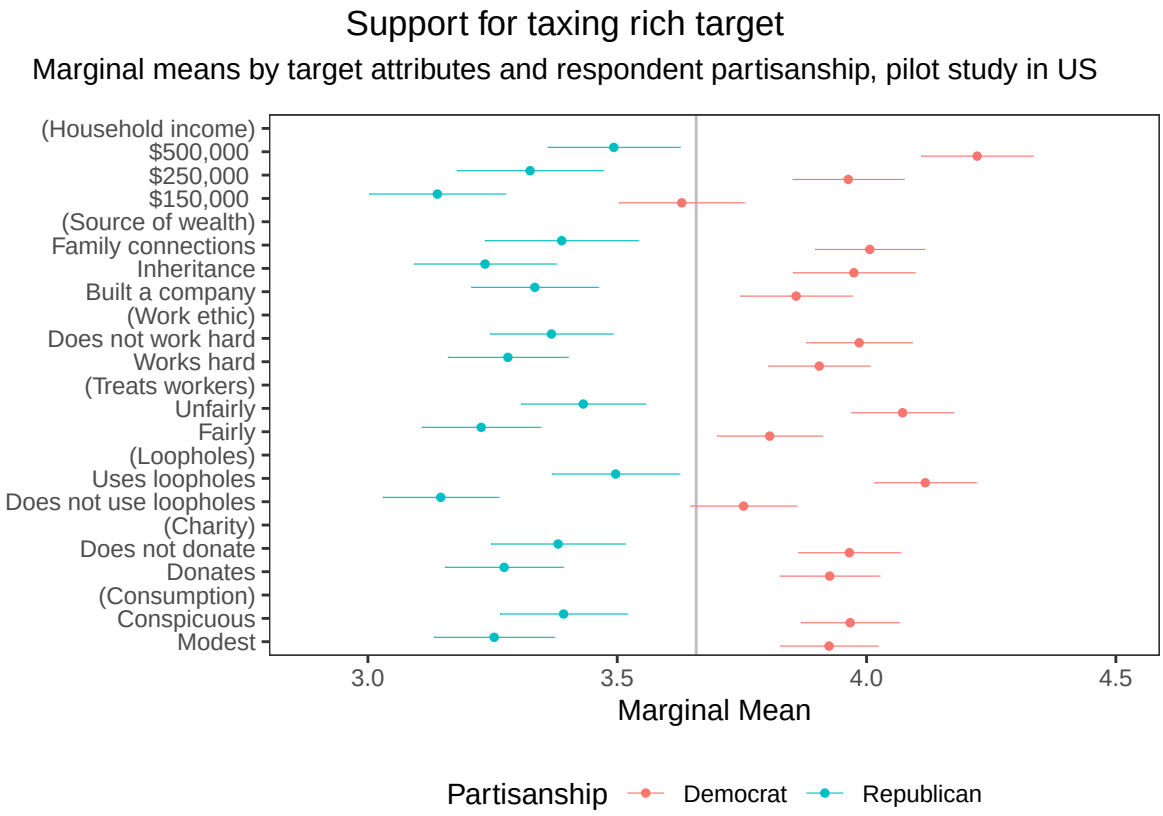

Figure A3: Results from pilot conjoint experiment. Showing marginal means by partisanship. Larger numbers on the outcome variable indicate support for increasing the target's income taxes (1-5 scale).

the most relevant to this study is the risk of preference cycling hiding a majority preference when a treatment feature has more than two levels which are not unambiguously ordinal. In this study, this raises the question of whether the three-level source of income feature may exhibit preference cycling. Using the random forest analysis recommended in Scott F. Abramson et al. (2020), I explored the data for indications of such cycling and found none. I additionally explored heterogeneous effects across demographic characteristics and similarly did not find reasons for concern. Nonetheless, out of caution for the points raised in these articles, the follow-up studies are simplified so that each feature only has two treatment levels, eliminating the risk of preference cycling. In addition to this methodological improvement, this also allows a clearer theoretical distinction between family background and individual work ethic in the follow-up experiments.

Appendix A3: Details for Studies 2 and 3

Tables A7 and A8 show the demographic characteristics of the Danish and American samples, respectively. The tables show the demographic breakdown of the sample, and the target breakdown for a representative sample as defined by Lucid for the United States and YouGov for Denmark, respectively. The Danish sample is representative on all targeted demographic criteria. The US sample is broadly representative, but with a slight over-representation of Hispanic respondents and a slight under-representation of white respondents.

Table A9 shows the full wording of each treatment in the vignettes used in Studies 2 and 3. An example vignette with full wording is presented in the main text of the article.

These studies were approved by the University of Memphis IRB. Study 2 included a consent form that was approved as part of the IRB application, and which informed the participants they were participating in research. The consent form is available as part of the survey protocol in the replication materials. Study 3, which was entirely run on YouGov's proprietary survey platform, did not include a separate consent form, as the IRB

Table A7: Sample demographic characteristics, United States sample

Variable	Percent in sample	Percent target, US
Gender		
Male	48.55	50
Female	51.45	50
Age		
18-24	15.29	13
25-34	19.40	20
35-44	20.60	21
45-54	14.36	15
55-64	14.25	15
65+	16.11	16
HH income		
Less than 25k	22.58	25
25-49k	25.32	25
50-74k	18.79	18
75-99k	11.95	11
100-124k	8.71	8
>125k	12.66	13
Region		
Midwest	22.85	22
Northeast	17.75	18
South	37.15	37
West	22.25	23
Ethnicity		
Hispanic	28.71	11
Not Hispanic	71.29	89
Race		
Black	10.47	13
White	64.55	72
Other	22.90	15
Education		
Postgraduate	10.85	10
Associates/BA	34.36	34
Some College	17.70	16
High School	27.95	29
Less than high school	9.15	11

Note:

United States column shows demographic targets used by Lucid.

Table A8: Sample demographic characteristics, Danish sample

Variable	Sample	Denmark
Gender		
Male	48.6	50.6
Female	51.4	49.4
Age		
18-34	27.0	27.4
35-54	33.6	32.6
55-69	23.0	22.4
70+	16.4	17.5
Region		
Capital Region	33.2	31.6
Zealand	14.3	14.5
Southern Denmark	19.9	21.1
Central Denmark	22.0	22.6
North Denmark	10.5	10.3
Education		
Low	66.9	68.2
Middle	22.7	22.4
High	10.4	9.4

Note:

Denmark column shows demographic targets used by YouGov.

determined that YouGov's own informed consent procedure when recruiting participants into the respondent panel was sufficient to ensure informed consent. The study involved no deception. The participants were paid by Lucid and YouGov at rates determined by these companies; while the researcher paid the companies for the service of recruiting participants, the researcher did not directly pay any participants and had no control over the compensation level for the individual participants.

Table A9: Full wording of conjoint experiment conditions in studies 2 and 3

	Variation 1	Variation 2
Gender	He	She
Income (US only)	$250,000 per year, which is more than what 9 out of 10 American households earn	$500,000 per year, which is more than what 99 out of 100 American households earn
Income (DK only)	50,000 kroner gross per month, which is more than what 9 out of 10 Danes earn	130,000 kroner gross per month, which is more than what 99 out of 100 Danes earn
Family background	a modest family background and did not inherit money	a well-off family, and inherited a large amount of money
Work ethic	known as someone who works hard and always puts in the effort	not known as someone who works hard, but rather has a tendency to slack off
Treat workers	treating workers fairly and paying them a good wage	treating workers unfairly and paying them very low wages

	Variation 1	Variation 2
Use of loopholes	he employs an accountant who helps him make the most use of tax loopholes, so that he pays the smallest amount of taxes legally possible	even though he could employ an accountant to help him make use of tax loopholes and pay less taxes, he chooses not to do this and just pays the tax rates as they are
Charitable giving	regularly gives money to different charities	does not give money to charities
Conspicuous consumption	modest habits and does not often buy expensive and flashy things	the habit of buying expensive and flashy things for himself

Appendix A4: Additional pre-registered analysis of the 'hard work' variable

The pre-registration plan for studies 2 and 3 specified a second analysis of the 'hard work' treatment variable. This was pre-specified due to a concern that a combination of a poor family background and poor work ethic may send contradictory/confusing signals about how the target individual has acquired a high income. Therefore, the impact of the work ethic variable was to be additionally estimated only among vignettes where the target is described as being from a wealthy background. A significant effect of 'hard work' found when restricting the analysis to rich targets would be considered evidence in support of the hypothesis that perceptions of "hard work" matter.

Figure A4 reproduces the conjoint analyses in studies 2 and 3 with both datasets subset only to rankings of targets that were described as being from a well-off family background.

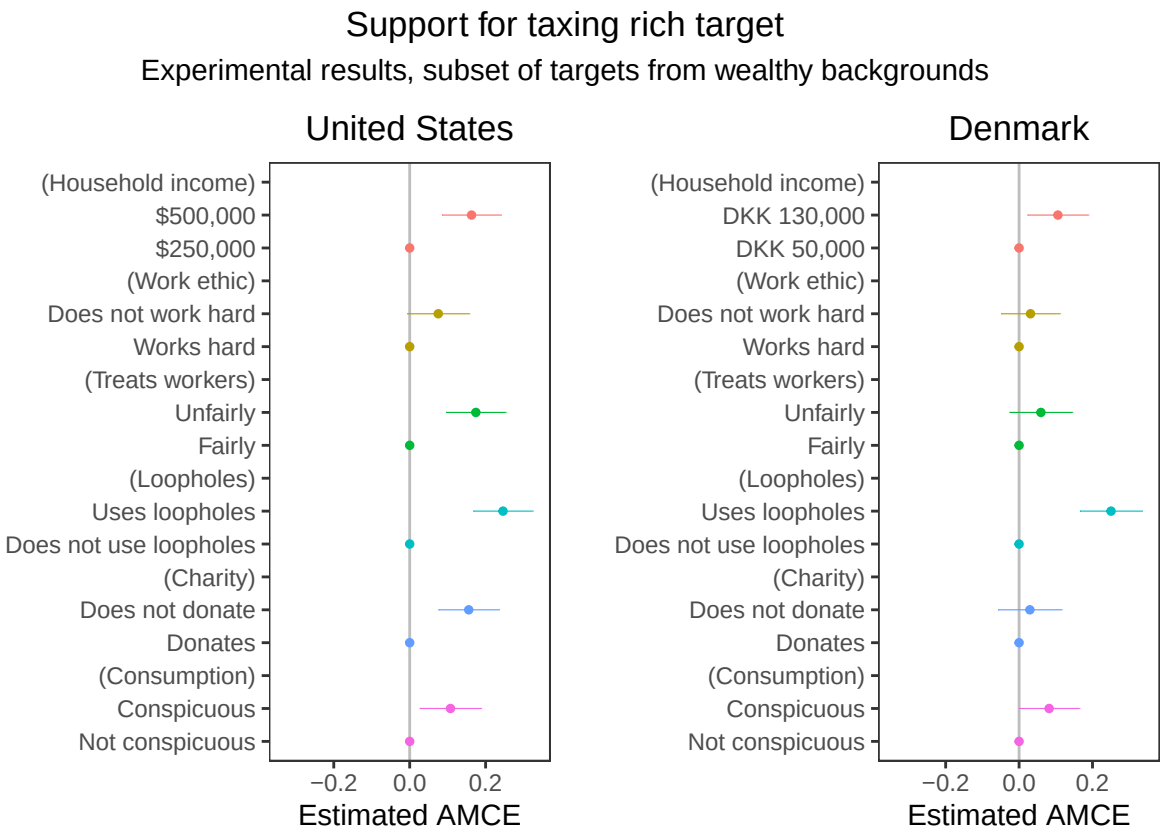

Figure A4: Results from conjoint experiments in Studies 2 and 3. Showing pre-registered analysis when subsetting to only vignettes where the target was from a well-off family background.

In both studies, support for increasing taxes on the target from a well-off family is not statistically significantly higher if the target is described as "not working hard". Therefore, no evidence of a significant effect by which hard work justifies lower taxes was found in this pre-registered analysis.

Appendix A5: Equivalence tests

Tables A10-A12 show equivalence tests for the main experimental findings in the US and Denmark (Table A10), as well as for the interaction analysis comparing results across countries (Table A11) and across political leaning within countries (Table A12). The implementation of these tests relies on instruction materials made publicly available by Heiss (2023) and Rainey (2023).

Equivalence tests are a way to explore whether findings fall into a sufficiently narrow range around zero so as to be considered negligible. The null hypothesis in an equivalence test is that the coefficient falls *outside* a researcher-defined range of negligible effects. Low p-values indicate that data is inconsistent with the null hypothesis, or in other words that data is consistent with the effect size being negligible. High p-values indicate that data is inconsistent with the effect size being negligible.

In all analyses in this section, I have chosen to specify an effect size in the range (-0.1 to 0.1) as negligible. For the interaction effects, I use the same negligible range; *differences* in effects that are smaller than (-0.1 to 0.1) are considered negligible.

Table A10: Equivalence test results for the main experimental effects, US and DK. Showing p-values for the null hypothesis that the causal effect is zero, and equivalence p-values for the null hypothesis that the causal effect falls outside the negligible effect interval defined as (-0.1, 0.1).

Term	US Estimate	US Std. Error	US p-value	US Equiv. p-value	DK Estimate	DK Std. Error	DK p-value	DK Equiv. p-value
Charity	0.10	0.03	0.00	0.56	0.06	0.03	0.07	0.07
Consumption	0.12	0.03	0.00	0.76	0.10	0.03	0.00	0.51
Family background	0.05	0.03	0.06	0.05	0.04	0.03	0.22	0.01
Hard work	0.07	0.03	0.03	0.12	0.01	0.03	0.61	0.00
Income	0.14	0.03	0.00	0.94	0.15	0.03	0.00	0.94
Loopholes	0.25	0.03	0.00	1.00	0.25	0.03	0.00	1.00
Treat workers	0.16	0.03	0.00	0.98	0.10	0.03	0.00	0.56

Table A11: Equivalence test results for the difference in experimental effects between respondents in the US and DK. Showing p-values for the null hypothesis that the difference in causal effects is zero, and equivalence p-values for the null hypothesis that the difference in causal effects falls outside the negligible difference interval defined as (-0.1, 0.1).

Term	US vs. DK Estimated Difference	Std. Error	P-value	Equiv. p-value
Charity	-0.05	0.04	0.25	0.11
Consumption	-0.02	0.04	0.64	0.03
Family bgr	-0.02	0.04	0.69	0.02
Hard work	-0.05	0.04	0.23	0.12
Income	0.00	0.04	0.97	0.01
Loopholes	-0.01	0.04	0.90	0.01
Treat workers	-0.06	0.04	0.17	0.16

Table A12: Equivalence test results for the difference in experimental effects between left-wing and right-wing respondents in the US and DK. Showing p-values for the null hypothesis that the difference in causal effects is zero, and equivalence p-values for the null hypothesis that the difference in causal effects falls outside the negligible difference interval defined as (-0.1, 0.1).

Term	US Estimated Difference	US Std. Error	US p-value	US Equiv. p-value	DK Estimated Difference	DK Std. Error	DK p-value	DK Equiv. p-value
Charity	-0.12	0.06	0.05	0.62	0.00	0.07	0.99	0.07
Consumption	0.07	0.06	0.21	0.34	-0.01	0.07	0.93	0.08
Family bgr	0.03	0.06	0.56	0.14	-0.04	0.07	0.57	0.18
Hard work	0.00	0.06	0.99	0.05	0.09	0.07	0.17	0.46
Income	-0.07	0.06	0.25	0.30	-0.13	0.07	0.06	0.66
Loopholes	-0.18	0.06	0.00	0.91	-0.08	0.07	0.23	0.39
Treat workers	0.14	0.06	0.02	0.76	0.11	0.07	0.11	0.54

Appendix A6: Heterogeneous effects by respondent income

Figures A5 and A6 show an exploratory analysis of differences in treatment effects by respondent's household income, separately for the United States and Denmark. Respondents are divided into two groups: household incomes above the sample median, and household incomes at or below the sample median. Results show similar patterns across respondent income groups.

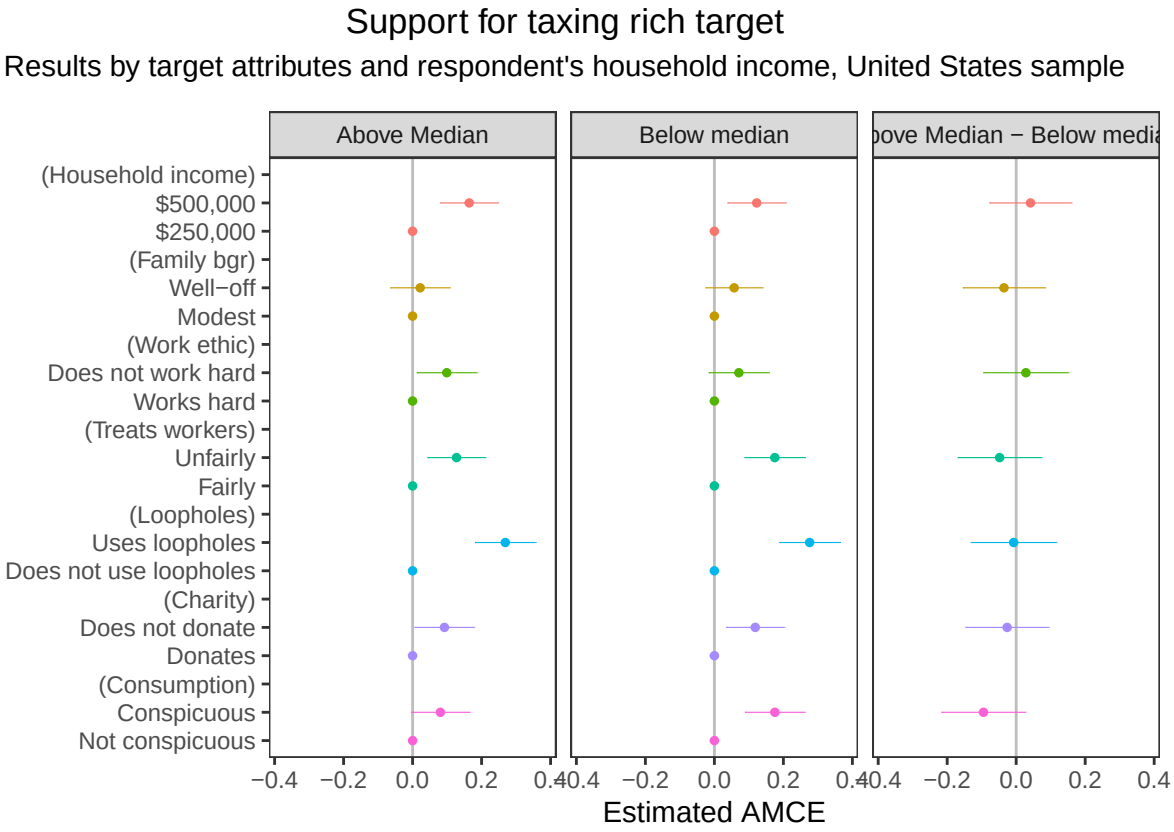

Figure A5: AMCE by target attributes and respondent household income.

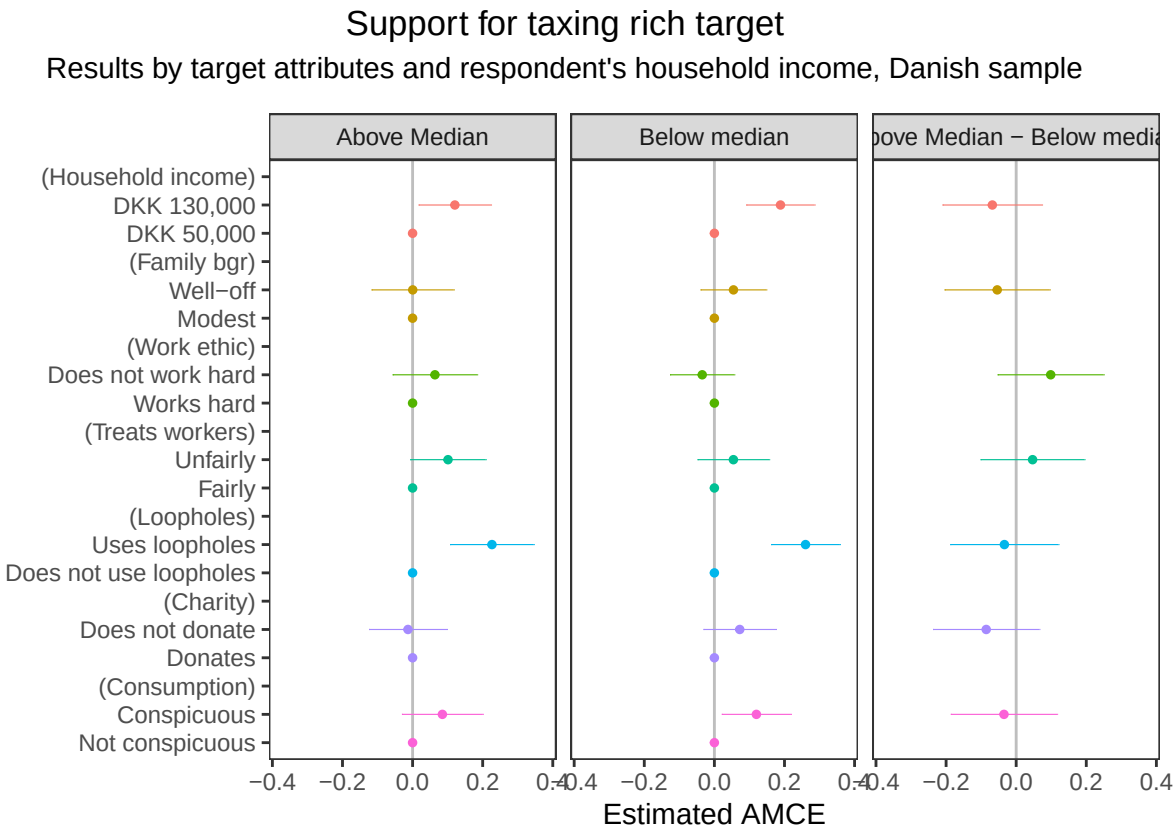

Figure A6: AMCE by target attributes and respondent household income.

Appendix A7: Comparison of marginal means across US and Danish results

Figure A7 displays the marginal means for the United States and Denmark side by side and on a 3.0 to 4.0 axis to visualize cross-national differences in underlying support for progressive taxes. The national differences in marginal means co-exist with similar size causal effects in the conjoint experiment, meaning that the national differences persist in each treatment condition, and that reactions to the treatments in the vignettes are similar across countries.

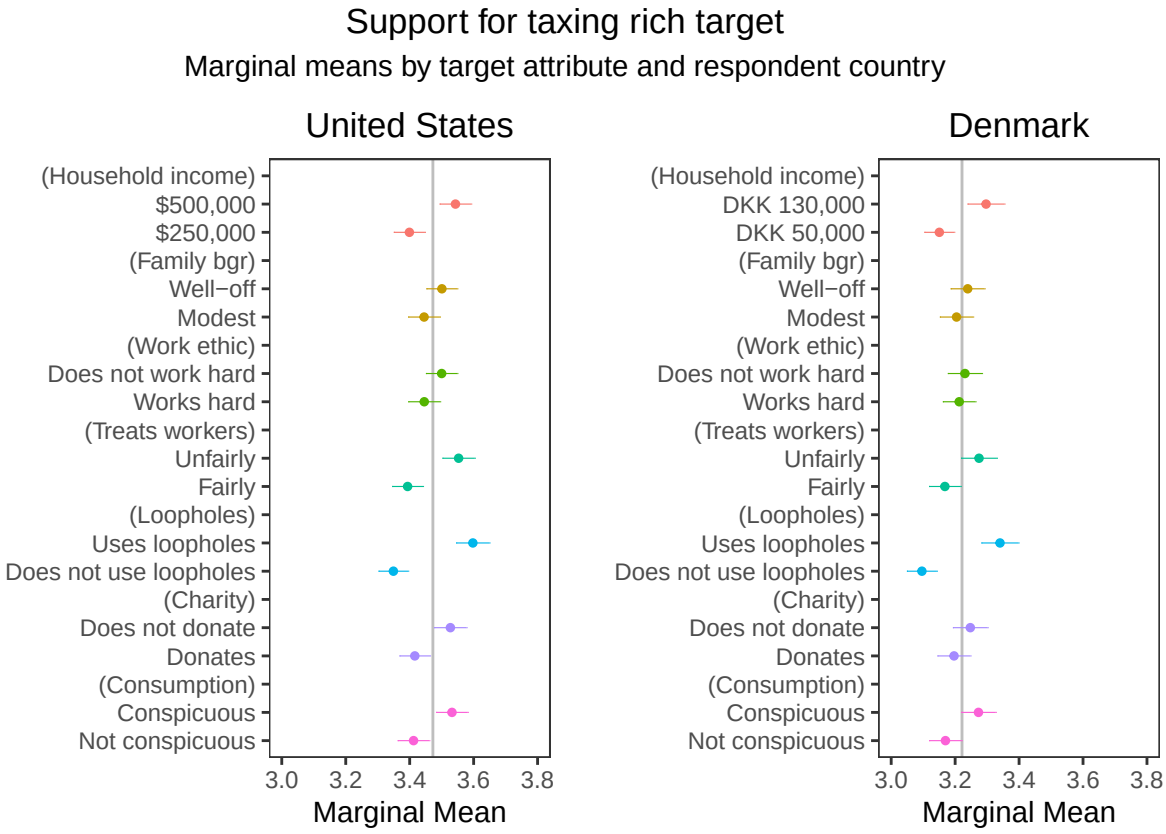

Figure A7: Results from conjoint experiments in Studies 2 and 3. Showing marginal means for the United States on the left, and Denmark on the right. Higher values of the outcome variable indicate stronger support for increasing income taxes for the target individual. The grey lines in the marginal means plots are the grand mean for each sample.

Appendix A8: Differences in marginal means by party

Figure A8 shows differences in marginal means by party and country. As the visualization shows, overall support for increasing the income taxes of the hypothetical targets is higher in the United States than in Denmark. Differences between Danish and American respondents, and left-wing and right-wing respondents, remain significant in all treatment conditions, while reactions to experimental treatments are similar across groups.

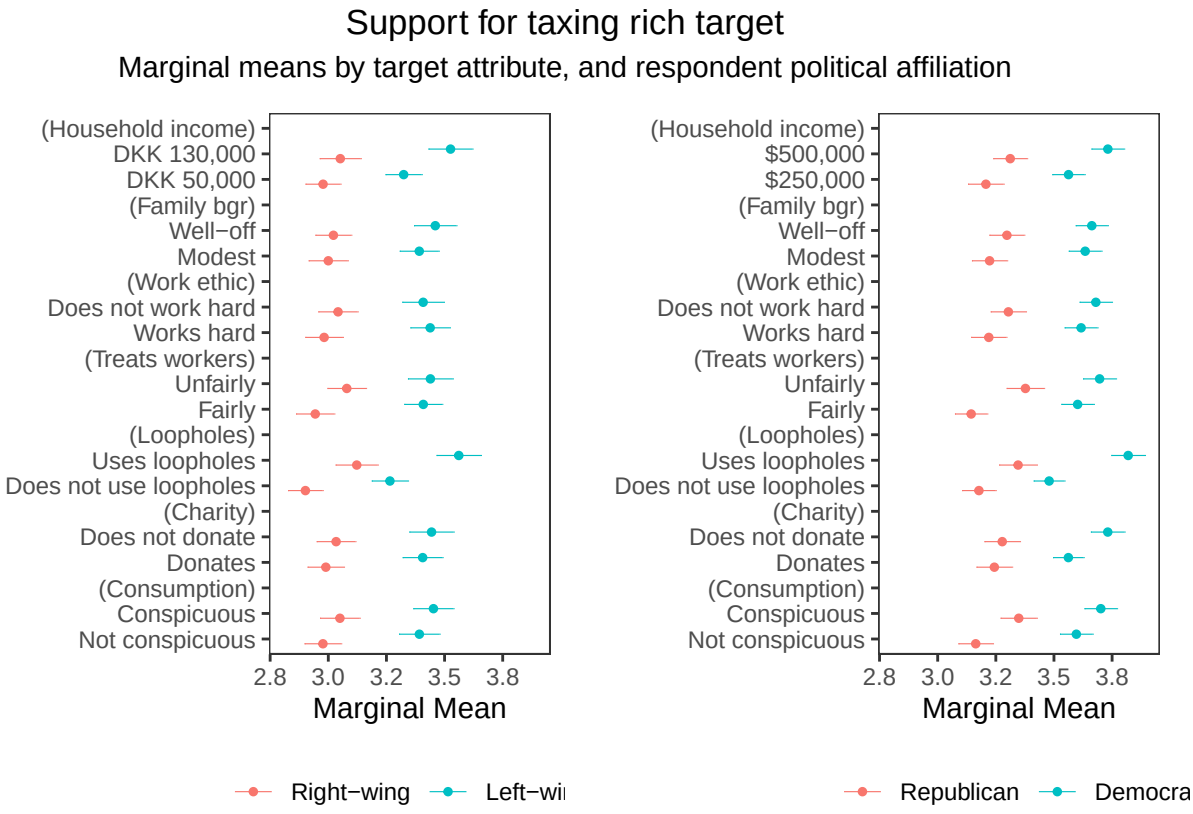

Figure A8: Marginal means by treatment condition, displayed separately by political affiliation and country.

References

Abramson, Scott F., Korhan Kocak, and Asya Magazinnik. 2022. "What Do We Learn about Voter Preferences from Conjoint Experiments?" *American Journal of Political Science* 66 (4): 1008–20. https://doi.org/10.1111/ajps.12714.

Abramson, Scott F, Korhan Kocak, Asya Magazinnik, and Anton Strezhnev. 2020. "Improving Preference Elicitation in Conjoint Designs Using Machine Learning for Heterogeneous Effects," 52.

Heiss, Andrew. 2023. "The Ultimate Practical Guide to Conjoint Analysis with R." https://www.andrewheiss.com/blog/2023/07/25/conjoint-bayesian-frequentist-guide/.

Leeper, Thomas J., and Matthew Barnfield. 2020. "Cregg: Simple Conjoint Tidying, Analysis, and Visualization."

Rainey, Carlisle. 2023. "Equivalence Tests with {Marginaleffects}." https://www.carlislerainey.com/blog/2
08-18-equivalence-tests/.