

Preferences for Rights*

Aviv Caspi

Julia Gilman

Charlie Rafkin

December 2024

Abstract

Political debates often invoke “rights” to justify public transfers (e.g., the right to health care), whereas economists use welfarist frameworks which evaluate transfers’ impacts based on how they affect people’s utility. We conduct real-stakes online experiments that isolate non-welfarist from welfarist motives, and find sizable non-welfarist preferences to provide health care and legal aid to the indigent. 73% of participants make choices which are incompatible with welfarism. Non-welfarist concerns are weaker but still pervasive with neutral comparison goods. Additional experiments highlight drivers of non-welfarist motives and a key policy implication: non-welfarist concerns make Social Welfare Functions less progressive.

*Stanford, MIT, and U.C. Berkeley (caspi@stanford.edu, jhgilman@mit.edu, and crafkin@berkeley.edu). We are grateful to Abhijit Banerjee, Amy Finkelstein, Jacob Goldin, Daniel E. Ho, Jim Poterba, and Frank Schilbach for their guidance and support. For helpful suggestions, we thank Jenna Anders, Peter Andre, Doug Bernheim, Alexander Cappelen, John Conlon, Stefano DellaVigna, Esther Duflo, Joel Flynn, Luca Henkel, Peter Hickman, Louis Kaplow, Muriel Niederle, Ted O’Donoghue, Alex Rees-Jones, Chris Roth, Emmanuel Saez, Ian Sapollnik, Jesse Shapiro, Adam Solomon, Krishna Srinivasan, Dmitry Taubinsky, Bertil Tungodden, Pierre-Luc Vautrey, Matthew Weinzierl, and particularly Shakked Noy, along with seminar participants at BERG, CELS, MIT, SITE Psychology and Economics, and Stanford. We thank The Works, Inc. (TWI) for implementing incentives in this study. Rafkin gratefully acknowledges financial support from TWI to cover graduate school tuition and release time from teaching. We also gratefully acknowledge funding from The Institute of Consumer Money Management (ICMM) Pre-doctoral Fellowship on Consumer Financial Management, awarded through the National Bureau of Economic Research; Stanford Law School’s John M. Olin Program in Law and Economics; the National Science Foundation Graduate Research Fellowship under Grant No. 1122374; the MIT Shultz Fund; Stefano DellaVigna; and Frank Schilbach. This study was approved by MIT’s Committee on the Use of Humans as Experimental Subjects under protocol #2112000534 and #E-5131, and pre-registered at the AEA RCT Registry under #0012065 and #0014218. First version: November 8, 2023.

Economists often use “welfarist” Social Welfare Functions, which evaluate policies based on their consequences for individuals’ utilities (Kaplow, 2024). A criticism of welfarism is that it fails to engage with non-consequentialist or other moral concerns that animate real-world debates. Political arguments for redistribution, for instance, focus on moral imperatives as much as marginal utility. In principle, welfarism can accommodate concerns for fairness (Rabin, 1993; Cappelen et al., 2007, 2013, 2020, 2023), liberal institutions (Dal Bó et al., 2010; Dal Bó, 2014), rights (Suzumura, 2011), or social preferences (Becker, 1974; Fehr and Schmidt, 1999): if these concerns enter people’s utility functions, welfarists value them too. Yet in practice, scholars who see welfarism as too detached from moral reasoning have sought alternatives (Tobin, 1970; Rawls, 1971; Sen, 1985; Weinzierl, 2014, 2017; Saez and Stantcheva, 2016). Still, welfarism remains popular in applied research, in part because non-welfarism faces its own objections (e.g., Kaplow and Shavell, 2001; Sher, 2024).

The disconnect between economists’ welfarist evaluations and real-world non-welfarist rhetoric is especially stark in debates about public transfers of goods to needy people. Here, policymakers often appeal to “rights,” like the rights to health care, food, or legal counsel, but only for certain goods.¹ This divergent rhetoric is consistent with Walzer (1983)’s influential *Spheres of Justice*, which holds that non-welfarist principles are domain-specific and different goods require distinct welfare criteria. In contrast, economists’ textbook arguments for such government transfers (e.g., Currie and Gahvari, 2008) rely on a one-size-fits-all welfarism that emphasizes efficiency, market failures, and paternalism, while giving less attention to non-welfarist concerns. Incorporating non-welfarist preferences into economic analysis of transfers could better ground welfare assessments in real-world debates, and explain motives for redistribution across domains.

This paper takes an empirical step forward, in the spirit of the “positive welfare economics” agenda, documenting the prevalence and implications of non-welfarist preferences to provide certain transfers.² How common are such non-welfarist concerns, how do they differ across domains, and how do they influence Social Welfare Functions and redistributive policy? To answer these questions, we conduct incentivized experiments in which “Spectators” transfer real goods to needy, low-income “Recipients.” We detect high rates of non-welfarist behaviors even in neutral domains, like allocating bus tickets, where societies rarely ascribe rights. These non-welfarist preferences become yet more prevalent when Spectators allocate “rights goods” — in the experiments, health care or lawyers — aligning with ideas in *Spheres of Justice*. Going further, we show how these non-welfarist preferences reduce Social Welfare Functions’ progressivity,

¹President Barack Obama said, “Health care is not a privilege for the fortunate few — it is a right” (Obama, 2013). Since 2017, 17 U.S. cities and four states passed “right to counsel” policies, which give free lawyers to defendants in eviction cases. The American Civil Liberties Union argues, “Tenants’ right to legal representation in eviction cases is a civil liberties issue, a gender justice, racial justice, and economic justice issue” (ACLU, 2022).

²Recent work in this agenda includes Ambuehl et al. (2021), Chen and Schonger (2022), Arrieta and Bolte (2023), Bartling et al. (2023), Hvidberg et al. (2023), Ambuehl and Bernheim (2024), Andre (2024), Bénabou et al. (2024), Bernheim et al. (2024), Capozza and Srinivasan (2024), and Exley and Kessler (2024).

leading Spectators to favor universal provision rather than policy targeting, and correlate with policy support. Additional experiments probe which mechanisms trigger non-welfarist concerns.

Understanding whether non-welfarist preferences exist in the population, and if they vary across goods, is important for three reasons. First, their existence would suggest a middle-ground between non-welfarism and welfarism (Saez and Stantcheva, 2016; Fleurbaey and Maniquet, 2018): if non-welfarist concerns are inside people’s utility functions, as our experiments show, then even welfarist social planners respect such concerns. Second, understanding non-welfarism across domains helps make predictions about support for redistributive policy. Third, understanding real-world reasoning about policies is intrinsically valuable and sheds light on economic primitives (Hausman and McPherson, 1993; Stantcheva, 2021, 2022; Andre et al., 2022, 2023).

A simple framework guides the experiments (Section 1). Individuals are welfarist if they derive utility over societal allocations of a good based only on the utility that allocations provide to recipients. Non-welfarists, however, derive utility from allocating a good even fixing recipients’ utility. Such concerns resemble warm glow in giving (Andreoni, 1989, 1990) or intrinsic preferences not to lie (Abeler et al., 2019), among others. For instance, preferences for the “right to legal counsel” would be non-welfarist if people get a utility bonus from allocating lawyers, even when lawyers do not affect recipients’ utility. In contrast, motives like other-regarding preferences are welfarist, as these require recipients to get utility from lawyers. We prove that a purely welfarist Social Welfare Function (SWF), which aggregates people’s utilities, can replicate a non-welfarist SWF if non-welfarist concerns directly enter people’s utility functions. But whether people in society are actually non-welfarist is an empirical question, setting the stage for our experiments.

In these experiments, $N = 1,800$ Spectators on Prolific face incentivized choices about allocating goods to low-income Recipients in Memphis, Tennessee (Section 2). We study how Spectators make allocation decisions, isolating non-welfarist concerns by carefully shutting down instrumental reasons to give goods. We emphasize the difference in behaviors with rights goods (lawyers, health care) versus neutral benchmark goods (YMCA memberships, bus tickets). Beyond capturing economically important heterogeneity across policy domains, these differences offer a lower bound on non-welfarist preferences: they purge elicitation errors and other issues that may affect levels of behaviors in an online sample. Our field implementation adds realism to the laboratory context, as we partner with a Memphis nonprofit to enact some Spectator choices about lawyers and benchmarks. Health care choices are always hypothetical, so we embed several tests for the impact of incentives and reject even small differences in behavior from lack of incentives.

Our first three experiments test for the existence and magnitude of non-welfarist concerns (Section 3). Experiment 1 is motivated by the philosophical principles of “inalienability” — i.e., that there is a particular harm from taking away a good or freedom — and procedural justice.

Specifically, Experiment 1 creates a situation where only non-welfarist Spectators would pay money to prevent the experimenter from changing how goods are allocated. We tell Spectators

that a lottery allocated a good to one of two *ex ante* identical Recipients. We then elicit Spectators' willingness to pay (WTP) to preserve the results of the lottery. We tell them that if they pay nothing, the experimenter will rerun the lottery, which risks taking from the first lottery-winner and giving to another Recipient instead. Critically, we shut down reference-dependence or endowment effects, as we inform Spectators — and emphasize with confirmation checks — that neither Recipient knows the initial lottery result. Welfarists' WTP to preserve the lottery, then, should be zero.

Yet Spectators often pay to preserve the lottery, even though welfarists would not. When the lottery involves the benchmark goods, 46% of Spectators have positive WTP to preserve the lottery. Supporting that non-welfarist behaviors are partially domain-specific, we find that these preferences are even more pronounced with rights goods. With rights goods, 53% of Spectators exhibit a positive WTP. The 7.8 pp difference (s.e.: 2.5 pp; p -value = 0.002) also rejects the null hypothesis that all Spectators are welfarists, even under the adversarial assumption that non-welfarism with benchmarks only reflects elicitation errors.

Experiment 2 examines the non-welfarist principle that people should have the freedom to choose (related to, e.g., the capabilities approach in Sen, 1985). We create a situation where only non-welfarist Spectators have positive WTP to put a good in a Recipient's choice set. Consider a Spectator who is certain that a Recipient would choose $\$y$ in cash over a lawyer. The Spectator should never pay to let the Recipient choose between the two rather than giving them $\$y$ directly, unless the Spectator intrinsically values providing choice. Experiment 2 elicits Spectators' beliefs that Recipients choose $\$y$ in cash over the good. Focusing on Spectators who are certain the Recipient chooses cash, we elicit Spectators' WTP to provide choice rather than $\$y$ directly. Thus, Experiment 2 trades off pecuniary costs of giving choice against possible non-welfarist benefits.

Consistent with non-welfarism, Spectators have positive WTP to give choices that they believe will not be exercised. As in Experiment 1, we detect non-welfarist concerns even for the benchmark goods. Among Spectators who are at least 90% sure the Recipient will choose cash, 39% are still willing to pay to provide the choice, meaning they burn money if the choice is not exercised. These preferences are stronger with rights goods, where 55% of these Spectators have positive WTP for choice (s.e. of difference: 2.8, $p < 0.001$). In fact, WTP for choice is 0.52 s.d. higher (s.e.: 0.11) for rights goods than benchmarks even among Spectators who express 100% certainty that the Recipient would not choose the good. Such large and significant differences again suggest domain-specificity. These differences also yield a meaningful lower bound on non-welfarist preferences if one is worried about elicitation issues, noise in beliefs, or Spectators internalizing Recipients' intrinsic value of *making* choices (Bartling et al., 2014; Freundt et al., 2023).

To examine whether the principle of equal rights differs across domains, Experiment 3 tests for domain-specific inequity aversion. Here, we find more mixed evidence. Consistent with preferences for equality, Spectators have higher WTP to give another Recipient the good when we tell them that 9 out of 10 possible Recipients already have the good, versus when 1 or 5

out of 10 do. But, we do not find that the increase in WTP differs across domains, and cannot reject the hypothesis that all behaviors reflect inequity aversion rather than non-welfarism. To be conservative, our interpretation is that we detect evidence of non-welfarism in Experiments 1 and 2 only, and focus on those behaviors in the remainder of the paper. However, the evidence is also consistent with equal non-welfarist preferences across goods in Experiment 3.

Altogether, non-welfarist concerns are prevalent yet also show meaningful domain-specificity. 73% of Spectators who do experiments with rights goods have positive WTP in either Experiments 1 or 2. 31% have positive WTP in *both* experiments, a sizable 12 pp (65%) increase off the 19 pp who have positive WTP in both experiments with benchmarks ($p < 0.001$).

Having found evidence of non-welfarist preferences, we next consider their implications for Spectators' redistributive choices and SWFs' incidence (Section 4). When thinking about SWFs, we view the Spectators as the people in society whose decisions SWFs aggregate; Spectators are not social planners themselves.

In Experiment 4, we show how non-welfarist preferences influence the progressivity of Social Welfare Functions. We design a novel redistribution task where Spectators allocate goods among 10 anonymous Recipients with varying need, as indicated by their incomes. Spectators choose between giving the good to everyone, or the good plus cash to people with lower incomes. Spectators express more or less progressivity by how much they favor the lowest-income Recipients.

Spectators choose less progressive, more "anti-targeted," provision of rights goods relative to benchmarks. Spectators are 64% (s.e.: 8.9%) more likely to provide rights goods to all 10 Recipients than when doing the task with cash or benchmarks. The experiment thus gives new evidence of domain-specific redistributive preferences.

Non-welfarist preferences explain Spectators' reduced progressivity when allocating rights goods. Spectators' non-welfarist behaviors in Experiments 1 and 2 explain more than half the gap in progressivity between rights and benchmark goods in Experiment 4. Non-welfarist preferences in Experiments 1 and 2 also correlate with stated support for public transfers in these and other domains. In sum, non-welfarist preferences push toward more universal, less progressive provision of these transfers — a practical policy take-away, and an economic case for universal provision that complements standard rationales like market failures or paternalism.

A second policy implication concerns the incidence of non-welfarist preferences. If the rich value abstract concerns like rights but the poor value practical welfarist issues, then prioritizing non-welfarist concerns would be regressive. We find inverted U-shaped incidence, as non-welfarist preferences are 42–65% more common in middle-income versus low- or high-income Spectators. Simulations calibrated to these findings suggest a policy trade-off: SWFs that promote non-welfarist concerns must sacrifice progressivity to down-weight poor Spectators' welfarist views.

We conclude by studying mechanisms that activate non-welfarist concerns (Section 5), toward a better understanding of the conditions under which such preferences matter most. We focus

on the decision’s welfarist consequences (stakes). One hypothesis is that non-welfarist concerns only emerge in small-stakes situations. Alternatively, larger welfarist stakes could force reflection and direct attention to non-welfarist concerns. In experiments with 6,000 new Spectators, we vary stakes in the context of Experiment 1’s provision of lawyers. Across treatments, non-welfarist concerns rise with larger stakes: non-welfarist WTPs are 78% to 124% higher when providing more impactful lawyers and 16% higher when Recipients are needier ($p \leq 0.001$ for each). Non-welfarist concerns emerge when decisions are consequential, underscoring their relevance.

Contribution. This paper stakes a compromise position in applied welfare economics. We are of course unlikely to settle the longstanding debate between welfarists and non-welfarists. Indeed, one can construe the results as supporting either point of view. Our findings buttress welfarist frameworks, showing they are compatible with real-world moral considerations. Yet, the results likewise support non-welfarist frameworks, as those frameworks become more compelling if non-welfarist concerns are indeed prevalent in society. However one interprets the findings philosophically, they have important implications for practice: they argue against the prevailing approach of ignoring non-welfarist concerns in welfare assessment.

We bridge the literature using survey experiments to explore how people think about policies, with research detecting policy preferences using more laboratory-experimental manipulations. We contribute to the economics literature on moral preferences (Rabin, 1993; Fehr and Schmidt, 2006; Dana et al., 2007; Bénabou and Tirole, 2011; Alger and Weibull, 2013; Enke, 2024; Saccardo and Serra-Garcia, 2023), documenting non-welfarist concerns in important policy settings along with meaningful domain-specificity.³ We build on the experimental literature on redistributive decisions (Levitt and List, 2007; List, 2007; Alesina and Giuliano, 2011; Cappelen et al., 2013, 2020; Capozza and Srinivasan, 2024; Exley and Kessler, 2024), as well as distinguished literatures in empirical ethics (Appiah, 2010; Knobe et al., 2012) and moral psychology (Doris et al., 2006).^{4,5} We particularly add to the strands of the economics literature detecting non-instrumental morals (e.g., Abeler et al., 2019; Chen and Schonger, 2022; Sullivan, 2023; Bénabou et al., 2024) or non-welfarist principles in taxation (Konow, 2001; Cappelen et al., 2010; Sheffrin, 2013; Weinzierl, 2014, 2017; Saez and Stantcheva, 2016; Fleurbaey and Maniquet, 2018; Stantcheva, 2021).

Relative to this work, we test for non-welfarist concerns about government transfers in policy-relevant areas where isolating the nature of moral concerns is key.⁶ Specifically, we conduct new

³Non-welfarist concerns strongly correlate with preferences for redistribution, complementing other predictors in the literature (e.g., Luttmer and Singhal, 2011; Alesina et al., 2018; Roth and Wohlfart, 2018).

⁴Examples of moral concerns range from preferences not to lie (e.g., Fischbacher and Föllmi-Heusi, 2013; Gneezy et al., 2018; Abeler et al., 2019) or steal (e.g., Cohn et al., 2019), to religion and debt (e.g., Bursztyn et al., 2019), to voting patterns (e.g., Enke, 2020), among many others.

⁵Like Fisman et al. (2007), we study how bystanders trade off efficiency for redistributive considerations, but focus on willingness to pay for satisfying moral concerns rather than altruistic giving. Like Alatas et al. (2012), we quantify how people target, but document a different phenomenon — anti-targeting — and propose explanations. Like Charité et al. (2022), we consider how non-classical forces affect Spectators’ choices for others.

⁶As non-welfarist concerns imply different motives for providing certain goods, beyond their instrumental effects.

experiments testing for non-welfarist principles for in-kind transfers (Experiments 1–3); show how these concerns relate to redistributive preferences (Experiment 4); and manipulate goods to isolate domain-specificity and rule out elicitation errors as a complete explanation for non-welfarism.^{7,8} Throughout, we complement experimental outcomes using Spectators’ choices with open-ended text responses, which confirm our interpretation of participants’ motives (Haaland et al., 2024).

We additionally contribute to the economics of government transfers (Currie and Gahvari, 2008), examining the non-welfarist rationales that economists often overlook. We relate to Liscow and Pershing (2022)’s experiments that decompose hypothetical demand for in-kind redistribution. Relative to this work, we conduct experiments that quantify non-welfarist concerns without relying on stated attitudes, and we show how these preferences relate to redistributive decisions.⁹

1 Conceptual Framework

Our experiments are concerned with the existence and implications of non-welfarist preferences. To motivate this work, we begin by formalizing when the presence of non-welfarist preferences among individuals in society can lead welfarist Social Welfare Functions to behave like non-welfarist Social Welfare Functions. Then, we propose a simple functional form for these preferences, which yields predictions that we test in experiments. Appendix C provides details and proofs.

1.1 Take-Aways for Empirical Work

Readers who wish to skip the details can read these take-aways and proceed directly to Section 2.

1. The Existence of Non-Welfarist Preferences Is Relevant to Welfarists. Proposition 1 clarifies the conceptual distinction between welfarist and non-welfarist Social Welfare Functions (SWFs). When people in society have preferences that are themselves non-welfarist and “distinct enough,” there exists a *welfarist* SWF that behaves exactly like a given *non-welfarist* SWF. The idea is simple.

Our focus on affirmative rights to transfers also distinguishes us from the rich literature, reviewed by Suzumura (2011), which explores the relationship between welfarism and libertarian rights to avoid interference from others. Our conception of rights, where people intrinsically care about providing goods to others, has implications for redistributive policy and embeds rights within even welfarist frameworks. This discussion is one example of how the distinction between “negative” and “positive” rights can become blurred (Holmes and Sunstein, 2000): the right to make choices is central to libertarian rights (Experiment 2), but also has implications for redistribution (Experiment 4).

⁷We build off Polman (2012), who studies demand to change initial allocations for others, but who cannot identify non-welfarist concerns. First, in some of Polman (2012)’s experiments, the recipients know the initial allocation, so the Spectators may be welfarists who aggregate others’ loss aversion. Second, several experiments are thought experiments as they involve WTP for non-quantifiable outcomes (e.g., “ability to get dates”). Third, several experiments measure WTP to improve an outcome versus stop the outcome from getting worse. If Spectators perceive Recipients as having concave utility, then higher WTP to avoid losses is consistent with welfarist preferences.

⁸Related to Experiment 2, Bartling et al. (2014), Bobadilla-Suarez et al. (2017), and Freundt et al. (2023) find subjects value choice for themselves, but do not study how subjects value choice for others. Andreoni et al. (2020) study fairness in the presence of uncertainty and find subjects apply deontological principles when allocating lottery tickets.

⁹One of Liscow and Pershing (2022)’s treatments makes the right to in-kind goods salient and studies hypothetical support for redistribution. They find no impact of the salience framing. This null result of salience could arise because participants already have non-welfarist preferences in both treatment and control, consistent with our findings.

Welfarists just aggregate up people’s preferences. So, with suitable welfare weights and sufficiently varied preferences, welfarist SWFs can mimic non-welfarist SWFs.

Thus, the existence of non-welfarist concerns is informative about whether non-welfarist SWFs are necessary to capture moral concerns. To claim they are necessary, one must argue that the non-welfarist part of the SWF is *not* shared by anyone in society.

2. The Incidence of Non-Welfarist Preferences Matters. Proposition 1 highlights that non-welfarist SWFs may be equivalent to welfarist SWFs that are regressive or unfair. For instance, if only rich people have non-welfarist preferences, then welfarists’ SWFs can indeed behave like non-welfarists’ — albeit, by up-weighting the rich. This point motivates us to study the empirical incidence of non-welfarist preferences.

3. Non-Welfarists Favor Anti-Targeting. Proposition 2 shows that non-welfarists choose less progressive (more anti-targeted) allocations. This empirical prediction provides a link between Experiments 1–3, which test for the existence of non-welfarist preferences, and Experiment 4, which tests for anti-targeting.

1.2 When Are Non-Welfarism and Welfarism Distinct?

Notation and Definitions. Suppose individual in society (“Spectator”) $i \in \mathcal{I}$ has preferences for allocating goods $y \equiv (y_1, \dots, y_J) \in \mathbb{R}^J$ to Recipients $j \in \mathcal{J}$ as follows:

$$v_i(x, y; \vec{u}) = f_i(u_1(x_1, y_1), \dots, u_J(x_J, y_J)), \quad (1)$$

where $f_i : \mathbb{R}^J \rightarrow \mathbb{R}$ aggregates others’ utilities $u_j(\cdot)$, $|\mathcal{I}| = I$, $|\mathcal{J}| = J$, and $x \equiv (x_1, \dots, x_J) \in \mathbb{R}^J$ captures other aspects of utility or endowments (e.g., income or other consumption besides y). We denote the vector of Recipients’ utility functions as $\vec{u} \equiv (u_j(x, y))_{j \in \mathcal{J}}$. We consider finite \mathcal{I} and \mathcal{J} , though ideas easily extend.

In Equation (1), i ’s preferences depend only on the utilities of other people in society. This benchmark captures utilitarian, altruistic, Rawlsian (maximin), inequity averse, paternalistic, and many other commonly used social preferences in economics.¹⁰ In all these cases, Spectators want to allocate goods to maximize a function of Recipients’ utilities. For example, utilitarian preferences are $v_i = \sum_j u_j$.

Spectators are people in society, and in the experiments, correspond to the Spectators who make choices over Recipients’ allocations. Spectators are not Social Welfare Functions (SWFs), which instead aggregate Spectators’ preferences. This convention departs from some prior work in which experimenters have treated Spectators as the social planners themselves. However, while

¹⁰Paternalism is not welfarist in cases where Spectators do not respect Recipients’ preferences (e.g., some flavors of the ideals-projective paternalism in Ambuehl et al., 2021).

v_i is an individual preference, if we delete the i subscript, Equation (1) shows a welfarist SWF.

To link to our experiments, the above notation emphasizes cases where Spectators allocate goods to others; i.e., Spectators are not Recipients, and \mathcal{I} and \mathcal{J} are disjoint. However, Spectators *can* also be Recipients, if f_i includes the standard consumption utility i gets from the good.¹¹

Now augment individual i 's preferences:

$$v_i(x, y; \vec{u}) = \hat{f}_i(u_1(x_1, y_1), \dots, u_J(x_J, y_J), \mathbf{x}, \mathbf{y}), \quad (2)$$

for $\hat{f}_i : \mathbb{R}^{3J} \rightarrow \mathbb{R}$ where the independent dependence on \mathbf{x} and \mathbf{y} captures how i may have non-instrumental preferences to change x and y . Then, allocating y can affect i 's utility, fixing how y affects j 's utility. If we delete the i subscript from Equation (2), the expression shows a non-welfarist Social Welfare Function.

Definition 1. We say i is *non-welfarist* if there exists a vector of utility functions \vec{u} and bundles x, y, x', y' such that $u_j(x, y) = u_j(x', y')$ for all j , but $v_i(x, y; \vec{u}) \neq v_i(x', y'; \vec{u})$. Otherwise, i is *welfarist*.

That is, a Spectator is non-welfarist if x and y can affect i 's preferences, fixing utility in society.¹²

Social Welfare Functions. A welfarist Social Welfare Function (SWF) aggregates Spectators' preferences:

$$W_{\text{welfarist}} = F(v_1(x, y; \vec{u}), \dots, v_I(x, y; \vec{u})), \quad (3)$$

for $F : \mathbb{R}^I \rightarrow \mathbb{R}$. Non-welfarist Social Welfare Functions may independently depend on x and y :

$$W_{\text{non-welfarist}} = \hat{F}(v_1(x, y; \vec{u}), \dots, v_I(x, y; \vec{u}), x, y), \quad (4)$$

for $\hat{F} : \mathbb{R}^{I+2} \rightarrow \mathbb{R}$. By analogy to Definition 1, an SWF is non-welfarist if there exist x, y, x', y' such that $v_i(x, y) = v_i(x', y')$ for all i , but $W(x, y) \neq W(x', y')$. Otherwise, the SWF is welfarist.

When Can Welfarism Replicate Non-Welfarism? A simple way of incorporating non-welfarist concerns into a fully welfarist SWF is to put non-welfarist concerns into people's utility functions directly, as Kaplow and Shavell (2001) (p. 285), Fleurbaey et al. (2003) (footnote 4), and others note. However, it remains unclear when *arbitrary* non-welfarist SWFs can be replicated by welfarist SWFs. We therefore identify necessary and sufficient conditions on Spectators' utilities such that a carefully chosen welfarist SWF exactly replicates the non-welfarist SWF.

¹¹For classical Spectators, f_i is the identity function over the element corresponding to i 's utility and zero over other elements. We index both \mathcal{I} and \mathcal{J} starting at 1, so the i th Spectator could be the j th Recipient. We prevent circularity by imposing that the u_j corresponding to i 's private utility is only her private utility, although other u_j 's include all social preferences for the other people in society. That is, f_i is not a function of f_i itself.

¹²This definition comports with Kaplow and Shavell (2001), but other definitions impose an independence principle (d'Aspremont and Gevers, 2002; Fleurbaey et al., 2003; Fleurbaey and Maniquet, 2018).

Definition 2. Fixing a vector of Spectator utility functions $\vec{v} \equiv (v_i(\cdot))_{i \in \mathcal{I}}$, a welfarist Social Welfare Function F replicates a non-welfarist Social Welfare Function \hat{F} if and only if for all x and $y \in \mathbb{R}^J$,

$$\hat{F}(v_1(x, y; \vec{u}), \dots, v_I(x, y; \vec{u}), x, y) = F(v_1(x, y; \vec{u}), \dots, v_I(x, y; \vec{u})). \quad (5)$$

That is, the replicating SWF is chosen to remove the non-welfarist SWF's separate dependence on x and y . The non-welfarist SWF can only be replicated for a given utility vector \vec{v} . If the non-welfarist SWF could be replicated for all possible \vec{v} , then it would in fact be a welfarist SWF.

A basic example shows how the welfarist might replicate the non-welfarist, when one Spectator has a utility function that copies the non-welfarist part of the SWF:

Example 1. Suppose the SWF is:

$$\hat{F}(v_1(y), \dots, v_I(y), y) = \sum_{i=1}^I v_i(y) + \mathbb{1}(y_i \geq 1 \text{ for all } i). \quad (6)$$

The SWF is non-welfarist because it depends on y_i directly, so there exist utility functions where the SWF gets utility from giving y_i but individual preferences are constant. However, suppose Spectator 1's utility is $v_1(x, y) = \mathbb{1}(y_i \geq 1 \text{ for all } i)$. Then, there exists a welfarist Social Welfare Function that replicates the non-welfarist SWF:

$$F(v_1, \dots, v_I) = 2v_1 + \sum_{i=2}^I v_i. \quad (7)$$

Intuitively, if Spectator 1 has the right non-welfarist preferences, then a purely welfarist social planner copies the non-welfarist SWF by up-weighting Spectator 1's preferences.

We now generalize the example.

Definition 3. A function $\rho(x, y)$ distinguishes the utility vector \vec{v} if the following condition is satisfied: for all x, x', y , and $y' \in \mathbb{R}^J$, whenever $(v_1(x, y), \dots, v_I(x, y)) = (v_1(x', y'), \dots, v_I(x', y'))$, then $\rho(x, y) = \rho(x', y')$. We also call $\rho(x, y)$ a distinguishing function.

Proposition 1. For any non-welfarist Social Welfare Function \hat{F} and utility vector \vec{v} , there exists a welfarist Social Welfare Function that replicates \hat{F} if and only if there exists a distinguishing function $\rho(x, y)$ and Social Welfare Function \tilde{F} such that $\hat{F}(v_1, \dots, v_I, x, y) = \tilde{F}(v_1, \dots, v_I, \rho(x, y))$.

Proposition 1 formalizes the intuition that a welfarist SWF can mimic an arbitrary non-welfarist SWF, if Spectators' utilities proxy for non-welfarist concerns. The distinguishing condition on ρ gives the welfarist SWF exactly the right information to reproduce the mapping between the non-welfarist SWF's separate dependence on (x, y) and Spectators' utility. That is, the welfarist can copy the non-welfarist, by up-weighting utilities that covary with \hat{F} in the same way as moving (x, y) independently. Another intuition is that $\rho(x, y)$ parameterizes the non-welfarist part of the

SWF, captured in \tilde{F} . If $\rho(x, y)$ is sufficiently low-dimensional, then utilities in society can proxy it. Conversely, no welfarist SWF can replicate a non-welfarist SWF when the utility vector is not “diverse enough” to capture the non-welfarist part of the SWF in ρ .

Two standard settings satisfy the distinguishing condition, and thus guarantee that suitably chosen welfarist SWFs replicate non-welfarist SWFs. First, generalizing Example 1, if a single “virtuous” Spectator has non-welfarist utility exactly like the SWF.¹³ Second, if Spectators’ preferences are distinct enough that every allocation (x, y) induces unique utilities $(v_1(x, y), \dots, v_I(x, y))$.¹⁴

In practice, how can non-welfarist preferences capture real-world non-welfarist concerns? Consider the idea of a basic right to health care, separate from how health care affects utility. Or, as in Saez and Stantcheva (2016), the idea that society wants to redistribute to the “deserving” but not “undeserving” poor. Proposition 1 says that as long as one person in society has these preferences, a purely welfarist SWF can capture such concerns.

In fact, the result is actually much stronger, as a society with purely welfarist or even selfish Spectators can still replicate a non-welfarist SWF. That is, having non-welfarists in society is sufficient but not necessary to replicate the non-welfarist SWF. Returning to Example 1, imagine each i is completely selfish with utility given by $v_i = y_i$. Then, a welfarist SWF that replicates \hat{F} is

$$F(v_1, \dots, v_I) = \sum_{i=1}^I v_i + \mathbb{1}(v_i \geq 1). \quad (8)$$

The Proposition reflects this, as we can define $\rho(x, y) = \mathbb{1}(y_i \geq 1 \text{ for all } i)$, and $\rho = 1$ and $\rho = 0$ correspond to distinct utility vectors.

Relation to the Literature. Saez and Stantcheva (2016) propose generalized social marginal welfare weights, which result in non-welfarist SWFs. In contrast, Proposition 1 shows when putting non-welfarism into the *utility function* yields “as-if non-welfarism,” even though the SWF and welfare weights are completely welfarist. The Proposition thus sidesteps Sher (2024)’s objections about SWF cycles in generalized social marginal welfare weights, and Kaplow and Shavell (2001)’s point that non-welfarism violates Pareto. For this reason, our approach reconciles moral concerns with policy evaluation, while avoiding some theoretical pitfalls.

The advantages of this strategy do not come for free. Welfarist SWFs and suitable weights cannot replicate any possible non-welfarist SWF. This as-if non-welfarism remains disciplined by whether utilities are distinct enough to copy the non-welfarist SWF. If not, then some non-welfarist

¹³Or, generalizing this case, if $\rho(x, y) = g(\bar{v}(x, y))$ for some aggregator $g(\cdot)$. Another example is if one or a group of “evil” Spectators proxy for the non-welfarist part of the SWF via $\rho(x, y) = -g(\bar{v}(x, y))$. In the latter case, that would imply the replicating SWF is not an increasing function of Spectators’ utilities.

¹⁴The more diverse are Spectators’ utilities, the more non-welfarist SWFs the welfarist can replicate. Each vector of utilities \bar{v} induces a collection $\mathcal{R} = \{\rho(x, y) \mid \bar{v}\}$ of functions $\rho(x, y)$ that distinguish utility vector \bar{v} . The welfarist replicates any non-welfarist SWF characterized by a given $\rho(x, y) \in \mathcal{R}$.

SWFs are not replicated. This is an empirical question, motivating Experiments 1–3.¹⁵

Proposition 1 generalizes An (2024), who gives a special example of a parametric non-welfarist SWF that can be recast as a welfarist SWF and similar intuition. Moreover, the idea of a replicating SWF relates to research that infers SWFs from which policies get implemented (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016; Hendren, 2020; Björkegren et al., 2022).

Incidence. Proposition 1 mimics non-welfarist SWFs by up- or down-weighting preferences of people in society. There is no guarantee that the resulting weights are progressive or even reasonable. For instance, Example 1 placed double weight on Spectator 1. If Spectator 1 is rich, the replicating SWF sacrifices progressivity to accommodate non-welfarist preferences. In our empirical settings, it could be that only rich people care about the right to health care, whereas poor people just want more consumption. Hence we examine incidence empirically.

Domain-Specificity. A key point of the paper is that non-welfarist preferences could be domain-specific, differing by good (e.g., the idea in *Spheres of Justice*). The notation implies that the good y_j is one-dimensional for simplicity. Ideas generalize if each y_j is itself a vector.

1.3 Anti-Targeting

We now propose a specific empirical prediction. Under natural functional forms, we show how non-welfarist concerns reduce progressivity of Spectators' allocation decisions.

Benchmark. We specialize the environment. Impose homogeneous, selfish Recipients' utilities u_j . Consider a welfarist Spectator i , whose utility is separable in Recipients':

$$v_i(x, y; \gamma) = \sum_{j=1}^J \gamma_j u(x_j, y_j) \quad (9)$$

for exogenous weights $\gamma \equiv (\gamma_1, \dots, \gamma_J)$.

Consider allocating m indivisible goods among the population of J Recipients. Fixing the weights, optimal allocation $\{y_j^*\}$ solves

$$\max \sum_{j=1}^J \gamma_j u(x_j, y_j^*), \text{ such that } \sum_j y_j^* \leq m. \quad (10)$$

In this benchmark, the Spectator gives the goods to the $k \leq m$ people where the social marginal welfare gains are largest.¹⁶ If y is not too complementary with x , then the social marginal welfare

¹⁵If one normatively respects other-regarding preferences, certain paradoxes emerge, e.g., when other-regarding preferences conflict (Fleurbaey, 2019). Proposition 1 shows when a targeted non-welfarist SWF can be replicated. The individual utilities may be weighted in counterintuitive ways, but the utilities do still aggregate to a well-posed SWF.

¹⁶The Spectator can give multiple of the good to one Recipient. The Spectator provides goods such that, for all $j, k \leq J$, $\gamma_j (u(x_j, y_j^*) - u(x_j, y_j^* - 1)) \geq \gamma_k (u(x_k, y_k^* + 1) - u(x_k, y_k^*))$. Similar results obtain with continuous goods.

gain is maximized by providing to those with smaller endowments and higher weights.

Introducing Reference Points. Turning to Spectators with non-welfarist preferences, we parameterize non-welfarist concerns as entering utility via an aggregator $\mu_i : \mathbb{R}^3 \rightarrow \mathbb{R}$. We posit they follow a form of Kőszegi and Rabin (2006)-type reference-dependence:

$$v_i(x, y; \gamma, r) = \sum_{j=1}^J \gamma_j (u(x_j, y_j) + \eta_y \mu_i(x_j, y_j \mid r_{ij})) \quad (11)$$

$$\mu_i(y_j \mid r_{ij}) = \begin{cases} u(x_j, y_j) - u(x_j, r_{ij}) & \text{if } y_j > r_{ij} \\ \lambda [u(x_j, y_j) - u(x_j, r_{ij})] & \text{if } y_j \leq r_{ij} \end{cases} \quad (12)$$

Here, $r_i \equiv (r_{i1}, \dots, r_{iJ})$ corresponds to Spectator i 's reference points for allocating each $y_j \in y$. The subscript i emphasizes that r_i is Spectator i 's reference point for allocating to another person j , and not a reference point that enters j 's utility. If η_y and $\lambda > 1$, the Spectator i experiences loss aversion over Recipient j 's utility relative to i 's reference point on good y . This formulation remains agnostic about whether u_j is itself reference-dependent. The case with $\eta_y = 0$ nests welfarist allocative utility.

Why Reference-Dependence? Although admittedly special, the formulation has several virtues. First, it is psychologically natural if non-welfarist concerns involve meeting basic rights. Second, it links to a political psychology literature on political preferences, which often involve heuristics and reference points (Ottati, 1990; Druckman and Lupia, 2000). Finally, it clarifies the empirics, especially Experiment 1, where we manipulate reference points while fixing Recipients' utilities.

Importantly, no aspect of our experiments requires reference-dependence for the study to be valid. Rather, reference-dependence gives an explanation for a correlation between behaviors in Experiments 1–3 and Experiment 4.

Targeting With Reference Points. We now derive the prediction. Introduce a notion of flatness:

Definition 4. Let \mathcal{J}_i^* represent the set of j such that optimal $y_j^* > 0$ under utility v_i . Allocation \mathcal{J}_2^* is *weakly flatter* than \mathcal{J}_1^* if and only if $\mathcal{J}_1^* \subseteq \mathcal{J}_2^*$.

In words, a set of Recipients \mathcal{J}_2^* is weakly flatter than \mathcal{J}_1^* if all individuals who get at least one good under \mathcal{J}_1^* also get at least one under \mathcal{J}_2^* . There may be someone in \mathcal{J}_2^* who is not in \mathcal{J}_1^* .

We compare i 's optimal allocations with and without non-welfarist concerns (Equation 9 versus 11). Fixing i and suppressing subscripts, put the optimal set of Recipients $\tilde{\mathcal{J}}$ as the one that maximizes i 's utility for a given set of welfare weights, utility functions, and endowments and places no value on non-welfarist preferences (that is, $\eta_y = 0$). Put the optimal set of Recipients \mathcal{J}^* as the one that maximizes i 's utility for the same set of welfare weights, utility functions, and endowments but which also places value on non-welfarist preferences with $r_{ij} = 1$ for all j , $\lambda > 1$, and $\eta_y > 0$. Then the Spectator with non-welfarist preferences exhibits anti-targeting:

Proposition 2 (Anti-Targeting Prediction). *The set of Recipients with non-welfarist preferences \mathcal{J}^* is weakly flatter than the set of Recipients without non-welfarist preferences $\tilde{\mathcal{J}}$.*

Without reference points, the Spectator maximizes welfare by allocating goods to those with the highest marginal utility, optionally weighted by social welfare weights. Absent too much complementarity between the good and the endowment, diminishing marginal utility over wealth causes targeting toward those with low endowments. If the Spectator redistributes relative to a reference point, that can reduce this targeting motive. Allocating to any individual in the loss domain achieves a higher marginal welfare gain than allocating to those already above the reference point. Loss aversion gives a kink in welfare around the reference point which pushes toward more universal provision.

We have derived this Proposition as predicting anti-targeting among Spectators. It likewise implies anti-targeting among welfarist SWFs that aggregate Spectators’ utilities.

Connection to Rights. We view rights as mapping onto reference points that Spectators or social planners might value, as non-welfarist preferences. Experiments 1–3 test for non-welfarist preferences, that is, $H_0 : \eta = 0$ for a particular r in the experiment. Experiment 4 tests the model prediction in Proposition 2. The reference points in our model could relate to any feature of provision. Some may relate to rights. Others could relate to concerns about procedural justice.

2 Experiment Overview

2.1 Sample and Design Overview

Sample and Design Overview. We recruit $N = 1,800$ participants from Prolific, a widely used online platform for survey experiments (Appendix D.1). As is common in Prolific studies, participants are higher-income, younger, and more educated than in the U.S. in general (Table 1).

Figure 1 presents the experiment flow. We randomize participants (“Spectators”) into one of four goods at the start. They then complete four experiments, each with the same good, before answering several questions about political preferences.¹⁷

To be included in the study, participants needed to pass at least two out of three attention checks throughout the survey (Appendix D.4). 7% fail one of three and are included in the sample. We routinely include comprehension checks after task instructions but before the elicitation. Participants passed these checks at rates exceeding 85%. We correct participants who fail.

We pre-registered the experiment before launch (AEA Registry #0012065) and ran the study on September 11–12, 2023. Appendix F gives a key between the pre-registered analysis and exhibits in the paper. Experiment instructions and screenshots are posted on the authors’ websites.

¹⁷Half the benchmark good participants were randomized into doing Experiment 4 with cash instead.

The survey took 19 minutes on average. We paid \$6 for participating, 56% above Prolific's suggested wage for a 19-minute survey.

Setup and Goods. We inform participants early in the survey that they will make allocation choices on behalf of a nonprofit in Memphis, Tennessee, which assists tenants facing eviction. All participants are informed that the clients are those facing eviction. We make this choice so that all between-good comparisons hold fixed the need and financial situation of the Recipients.

The four goods are: attorneys, who can provide legal assistance to tenants facing eviction; one year of fully subsidized health care at urgent care; a bus pass containing \$350 of prepaid fare; and an annual membership at the local YMCA, which can provide child care and wellness services. All goods except health care can be purchased for about \$350. Our main tests compare "rights goods" (attorneys and health care) to "benchmark goods" (bus and YMCA).

The choice of goods is important but challenging. There are many potential rights goods and comparison goods. Once a right legally exists, the experimenter cannot easily manipulate endowing the right in the lab. Even if the experiment involves hypothetical choices, if the right has no analog in the market, participants may have difficulty forming views about willingness to pay for the right. For instance, even a hypothetical choice about willingness to pay for a lawyer may be more ecologically valid than a hypothetical choice about willingness to pay for free speech.

We pick rights goods relevant to live policy debates about government transfers and where appeals to rights are common in the public discourse. Health care is the subject of among the most contentious debates about the U.S. safety net. Legal services for eviction are at the frontier of rights expansion. Evictions are civil cases and hence not covered by 6th Amendment right to counsel. Many states and cities recently implemented policies to give tenants access to lawyers in these cases, and the policy discussion heavily relies on moral appeals (see footnote 1). We pick benchmark goods that are clearly valuable to low-income Recipients, but to which Spectators are unlikely to attach special rights.

We focus on both the levels of non-welfarist behaviors with right and benchmark goods, as well differences between rights versus benchmarks. The levels are informative about the prevalence of non-welfarism. The differences speak to the heterogeneity suggested by works like *Spheres of Justice* and policymakers' appeals to rights in some domains but not others.

Focusing on differences between goods also nets out potential elicitation issues. Otherwise, levels of behaviors may conflate inattention, confusion, or trembles (e.g., mistaken clicking that generates positive willingness to pay) with non-welfarism. Examining differences between goods also addresses possible experimenter demand effects, unless they interact with good (Bénabou et al., 2023). Any story about why the large WTPs we find are not evidence of non-welfarism must explain why they are systematically larger among rights goods versus benchmarks.

Under the assumption that non-welfarist preferences exist but are weaker for benchmark

goods than for rights goods, comparing rights to benchmark goods leads us to understate the extent of non-welfarist preferences. This assumption is reasonable. At one extreme, Spectators could value the right to property, and therefore exhibit non-welfarist preferences for *any* potential benchmark good. In that case, the comparison between rights and benchmark goods might not be distinguishable from zero. Less extreme versions of this preference, in which Spectators value the right to transit or the right to exercise but not as much as the right to lawyers or health care, attenuate our results but not to zero.

Main and Secondary Elicitations. Our main sample participated in four primary experiments, testing for features of rights (Experiments 1–3, Section 3) and redistributive preferences (Experiment 4, Section 4). We also present two secondary experiments and elicitations (pre-registered as such). The first directly elicits (bounds on) the Spectator’s indifference point between providing the good and giving cash to an anonymous Recipient, using a multiple price list and the strategy method. We interpret this elicitation as a Spectator’s willingness to pay WTP_i for the good and use this value in extra tests throughout. The second is an information-provision sub-experiment that tests for welfarism (discussed briefly in Section 4, with details in Appendix D). Finally, we conduct a separate mechanisms experiment (Section 5).

We need to do several experiments within-subject to study correlations between non-welfarism and redistributive preferences. Spectators do Experiments 1 and 3 in a random order, 2 next, and 4 last. An important concern is that instructions in one experiment could influence responses for subsequent experiments. Between-good differences net out most such issues. Further mitigating this objection, the most important priming concern we had was that Experiment 2 tells participants that most tenants choose cash over a good (see Section 3.2). For this reason, participants always do Experiment 2 after Experiments 1 and 3. Participants always do Experiment 4 last, since we then include an information treatment, and we do not want this treatment to influence the main experiments. One limitation of the order we choose is that, since we elicit the direct WTP to provide the good last, this elicitation is most likely to be influenced by previous instructions.

Experiments 1–3 measure willingness to pay (WTP) to take some action. We measure WTP using a separate donation budget. That is, WTP is how much money the Spectator would pay to take some action, where the money comes from a budget that is otherwise spent on other programming for needy households. In Experiments 1–2, the donation budget is for future programming at the nonprofit. In Experiment 3, the donation budget is for donations to a food bank. Where exactly the budget would be donated is not conceptually important, except that we require wasting the donation budget to be costly to Spectators, e.g., due to altruism. If not, their choices reduce to noise. We reject this concern using our conservative comparison of rights goods and benchmarks, text analysis of Spectators’ stated explanations for their choices, and other tests.

Ethics. Spectators provide consent and face no risk of harm. All goods or cash allocated only

have the potential to help Recipients. All Recipients are needy, and there is not enough funding to assist everyone. Therefore, providing goods based on Spectator choices is a reasonable way of targeting. Indeed, the allocation choices in this study are similar to those in real-world political decisions about means testing, as well as other studies on how experiment participants allocate scarce resources (Alatas et al., 2012; Sullivan, 2023).

2.2 Incentives

A feature of this paper is that we incentivize choices for all goods except health care. We use the strategy method, informing Spectators that there is a chance their choices will be implemented. To implement choices, we implement choices over giving legal assistance, YMCA memberships, bus passes, or cash via a nonprofit partner in Memphis (The Works, Inc.). Participants are informed and must confirm that they face real choices which could affect allocations for needy Recipients. We introduce another incentive by telling all participants (including in the health-care treatment) that the study results could influence the nonprofit’s future programming. We incentivize belief elicitation by paying participants if they are accurate. See Section 4.5 and Appendix D for details on incentives.¹⁸

We embedded several tests to probe whether lack of incentives affects results for health care. We find no evidence that incentives affect Spectators’ choices (Section 4.5). That said, we still view incentivization as an important part of this paper. It was unclear that incentives would have small effects *ex ante*, and some readers may (reasonably) have been skeptical of results if all elicitation were hypothetical (see discussion in Sullivan, 2023).

2.3 Specification, Balance, and Attrition

We estimate the models:

$$y_i = \beta_0 + \beta_{\text{Right}_i} (+X_i\delta) + \varepsilon_i \quad (13)$$

$$y_i = \beta_0 + \beta_l \text{Lawyer}_i + \beta_h \text{HealthCare}_i (+X_i\delta) + \varepsilon_i \quad (14)$$

where β is the effect of being a rights good on an outcome y_i in Equation (13), and β_l and β_h are the effects for lawyers and health care respectively in Equation (14). We pool both benchmarks for power, and we show they are generally similar in the raw data. Our main specifications omit controls X_i and compare raw means between rights goods and benchmarks, but robustness checks include them. Inference uses robust standard errors.

Demographics are balanced across rights treatments versus benchmarks (Table 1, joint $p = 0.441$; Table A1 disaggregates by good). Attrition rates were 4% (Table A2).

¹⁸Incentives are already implemented or are in the process of being given to Recipients.

3 Existence and Magnitude of Non-Welfarist Preferences

We present the design and results for Experiments 1–3 in turn. Along the way, we highlight how each experiment makes an independent contribution to our understanding of Spectators’ non-welfarist principles. We include some robustness checks for Experiments 1–3 with those for Experiment 4 in Section 4.5.

3.1 Experiment 1: Inalienability

3.1.1 Design

Motivation. Suppose a good first is assigned to one person, then is transferred to an *ex ante* identical person. Suppose also that neither person is aware of the initial assignment. Welfarists care only about the final allocation, and do not care about the initial assignment. Non-welfarists may dislike transferring goods from one person to another because it requires removing the good from someone who has, unbeknownst to them, received it. Preferences not to reallocate the good may be especially prevalent if Spectators believe Recipients have inalienable rights to the good.

Design. Experiment 1 hews closely to the motivation. We tell Spectators that Recipient B was assigned a good in a lottery. Spectators have the choice of rerunning the lottery, which has the chance of taking the good from Recipient B and giving it to Recipient A. If they rerun the lottery, the Spectator saves \$ x , which the non-profit will use for future programming. We tell the Spectator that the money saved will assist other tenants. We tell the Spectator that \$ x can be saved because prices for the good may sometimes change after running the initial lottery.

We use a multiple price list to find (bounds on) the point at which Spectators are indifferent between saving \$ x and preserving the lottery result, enforcing monotonicity (see screenshots on authors’ websites). We refer to I_i , the midpoint of the elicited bounds, as willingness to pay (WTP) to preserve the lottery. If $I_i > 0$, the Spectator burns I_i to satisfy non-welfarist preferences. As I_i ’s units are unintuitive, we often normalize I_i so that pooled benchmarks have mean 0 and standard deviation 1. We also study the propensity to have positive WTP, forming $\mathbb{1}(I_i > 0)$ with unnormalized I_i , and the propensity to have the maximum WTP that we elicit.

Design Choices Rule Out Welfarist Reasons to Pay To Preserve The Lottery. We ensure Spectators do not believe Recipient A is needier or would derive greater benefit from the lottery. Otherwise, even welfarists might rerun the lottery. The experiment informs Spectators that the Recipients are *ex ante* identical: “Both tenants had evictions filed against them, applied around the same time, and are similarly needy.” We specifically choose the lottery conceit (rather than just saying Recipient B got the lawyer) because lottery assignment suggests that Recipient B received the good due to random chance, not greater need.

Another welfarist reason Spectators may hesitate to rerun the lottery is if they internalize Recipients' reference points from the initial allocation. Indeed, Charité et al. (2022) find in related redistribution experiments that redistributors sometimes incorporate receivers' reference points when told that receivers are aware of initial allocations, depressing demand for redistribution. We take significant steps to rule out this channel: we inform and remind the Spectator that both potential Recipients are unaware of the initial allocation. The experiment text says: "Remember that the **tenants will not know that the lottery was rerun**. They will just learn the final result, and the ultimate allocation will be anonymous" (emphasis in experiment text). We also include a confirmation check that asks participants whether the Recipients will know who was originally supposed to receive the good. 98% of participants get the question right. For these participants, we reiterate: "**That is correct**. Tenants will only learn the final result of the lottery." We correct the 2% of participants who get the question wrong: "**That is incorrect**. Tenants will only learn the final result of the lottery." Thus, Spectators do not internalize Recipients' reference-dependent utility in a way that would exaggerate estimated non-welfarism. This conclusion holds even if Spectators have uncertainty about exactly how Recipients form reference points.

Why This Experiment? While the experiment takes place in a specific context, it is motivated by, and sheds light on, broader ideas in the literature. For instance, Experiment 1 borrows from thought experiments like the trolley problem (Foot, 1967; Thomson, 1976; Bénabou et al., 2024), which are used to examine non-consequentialist intuitions. It also links to the philosophical notion of rights being inalienable. More broadly, it is sensible to think that Spectators could care about shuffling allocations. The experiment thus relates to ideas like the sunk-cost fallacy.

3.1.2 Results

We begin by showing cumulative distribution functions of the raw WTP data for each good (Figure 2A). A first observation is that WTP is positive for between 40–60% of participants, for all goods. In fact, a remarkable 46% of Spectators exposed to our neutral benchmark goods still have a positive WTP to preserve the lottery. If Spectators were welfarists, their WTP should be 0.

One may be concerned that high levels of non-welfarist behaviors with benchmark goods reflect inattention, confusion, or elicitation issues. But the magnitudes cast doubt on these as a complete explanation. It is unlikely that a full 40% of participants are confused.

Second, the rights goods' WTP distributions first-order stochastically dominate the benchmark goods. This result strongly suggests some domain-specificity in non-welfarist preferences (pooled Wilcoxon p -value of rights goods versus benchmarks < 0.001).

Turning to formal tests of differences in WTP, Spectators have 0.29 s.d. higher average WTP to preserve the lottery for rights goods than for benchmarks (Figure 2B, s.e. of difference: 0.05). Converted back to units of money saved for future programming, participants' WTP is \$20 higher

for rights goods than benchmarks. Inspecting the propensity to pay anything (Panel C), we find that 53% of participants have positive WTP to preserve the lottery for rights goods, 7.8 pp higher than benchmarks (s.e. of difference: 2.5). These differences reject the null of no domain-specificity. As discussed, the differences also reject the null of no non-welfarism under the adversarial hypothesis that all non-welfarism in benchmarks owes to elicitation issues.

Domain-specificity within rights goods is modest. WTPs are slightly larger for lawyers than health care, but nearly identical on the extensive margin.

Table A3 collects results for the outcomes of WTP, having positive WTP, and having maximum WTP. On order effects, differences by good are similar if we look at Spectators who do Experiment 1 first, although levels fall somewhat for both rights and benchmark goods (Table A4).

Open-Ended Responses Support Our Interpretation of the Results. We asked participants why they made their WTP choice. One wrote: “It’s the principle of the matter. Even if the tenants wouldn’t know, you’d know.” Another wrote, “If I re-ran the lottery it would feel like I was removing the lawyer from the first winner, and it would feel wrong.” These explanations also manifest among benchmark goods: “I decided to keep the lottery results for each trial because rerunning the lottery and taking away the original winners’ YMCA seems very unfair.”

Text-analysis methods (Ferrario and Stantcheva, 2022; Haaland et al., 2024) reveal that these anecdotal patterns are systematic (Figure A7A). Responses from those with positive WTP to overturn the lottery disproportionately mention fairness and the original winner. Remaining participants largely focus on saving money.

3.2 Experiment 2: Choice

3.2.1 Design

Motivation. The logic behind Experiment 2 is that welfarists only value providing Recipients with the ability to choose insofar as the choice might be exercised. If welfarists are completely sure a Recipient would choose (a) over (b), then their WTP to give the Recipient choice between (a) and (b), versus giving (a) directly, is zero. The experiment relates to ideas about the freedom to make choices in Hayek (1960)’s classical liberalism or Sen (1985)’s capabilities approach, among others.¹⁹

Design. Suppose a Recipient facing the choice of \$ y in cash and a good chooses \$ y with probability p . Experiment 2 begins by eliciting Spectators’ beliefs about the probability p . Then, Spectators face the choice of: (i) providing \$ y to the tenant directly and saving \$ x for future programming for the nonprofit, versus (ii) giving the Recipient the choice between \$ y and the good. We elicit (bounds on) the value of \$ x that makes Spectators indifferent between (i) and (ii).

¹⁹One may be concerned that Recipients could have intrinsic value from making choices, an objection which we examine in detail in Section 3.2.3 and Appendix E.

We estimate willingness to pay C_i , the midpoint of these bounds, for rights and benchmark goods. We focus on C_i as p approaches 1. When the Spectator has high beliefs p , the Spectator burns surplus to provide choice, which welfarists would not do.

Our test requires conditioning on people with high beliefs p , which could lack power. To increase Spectators' beliefs, we provide all Spectators with truthful information from a randomly selected pilot sample. The information says that all tenants in the pilot sample chose cash over the good (see Appendix D.6 for details on the information). This information raises power by increasing the number of Spectators with high beliefs.²⁰ We therefore sometimes call the beliefs "posteriors," as we elicit them after giving information.

Alternative Measure. Appendix E defines and presents confirmatory results using a second measure of the value of choice. The idea is that we can infer a normative WTP for choice from the probability the choice will be exercised and the directly elicited value of giving the good directly.

3.2.2 Results

Spectators exhibit preferences for providing choice with both rights and benchmark goods (Figure 3). We start with cumulative distribution functions of the raw WTP data among Spectators with posterior beliefs exceeding 0.9. We find that between 30%–60% of Spectators have a positive WTP for providing choice, strongly suggesting the presence of non-welfarist preferences. Even more so than in Experiment 1, WTP for rights good stochastically dominates benchmark goods (Wilcoxon p -value < 0.001 , Panel A).²¹

Zeroing in on comparisons of rights versus benchmark goods, we present differences in mean WTP across beliefs bins (Panel B). Focusing on rights goods (blue series), we reassuringly find that WTP C_i is decreasing in beliefs that Recipients will choose the good over cash. This negative relationship reflects that choice is less valuable if it is unlikely to be exercised.

For every bin of beliefs, willingness to pay for rights goods is higher than for benchmark goods. Among Spectators with beliefs larger than 0.9, WTP is 0.52 sd (s.e.: 0.067) higher for pooled rights goods than benchmarks. Converted back to units of money for future programming, Spectators are willing to pay \$395 on average to give choices for pooled rights goods (\$223 for benchmarks). Spectators have higher WTP for choices over lawyers than health care, but both differ from benchmarks. Differences persist even if we condition on having beliefs larger than 0.95 (farthest right whiskers, Figure 3B). Rights versus benchmark differences also persist on the extensive margin, having a positive WTP for choice at all (Figure 3C).

²⁰Conditioning on beliefs could, in theory, affect experimental balance. First, Table A6 shows that balance persists conditioning on beliefs. Second, randomization is not required for this test. Any positive WTP as $p \rightarrow 1$ still indicates non-welfarist preferences. Third, our second measure of the value of choice is not subject to this concern.

²¹Raw dollar values are not comparable between experiments. We used a larger range for Experiment 2's elicitation, as a small chance of needing a lawyer could imply a large willingness to pay.

Tables A7 and A8 collect these results. In another test, instead of restricting to people with high posteriors, we perform our test within the whole sample and control for Spectators' posterior beliefs about whether the Recipient would choose the good. Doing so generates similar results (Table A7, Columns 1–3). Related to this test, results are similar if we use the sample with posteriors above 90 and control for their beliefs (Table A9, Columns 6–8).

Pushing our experiment to its logical conclusion, we examine the 459 Spectators who say there is a 100% chance the Recipient will choose cash. Rights CDFs first-order stochastically dominate benchmarks in this restricted sample (Figure A2, Wilcoxon p -value < 0.001). We focus on posteriors exceeding 0.9 for power, and because our incentive scheme rewards people equally if they had posteriors of 0.96–1, but results are similar with this restricted sample (Table A7, Columns 7–9). One difference is that among Spectators with posteriors = 1, results are stronger on the intensive versus extensive margin.

Open-Ended and Qualitative Questions Support Our Interpretation. An open-ended response question asks Spectators why they made their choice. A Spectator seeing lawyers who had the maximum WTP wrote, “I think the tenant has a right to choose what assistance to accept.” Another wrote, “The tenant has a right to choose, no matter what the monetary consequences.”

We again confirm these patterns with systematic text analysis (Figure A7B). Both groups mention saving money, but those with positive WTP discuss giving choice (e.g. “tenant choice”). Those with zero WTP mention tenant preferences (e.g., “prefer cash,” “tenants want”).

We also ask Spectators a qualitative question to explain their behavior. Specifically, we ask: “Which of the following reasons motivated your choice(s)? Select all that apply.” Options were: “I thought anyone who would choose the [good] would really want it”; “I did not think anyone would choose the [good] in reality”; “Saving is my priority”; “All tenants should be entitled to the choice of a [good]” (the right to choice option); and “None of the above.” The share of Spectators who say that Recipients have the right to choose when facing a rights good is about 42%, which is 16.1 pp (61%) more likely than with benchmarks (Table A9, Column 8).

Connection to Framework. To view this experiment through the lens of our framework, consider the choice itself as a good y . Suppose also that $r = 1$ for providing choice, that $\lambda = \lambda'$ for all goods, and that u_j does not itself depend on choice. Then our experiments test $H_0 : \eta_{\text{rights}} = \eta_{\text{benchmarks}}$, which is conservative for testing $H_0 : \eta_{\text{rights}} = 0$ as long as $\eta_{\text{benchmarks}} \geq 0$. Here, η 's relate to non-welfarist concerns over providing choice.

3.2.3 Addressing Concerns

Testing Selection on Gains. A concern is that the Spectators who benefit from exercising choice (even if low probability) may have high utility from choice. Such “selection on gains” is relevant only if beliefs p are mismeasured. That is because selection on gains still vanishes from welfarist

Spectator utility as $p \rightarrow 1$. And coefficient stability as we restrict to samples with higher posteriors, or control for posteriors, argues against this as a key concern. Still, one may worry that Spectators' beliefs are mismeasured due to noise, lack of numeracy, or elicitation issues. In Appendix E we discuss a sub-experiment manipulating the cash value in the choice bundle, which provides evidence against a meaningful selection on gains channel. The secondary outcome measure in the appendix also argues against this concern and does not involve measuring beliefs.

Intrinsic Value of Choice. Conditioning on $p \rightarrow 1$ intends to restrict to Recipients who do not value choice. But, Recipients could have intrinsic utility from choosing (Bartling et al., 2014; Freundt et al., 2023; Lenk, n.d.). Then, even if $p = 1$, welfarist Spectators may internalize Recipients' intrinsic choice utility and pay to give Recipients choice.

Our most compelling test of this concern is that we difference Spectators' WTP for choice with rights goods versus benchmarks. That is, the differences across goods remain valid if Spectators believe Recipients get an additive choice utility that does not depend on the good.

Still, our test overstates non-welfarism if Recipients get intrinsic choice utility that is larger for rights goods and if Spectators internalize their choice utility. Appendix E discusses two other robustness checks, and interpretations if our exercise does include some intrinsic value of choice.

3.3 Experiment 3: Equal Rights (Egalitarianism)

3.3.1 Design

Spectators could believe in equal rights to certain goods, and get a payoff from upholding such rights. Such views might support policy proposals like universal health care or shelter.

To operationalize this logic, we observe that welfarists' utility should not depend on the share of people in society who already get a good. To be concrete, suppose z out of 10 people get a lawyer regardless. Welfarists' WTP to provide the $(z + 1)$ th Recipient a lawyer should not vary with z differentially for rights goods. But if Spectators derive non-welfarist value from sustaining "equal rights," they have a higher WTP to provide the good when $z = 9$ than when $z = 1$.

Our experiment tests this idea by informing participants that z out of 10 *ex ante* identical and anonymous Recipients were selected to receive lawyers. We then elicit participants' willingness to pay to provide the $(z + 1)$ th person with a lawyer. In this case, the donation budget is money donated from the nonprofit to a food bank.²² We randomize $z \in \{1, 5, 9\}$.

²²Had the outside option been "saving for future programs" as in Experiments 1–2, then choices in Experiment 3 could never be universal. Tenants in future programs would not be assisted. Put another way, "future programs" would have raised the denominator from 10 to a larger number.

We estimate the following difference-in-differences specification:

$$y_i = \delta_0 \text{Right}_i + \delta_1 \mathbb{1}(z_i = 9) + \beta_0 (\text{Right}_i \times \mathbb{1}(z_i = 9)) + \varepsilon_i \quad (15)$$

$$y_i = \delta_0 \text{Lawyer}_i + \delta_1 \mathbb{1}(z_i = 9) + \delta_2 \text{HC}_i + \beta_l (\text{Law}_i \times \mathbb{1}(z_i = 9)) + \beta_h (\text{HC}_i \times \mathbb{1}(z_i = 9)) + \varepsilon_i. \quad (16)$$

The coefficients of interest are β_0 , β_h , and β_l .

The difference-in-differences specification addresses an important concern that inequity aversion (Fehr and Schmidt, 1999, still a welfarist concern) generates higher WTP for *any* good if $z_i = 9$. To further account for inequity-averse preferences to reduce absolute differences in utility across Recipients, we control for Spectators' direct WTP to provide the good. Then, between-good difference-in-differences in WTP reflect only a differential “preference for equal rights.”

3.3.2 Results

We find no evidence of differential preferences for equal rights with rights goods compared to benchmarks (Figure A4, Table A14). Across all goods, WTP to allocate another good rises with z . That is, WTP to allocate $z + 1$ goods at $z = 9$ is larger than $z = 1$ or $z = 5$. But the difference-in-differences, comparing benchmarks to rights goods, is actually negative.

This null result can be interpreted in several ways. First, non-welfarist preferences may extend to all four goods, including benchmarks. Alternatively, as there is no differential preference for equal rights with rights goods, these tests cannot reject the presence of welfarist but inequity-averse preferences (see Table A14 for quantification). In fact, when we include a control for direct WTP for the good, which further accounts for inequity aversion, we find a significant *negative* coefficient, indicating preferences *against* equal rights for rights goods. Elicitation errors or inattention are a less persuasive explanation for the spike at $z = 9$ as they are unlikely to differ at $z = 9$ versus $z = 5$ and $z = 1$.

As a fundamental tenet of liberalism, preferences for equal rights were reasonable to examine *ex ante*. However, equal rights are also hard to manipulate in the lab. We cannot control the share of people in society who have the good. One explanation for the null result is that Spectators internalize that, no matter their choice, many people will lack lawyers or health care.

3.4 Taking Stock

Non-Welfarism Is Prevalent Across Experiments. Using the data from both Experiments 1 and 2, which we focus on moving forward, affirms our findings: non-welfarism is prevalent and especially concentrated among rights goods. A full 73% have positive WTP for rights goods in either of Experiments 1 or 2 (Table 2). In fact, 31% have positive WTP in both experiments, which is 12 pp (65%) more than the 19% who express positive WTP in both experiments with

benchmarks ($p < 0.001$). Thus, “intense” non-welfarism especially manifests for rights goods.

Multiple Hypothesis Corrections. As we find results consistent with our hypotheses in two of three experiments, we perform multiple hypothesis corrections for Experiments 1–3 (Table A15). Romano-Wolf adjusted p -values for Experiments 1–2 remain significant at $p < 0.01$.

4 Implications for Applied Welfare Economics

Having found evidence of non-welfarist preferences in Experiments 1 and 2, we turn to their implications for applied welfare and public economics.

Our first set of tasks examines how non-welfarist preferences correlate with Social Welfare Functions. Using both a new redistribution task and participants’ stated preferences, we find that non-welfarist concerns predict support for less-progressive provision.

Next, we turn to incidence. In a pre-registered heterogeneity analysis, we detect that non-welfarist preferences are much more common among middle-income versus low- or high-income households. Simulations show that this inverted U-shape induces a trade-off between non-welfarism and progressivity, as foreshadowed in the discussion of Proposition 1.

We generally focus on non-welfarist preferences within the rights goods and how they explain policy support. That is because we want to study how non-welfarist concerns influence real-world policies like allocating health care or eviction lawyers as public transfers. Appendix analyses show that some patterns differ for benchmark goods, in interesting ways.

4.1 Redistribution Decisions: Anti-Targeting (Experiment 4)

Our goal is to measure how Spectators redistribute lawyers, health care, and benchmarks, and correlate these redistributive tastes with the non-welfarist concerns expressed in Experiments 1–2. Such redistributive choices are of intrinsic interest. There is scant evidence about how redistributive concerns vary across domains.

4.1.1 Design

We want to measure redistributive preferences over indivisible goods where Recipients only benefit from provision at the extensive margin. It is not obvious how to do that. To see why, suppose there are 10 people who need lawyers; they can be uniquely sorted by income (i.e., no ties); there are $R < 10$ lawyers; and no one benefits from multiple eviction lawyers. Anyone with progressive redistributive preferences gives the lawyers to the R poorest people. Thus, we cannot simply ask Spectators how they would allocate R lawyers among 10 people.

Experiment 4 introduces smoothness into the problem as follows. We truthfully tell Spectators that 10 Recipients with annual household incomes ranging from \$0 to \$36,000, in increments

of \$4,000, have applied for assistance. Spectators may give all Recipients the good g , again randomized across the four goods. Alternatively, Spectators may give the poorest $R \in \{1, 2, \dots, 9\}$ people the good as well as cash. The value of the cash is decreasing in R . Thus, Spectators face a trade-off between (i) giving more money and the good to fewer, needier households, versus (ii) less money and the good to more households, where the marginal household is less needy.

To ensure that every choice considers the same budget, we fix the total redistributive budget for this choice at B . The task asks Spectators how many goods they want to “buy” out of the budget, versus how much cash to distribute. Good g ’s price is $p_g = B/10$. If the Spectator buys R goods, the poorest R Recipients get the good. Money not spent on buying the good is divided equally among the R Recipients. That is, Spectators choose between giving the good and $\frac{B-p_g R}{R}$ dollars to the R poorest Recipients, or all 10 Recipients the good. Using multiple price lists, we find the number of Recipients that makes a Spectator indifferent between giving the good to everyone versus money plus the good to fewer Recipients.

As an example for one good, we price lawyers at \$500 and fix the redistributive budget $B = \$5,000$. First, Spectators choose between giving (i) lawyers to everyone, versus (ii) five Recipients a lawyer and \$500 in cash each. If they choose (i), they face the choice of giving lawyers to everyone versus six Recipients a lawyer and \$333 in cash each. We iterate on these questions until we find R_i , the Spectator’s preference of how many Recipients should get the good, for $R_i \in \{1, \dots, 10\}$.

For simplicity, elicitation imposed monotonicity. We assume that preferring a universal bundle to giving to R Recipients means that the Spectator also prefers giving the universal bundle to $R - 1$ Recipients. We stress the outcome of preferring a universal allocation, which means that the Spectator explicitly reported preferring to give the good to 10 Recipients than 5, 6, 7, 8, or 9 Recipients. This outcome still has meaning even if one is concerned about monotonicity.²³

Design Considerations. It is important to choose the price of each good (equivalently, the budget) correctly. To see why, suppose p_{lawyer} were \$1 and p_{bus} were \$100. Then, since not much cash can be redistributed by giving lawyers to fewer people, and lawyers may be very effective, most Spectators would likely choose to give 10 Recipients the lawyer. We price health care at \$600, lawyers at \$500, YMCA at \$300, and bus passes at \$250. We selected these prices to be the median of pilot WTP elicitation. Making rights goods more expensive pushes toward allocating them less universally. This choice is conservative for our ultimate conclusions. Prices are not identical to the direct WTP we elicit for each good (Figure A8), a discrepancy that we return to below.

In addition to doing this exercise with rights and benchmark goods, we also randomize half of the Spectators assigned to benchmark goods into doing this exercise with cash, again with budget $B = \$5,000$. Spectators in this elicitation choose the value R at which they are indifferent between giving B/R in cash to R people or $B/10$ to 10 people.

²³Elicitation of WTP throughout the study also imposes monotonicity. Monotonicity in Experiment 4 is less standard, since the number of Recipients R as well as the monetary top-ups given to those Recipients change in each question.

4.1.2 Results

Spectators are more likely to “anti-target” — that is, give goods universally (to all 10 Recipients) with rights goods than benchmarks (Figure 4). Pooling lawyers and health care, 43% of Spectators who allocate rights goods anti-target, compared to 26% of Spectators allocating benchmarks or cash (s.e. of difference: 2.3 pp). Both lawyers and health care are significantly different than the benchmarks and cash. Spectators are more likely to anti-target with lawyers than health care. Lawyers are 26.3 pp more likely to be anti-targeted than benchmarks or cash (s.e. of difference: 2.7 pp), whereas health care is 7.1 pp more likely to be anti-targeted (s.e.: 2.6 pp). Table A16 aggregates the tests, along with similar results for a continuous outcome of the number of Recipients assisted.

These results do not imply that Spectators would give goods universally for any population of Recipients. However, fixing the Recipient population, Spectators’ targeting preferences appear flatter for rights goods than benchmarks or cash.

Open-Ended Responses Confirm Our Interpretation. We ask Spectators to explain the reasons for their allocation. Those who anti-target mention how all deserve assistance, regardless of income. Those who do not disproportionately discuss the poorest tenant (Figure A7C).

4.1.3 Non-Welfarist Preferences and Anti-Targeting

Our framework predicts that non-welfarist preferences lead to flatter allocation of goods (Proposition 2). Thus, we hypothesize that Spectators who demonstrate non-welfarist preferences in Experiments 1 and 2 provide rights goods to more Recipients in Experiment 4.

Non-welfarist preferences indeed have strong explanatory power for redistributive decisions in our incentivized allocation task. In what follows, we sometimes call WTP in Experiments 1–2 “rights preferences.” Rights preferences in Experiments 1 and 2 predict propensity to anti-target (Table 3). For instance, a 1 s.d. increase in WTP for choice in Experiment 2 correlates with a 9 pp increase in anti-targeting, a 20% increase off the mean of 44 pp.

As another way of presenting these correlations, we find that 53% of Spectators who express non-welfarist preferences in both Experiments 1–2 provide rights goods universally, compared with 42% (p -value of difference = 0.023) of Spectators who are non-welfarist in exactly one experiment and 36% of Spectators who do not express non-welfarist preferences in either experiment (Figure 5A). Meanwhile, 25% of Spectators provide benchmarks universally.²⁴

Thus, looking *within* rights goods, there is a 16 pp difference in anti-targeting for those with rights preferences in both experiments versus neither (2nd and 4th bars of Figure 5A). That is similar to the 17 pp *between*-good difference in anti-targeting of rights goods versus benchmarks

²⁴With benchmarks, correlations between anti-targeting and non-welfarism are directionally consistent but smaller in magnitude (Figure A11).

(Figure 4). Alternatively, the 16 pp difference explains more than half the 28 pp gap between anti-targeting in rights versus benchmarks (1st, 2nd, and 4th bars of Figure 5A).

Such between-experiment correlations give compelling evidence of non-welfarism's relevance for public policy. Nothing about Experiment 4 mechanically pushes non-welfarists, as measured in Experiments 1–2, to anti-target.

Addressing Objections. We purposefully set a price for each good that is conservative with respect to generating anti-targeting of rights goods — that is, Spectators could give more cash to the poorest if they chose not to anti-target lawyers and health care. Despite these efforts, the prices were not set perfectly, as WTP in the experiment for lawyers exceeds that from pilots. In particular, average WTPs for YMCA, bus, health care, and lawyers are: \$328, \$373, \$507, and \$765 respectively (Figure A8). Even if the price of each good differentially pushes toward universal provision, that would not bias our main result: the *within-good correlations* between anti-targeting and non-welfarist behaviors.

Another test controls for Spectators' elicited WTP (Table A17). This control accounts for how instrumentally valuable Spectators think the good is, which addresses concerns about incorrectly set prices. With this this control, non-welfarism still predicts anti-targeting propensity (Columns 1 and 2). We do observe modest attenuation, but Columns 5 and 6 show this comes from allocating *health care*, even though *lawyers* are the only good where elicited WTP exceeds the implied price. The correlation between I_i and C_i and universal provision also persists even with controls for direct WTP (Table A18). Controlling for WTP is a conservative exercise here, as non-welfarist concerns may raise WTPs directly, so WTP risks being a “bad control” (Angrist and Pischke, 2009).

A related objection involves paternalism. If Spectators believe low-income participants will misuse cash, they may prefer universal provision of any good. Rights goods versus benchmark differences in universal provision address this concern.

4.2 Non-Welfarist Preferences Predict Stated Policy Support

We conclude the experiment by asking Spectators if they support (i) “right to counsel” policies that provide lawyers to tenants facing eviction, (ii) rent control, and (iii) universal health care. We then regress support for these policies on exhibiting non-welfarist preferences in both Experiments 1–2. For right to counsel and rent control, we conduct this exercise among Spectators who did the experiments with lawyers. For universal health care, we conduct the exercise among Spectators who did the experiments with health care.

Non-welfarist preferences predict support for right to counsel and rent control, but not universal health care (Figure 5B). For right to counsel and rent control, the relationship survives adding a control for whether the Spectator believes the policy would be effective, as well as for whether the person is a liberal. Thus, non-welfarist preferences, at least in the context of providing

lawyers, predict policy support on top of welfarist/instrumental views about whether policy will help people. One explanation for why non-welfarist preferences could be less predictive of support for universal health care is that this issue is more politicized.

Toward examining domain-specificity, we also ask Spectators if they agree there is a right to several types of goods. Views vary, from more than 75% who believe in a “right to food” versus 45% who believe in a right to an eviction lawyer. Strongly agreeing there is a right to goods like food, education, and housing is correlated with having non-welfarist preferences (Figure A14). One exception is agreement with the view that there is a right to a lawyer in criminal cases. Spectators could view that question as a factual matter, as this right is guaranteed in the U.S.

4.3 Incidence and Demographic Heterogeneity

If Spectators with non-welfarist preferences are mostly rich, then SWFs face a trade-off between incorporating non-welfarist concerns and progressivity. To examine the incidence of non-welfarist preferences, we consider someone a non-welfarist if they express non-welfarist preferences in both Experiments 1–2, focusing only on those who see lawyers or health care.

We find a stark inverted U-shaped pattern, where Spectators with annual household incomes of \$40–60,000 exhibit significantly higher rates of non-welfarism (Figure 6A). These middle-income Spectators are 42–65% more likely to be non-welfarist than those with incomes of \$100,000+ or \$0–20,000. Non-welfarism exhibits mostly small correlations with other characteristics (Panel B).

Table A21 quantifies and shows incidence of non-welfarism in Experiments 1–2 separately. The inverted-U holds in each experiment, but statistical significance benefits from pooling them.

This pattern suggests the possible validity of a key objection to non-welfarist SWFs: they are replicated by welfarist SWFs that implicitly downweight the poor. We explore this concern with two illustrative simulation exercises.²⁵

Incidence Quantification 1: Implied Weights. Our first exercise extracts the welfare weights that a welfarist SWF uses to replicate an example non-welfarist SWF, akin to Proposition 1 or Example 1. We imagine that the non-welfarist social planner, if she participated in Experiments 1 or 2, would have “normatively correct” indifference points (I^*, C^*) . We simulate the welfare weights on income groups that minimize the difference between the normatively correct indifference points and the average indifference point in society. For simplicity, we collapse income groups into three bins i : \$0–40,000, \$40,001–\$60,000, and more than \$60,000. Then, we solve the system of equations:

$$\frac{\sum_i p_i w_i \bar{I}_i}{\sum_i p_i w_i} = I^* \quad (17)$$

²⁵Meanwhile, non-welfarist preferences for non-rights goods show no such U-shaped pattern. So, this objection may be particularly relevant for non-welfarism and policy-relevant transfers, rather than a concern about non-welfarist behaviors more generally (Figure A12).

$$\frac{\sum_i p_i w_i \bar{C}_i}{\sum_i p_i w_i} = C^* \quad (18)$$

$$\sum_i w_i = 1, \text{ and } w_i \geq 0 \text{ for all } i \quad (19)$$

for welfare weights w_i , share in each bin p_i , preferred social allocation (I^*, C^*) , and average indifference point in Experiments 1 and 2 \bar{I}_i and \bar{C}_i . We minimize the sum of squared deviations from Equations (17)–(18), standardizing each equation by I^* and C^* respectively. We extract weights w_i for many possible non-welfarist concerns (I^*, C^*) lying in the convex hull of Spectators' average group-level values (\bar{I}_i, \bar{C}_i) .

SWFs which value non-welfarist concerns place less weight on the poorest (low w_1), and more weight on the middle-income group (high w_2). Each dot in Figure 6C corresponds to weights (w_1, w_2) induced by a posited (I^*, C^*) pair. The dotted line between $w_1 = 1$ and $w_2 = 1$ is where $w_3 = 1 - w_2 - w_1 = 0$, so interior points have positive w_3 . For instance, the green dot with an arrow shows a non-welfarist SWF that pays \$85 in Experiment 1 and \$375 in Experiment 2. For this dot, the replicating welfarist SWF has weights $(w_1, w_2, w_3) = (0.19, 0.81, 0)$: that is, most weight on middle-income, then low-income. Moving up or to the right on the graph indicates more income progressivity (less weight on w_2 or w_3).

As Panel C shows, when the SWF is more non-welfarist, weights on the first income group are lower. The most progressive SWFs place relatively less weight on non-welfarist concerns (yellow dots toward the bottom right of the graph).

Incidence Quantification 2: Redistribution. Consider a non-welfarist SWF that faces the choice of taking \$1 from a reference Spectator j and redistributing to Spectator i . The SWF trades off redistributive motives (is i poorer than j ?) with non-welfarist concerns for valuing donations to people who are non-welfarist (is i more non-welfarist than j ?) in the spirit of Rabin (1993).

We formalize this trade-off and present detailed results in Appendix E. We consider the choice of redistributing from the middle-income Spectator to the lowest-income Spectator. With low enough weight on non-welfarist concerns, the SWF gets positive utility from such redistribution. But this preference flips with reasonable weight on non-welfarist concerns, and then the SWF prefers keeping money with the middle-income person.

Both exercises are simple yet illustrative. The SWFs we chose are not necessarily favored by real-world non-welfarists. However, the trade-off we found likely extends, as inverted-U incidence indicates far less non-welfarism among low-income households.

4.4 Additional Analysis: Quantifying Welfarists via Information Provision

An additional information sub-experiment, described in detail in Appendix E, quantifies the share of welfarists and contrasts with non-welfarists. Even when we surprise Spectators with relevant

information about the efficacy of lawyers and health care that disagrees with their priors, 40% or less change their allocations in Experiment 4’s targeting task — evidence that some Spectators do not exclusively maximize welfarist goals. Classifying Spectators into welfarists and non-welfarists using this extra treatment, we detect similar rates of non-welfarism and welfarism.

The sub-experiment replicates and rationalizes the moderate effect sizes of information treatments on policy support in the literature (see Haaland et al., 2023). With non-welfarist concerns, support depends less on beliefs about policy’s consequences.

4.5 Robustness Checks

Incentives. As health care was not incentivized, we embedded two tests to measure the importance of incentives (see details in Appendix D.3). First, we randomize benchmark goods into being incentivized throughout all experiments (as in lawyers) or not incentivized (as in health care). In particular, we randomize benchmark goods into receiving identical language as those who see health care. Comparing benchmark participants who see the health care language to those who see the lawyer language therefore jointly tests for the importance of incentives and other minor language differences that differ between health care and the other studies. If incentives matter for health care, we would expect them also to matter for benchmark goods. Yet we find no effect of incentives among benchmark participants (Table A22). As a consequence, rights goods versus benchmark goods differences in Experiments 1–4 also persist using only incentivized or only unincentivized benchmark participants (also Table A22).²⁶

Second, we elicit Spectators’ WTP to give Recipients the good directly. This is the secondary WTP elicitation, distinct from Experiments 1–4. We randomly assign half of the participants assigned to the three incentivized goods to have this WTP elicitation be unincentivized and clearly labeled as hypothetical, whereas the other half sees that the elicitation is incentivized. If incentives mattered in Experiments 1–4, we would expect experimental manipulation of incentives to affect the WTP elicitation as well. Yet, we reject even small effects of incentives here (Table A23).

We thus find no evidence that lack of incentives for health care influences results. The advantage of the first test is that it examines incentives in Experiments 1–4, whereas the advantage of the second test is that it examines how incentives affect provision of a rights good (lawyers). These findings add to a conflicted literature on the importance of incentives in laboratory experiments (Andersen et al., 2011; Gneezy et al., 2011; Charness et al., 2021; Danz et al., 2022; Sullivan, 2023).

Reweighting. Reweighting the online sample to match demographics in the U.S. population has no impact on results (Table A24).

Inattention. Inattention cannot explain the large share who express non-welfarist preferences.

²⁶We did not randomize incentives among lawyers in the main experiments because we did not know *ex ante* that incentives would have small effects, and we did not want to reduce power if they proved important.

We regress outcomes on an indicator for failing one attention check (Table A25). Inattention is insignificantly correlated with small increases in non-zero WTP (5 and 2 pp in Experiments 1 and 2) and universal allocations (3 pp in Experiment 4; all p -values > 0.2). Importantly, the correlation with inattention is small relative to levels of non-welfarism and universality among the attentive (50%, 48%, and 37% in Experiments 1, 2, and 4). Thus, to generate our results, unobserved inattention would need to be much more important than observed inattention.

We now examine if inattention affects differences in non-welfarism by good, rather than levels. If anything, we find slightly higher inattention with benchmark goods (7.4%, versus 7.1% with rights goods). That pushes against observing differences by good due to inattention.

5 Mechanisms: Non-Welfarist Decisions and Welfarist Consequences

Having shown the relevance of non-welfarist concerns for economics and policy, a natural question is why some domains induce non-welfarism more than others. A full taxonomy of which “spheres” pertain to justice is beyond the scope of this paper. But, in a first step, we study a particular force that influences non-welfarist behaviors: namely, a decision’s welfarist consequences.

5.1 Motivation and Design

Motivation. One hypothesis is that welfarist consequences should attenuate non-welfarist reasoning, perhaps because non-welfarist reasoning reflects small-stakes phenomena. Another hypothesis is that welfarist consequences draw attention to the question, removing automaticity in decision-making, and thus heighten non-welfarist concerns. A separate, pre-registered mechanisms experiment (AEA Registry #0014218) tests these hypotheses.

Experiment Design. We modify Experiment 1, looking at the propensity to pay not to rerun the lottery under different conditions. We focus on providing lawyers and experimentally manipulate three aspects that influence the lottery’s welfarist consequences.

One way to think about these treatments is with the following identity:

$$\begin{aligned} \text{Welfarist Consequences of Providing Lawyer} & \quad (20) \\ &= \text{Lawyer Effectiveness} \times \text{Case Magnitude} \times \text{Social Marginal Welfare Weight of Recipient.} \end{aligned}$$

The three sub-experiments each vary one term on the right-hand side.

Specifically, we have three treatment groups and six arms. In the first two arms (“effectiveness group”), we vary whether the lottery assigned a tenant to receive an *effective* lawyer who almost always wins their cases, versus an *ineffective* lawyer who almost always loses their cases.

In the second two arms (“magnitude group”), we change the setting to be lawyers defending a

speeding-ticket case. We vary the magnitude of the ticket, from \$50 to \$2,000. We tell participants that the Recipients can versus cannot afford the fine. We changed to a speeding-tickets case since it was simpler to vary the magnitude of the fine in that context than for eviction.

In the third two arms (“identity group”), we manipulate the identity of the Recipient to be low- versus middle-income. Middle-income Recipients have lower social marginal welfare weights.

Experiment Logistics. We recruit 6,000 participants on Prolific (1,000 per arm). Table A26 shows the experiment sample, attrition, and balance, and Appendix D.8 provides details about recruitment.

A limitation is that this experiment is always hypothetical. Given incentives’ small impacts in prior experiments (Section 4.5), their absence here likely does not affect results. As in Experiments 1–4, we tell Spectators that the non-profit will be informed about the results, and we ask participants to take the elicitation seriously.

5.2 Results and Discussion

Across all three treatments, welfarist consequences increase the presence of non-welfarist concerns (Figure 7). Moving left to right, average WTP is \$45 when the lawyer is ineffective, versus \$79 when the lawyer is effective (p -value of difference < 0.001). Differences are even larger when the fine magnitude is low versus high ($p < 0.001$). We also find that welfarist stakes raise non-welfarist behaviors when the lottery involves low- versus middle-income Recipients, but the difference is smaller (\$12, $p = 0.001$). The difference-in-differences in treatment effects, across the first two groups versus the identity group, are significant.²⁷ Table A27 consolidates findings and reports similar results with the other outcomes from Experiment 1.

We draw several conclusions. First, welfarist consequences dramatically raise the propensity to make non-welfarist decisions. Non-welfarism’s relevance in high-stakes situations suggests that our study’s findings could extend to high-stakes environments like the rights to shelter or food, as the qualitative evidence in Figure A14 suggested.

Second, our results are more consistent with some cognitive models than others. In particular, the results provide some support for models of attention or salience, in which large welfarist consequences prompt soul-searching (Ottati, 1990; Druckman and Lupia, 2000). That is, large welfarist stakes, juxtaposed with difficult choice facing the Spectator, could induce an emotional response that draws attention “bottom-up,” leading the Spectator to emphasize moral reasoning (Goodenough and Prehn, 2004; Bordalo et al., 2022). Meanwhile, small choices may be made more automatically, via welfarist heuristics that suffice for most daily choices (Simon, 1955).²⁸

²⁷ $p < 0.001$ on each difference-in-differences with identity group; $p = 0.138$ for difference-in-differences of effectiveness and fine magnitudes groups.

²⁸ The data are less consistent with a pure rumination channel, as we do not find that Spectators spend more time on the task when exposed to high-stakes treatments (Table A28). Even if time spent does not increase, emotional

Third, stakes have heterogeneous effects, as the identity treatment has distinguishable but smaller impacts. One pre-registered interpretation, consistent with earlier results, is that manipulating identity yields countervailing forces. First, non-welfarist concerns are more important as stakes rise, which should cause higher rates of non-welfarist behaviors with poorer Recipients. However, non-welfarist concerns also push toward universal provision, downweighting the importance of identity and attenuating the treatment effect (Proposition 2 and Experiment 4). That is, the gap in social marginal welfare weights between rich and poor is small with rights goods like lawyers. That reduces the difference in WTP to preserve the lottery across Recipients. Thus, the smaller treatment effect in this manipulation is consistent with an important role for stakes as well as anti-targeting — an encouraging finding that affirms the anti-targeting results in Experiment 4.

6 Conclusion

This paper points toward a middle-ground in welfare economics. Our experiments document prevalent and economically meaningful non-welfarist preferences. We further find evidence of domain-specificity, where people have stronger non-welfarist preferences for goods with high rights valence like health care or legal assistance, consistent with *Spheres of Justice*. Non-welfarist preferences correlate with policy support and appear to explain why people prefer less progressive allocations of lawyers and health care relative to other goods.

Our results imply that non-welfarist justification for policy-relevant public transfers — appealing to rights to health care and legal counsel — has empirical grounding. More broadly, welfarist Social Welfare Functions, which depend only on the preferences of people in society, will still incorporate important non-welfarist concerns in these domains.

Both non-welfarists and welfarists can take comfort in the findings. The former, because we detect many non-welfarists in society, perhaps raising non-welfarist frameworks’ appeal; the latter, because welfarism can accommodate the moral concerns so common in policy debates. Whichever philosophical framework one adopts, one must weigh individuals’ non-welfarist preferences — an approach very seldom seen in applied welfare economics. At the same time, we provide empirical backing for an important trade-off: emphasis on non-welfarist concerns could sacrifice progressivity, as the poorest Spectators are less non-welfarist in our setting.

We deliberately chose two controversial domains where non-welfarist concerns have influenced policy. But our findings may apply more generally: advocates call for rights to housing, education, or even “the right to access to and participate in science.”²⁹ If taken seriously, these views have dramatic consequences for how economists perform welfare evaluations of government

engagement might.

²⁹The United Nations, <https://www.ohchr.org/en/special-procedures/sr-cultural-rights/right-access-and-participate-science>, accessed 9/27/2024.

interventions. Methodologically, our experimental designs are portable to these and other settings, and we hope this paper sparks future research on non-welfarist preferences across domains.

References

- Abeler, Johannes, Daniele Nosenzo, and Collin Raymond**, "Preferences for truth-telling," *Econometrica*, 2019, 87 (4), 1115–1153.
- ACLU**, "No Eviction Without Representation," Technical Report, American Civil Liberties Union 2022.
- Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias**, "Targeting the poor: evidence from a field experiment in Indonesia," *American Economic Review*, 2012, 102 (4), 1206–1240.
- Alesina, Alberto and Paola Giuliano**, "Preferences for redistribution," in "Handbook of social economics," Vol. 1, Elsevier, 2011, pp. 93–131.
- , **Stefanie Stantcheva, and Edoardo Teso**, "Intergenerational mobility and preferences for redistribution," *American Economic Review*, 2018, 108 (2), 521–554.
- Alger, Ingela and Jörgen W Weibull**, "Homo moralis—preference evolution under incomplete information and assortative matching," *Econometrica*, 2013, 81 (6), 2269–2302.
- Allen, Heidi, Katherine Baicker, Mira Bernstein, Amy Finkelstein, Jonathan Gruber, Joseph P. Newhouse, Eric Schneider, Song, Sarah Taubman, Bill Wright, and Alan Zaslavsky**, "The Oregon Health Insurance Experiment in the United States — The Abdul Latif Jameel Poverty Action Lab," 2013.
- Ambuehl, Sandro and Bert Douglas Bernheim**, "Interpreting the will of the people: Social preferences over ordinal outcomes," 2024.
- , **B Douglas Bernheim, and Axel Ockenfels**, "What motivates paternalism? An experimental study," *American Economic Review*, 2021, 111 (3), 787–830.
- An, Zhiyong**, "A note on welfarist versus non-welfarist social welfare function," *International Tax and Public Finance*, 2024, 31 (4), 1066–1073.
- Andersen, Steffen, Seda Ertaç, Uri Gneezy, Moshe Hoffman, and John A List**, "Stakes matter in ultimatum games," *American Economic Review*, 2011, 101 (7), 3427–3439.
- Andre, Peter**, "Shallow meritocracy," *Review of Economic Studies*, 2024.
- , **Carlo Pizzinelli, Christopher Roth, and Johannes Wohlfart**, "Subjective models of the macroeconomy: Evidence from experts and representative samples," *The Review of Economic Studies*, 2022, 89 (6), 2958–2991.
- , **Ingar Haaland, Christopher Roth, and Johannes Wohlfart**, "Narratives about the Macroeconomy," CESifo Working Paper no. 10535, 2023.
- Andreoni, James**, "Giving with impure altruism: Applications to charity and Ricardian equivalence," *Journal of Political Economy*, 1989, 97 (6), 1447–1458.
- , "Impure altruism and donations to public goods: A theory of warm-glow giving," *The Economic Journal*, 1990, 100 (401), 464–477.
- , **Deniz Aydin, Blake Barton, B. Douglas Bernheim, and Jeffrey Naecker**, "When Fair Isn't Fair: Understanding Choice Reversals Involving Social Preferences," *Journal of Political Economy*, 2020, 128 (5), 1673–1711.
- Angrist, Joshua D and Jörn-Steffen Pischke**, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2009.
- Appiah, Kwame Anthony**, *Experiments in ethics*, Harvard University Press, 2010.
- Arrieta, Gonzalo and Lukas Bolte**, "What You Don't Know May Hurt You: Preferences Over Mental And External States," 2023.

- Baicker, Katherine, Sarah Taubman, Heidi Allen, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Eric Schneider, Bill Wright, Alan Zaslavsky, and Amy Finkelstein**, “The Oregon experiment—effects of Medicaid on clinical outcomes,” *New England Journal of Medicine*, 2013, 368 (18), 1713–1722.
- Bartling, Björn, Alexander W Cappelen, Henning Hermes, and Bertil Tungodden**, “Free to fail? Paternalistic preferences in the United States,” CESifo Working Paper no. 10441 2023.
- , **Ernst Fehr, and Holger Herz**, “The intrinsic value of decision rights,” *Econometrica*, 2014, 82 (6), 2005–2039.
- Becker, Gary S**, “A Theory of Social Interactions,” *Journal of Political Economy*, 1974, 82 (6), 1063–1093.
- Bénabou, Roland and Jean Tirole**, “Identity, morals, and taboos: Beliefs as assets,” *The Quarterly Journal of Economics*, 2011, 126 (2), 805–855.
- , **Armin Falk, and Luca Henkel**, “Ends versus means: Kantians, utilitarians, and moral decisions,” NBER Working Paper no. 32073, 2024.
- , —, —, and **Jean Tirole**, “Eliciting moral preferences under image concerns: Theory and experiment,” 2023.
- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo**, “quanteda: An R package for the quantitative analysis of textual data,” *Journal of Open Source Software*, 2018, 3 (30), 774.
- Bernheim, B Douglas, Kristy Kim, and Dmitry Taubinsky**, “Welfare and the Act of Choosing,” NBER Working Paper no. 32200, 2024.
- Björkegren, Daniel, Joshua E Blumenstock, and Samsun Knight**, “(Machine) Learning What Policies Value,” 2022.
- Bó, Pedro Dal**, “Experimental evidence on the workings of democratic institutions,” *Institutions, property rights, and economic growth: The legacy of Douglass North*, 2014, 266.
- , **Andrew Foster, and Louis Putterman**, “Institutions and behavior: Experimental evidence on the effects of democracy,” *American Economic Review*, 2010, 100 (5), 2205–2229.
- Bobadilla-Suarez, Sebastian, Cass R Sunstein, and Tali Sharot**, “The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses,” *Journal of Risk and Uncertainty*, 2017, 54, 187–202.
- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer**, “Salience,” *Annual Review of Economics*, 2022, 14 (1), 521–544.
- Bourguignon, François and Amedeo Spadaro**, “Tax–benefit revealed social preferences,” *The Journal of Economic Inequality*, 2012, 10, 75–108.
- Bursztyn, Leonardo, Stefano Fiorin, Daniel Gottlieb, and Martin Kanz**, “Moral incentives in credit card debt repayment: Evidence from a field experiment,” *Journal of Political Economy*, 2019, 127 (4), 1641–1683.
- Capozza, Francesco and Krishna Srinivasan**, “Who Should Get Money? Estimating Welfare Weights in the US,” 2024.
- Cappelen, Alexander W, Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden**, “The pluralism of fairness ideals: An experimental approach,” *American Economic Review*, 2007, 97 (3), 818–827.
- , **Cornelius Cappelen, and Bertil Tungodden**, “Second-best fairness: The trade-off between false positives and false negatives,” *American Economic Review*, 2023, 113 (9), 2458–2485.
- , **Erik Ø Sørensen, and Bertil Tungodden**, “Responsibility for what? Fairness and individual

- responsibility," *European Economic Review*, 2010, 54 (3), 429–441.
- , **James Konow, Erik Ø Sørensen, and Bertil Tungodden**, "Just luck: An experimental study of risk-taking and fairness," *American Economic Review*, 2013, 103 (4), 1398–1413.
- , **Ranveig Falch, and Bertil Tungodden**, "Fair and unfair income inequality," *Handbook of Labor, Human Resources and Population Economics*, 2020, pp. 1–25.
- Caspi, Aviv and Charlie Rafkin**, "Legal Assistance for Evictions: Impacts and Demand," 2023.
- Charité, Jimmy, Raymond Fisman, Ilyana Kuziemko, and Kewei Zhang**, "Reference points and redistributive preferences: Experimental evidence," *Journal of Public Economics*, 2022, 216, 104761.
- Charness, Gary, Uri Gneezy, and Vlastimil Rasocha**, "Experimental methods: Eliciting beliefs," *Journal of Economic Behavior & Organization*, 2021, 189, 234–256.
- Chen, Daniel L and Martin Schonger**, "Social preferences or sacred values? Theory and evidence of deontological motivations," *Science Advances*, 2022, 8 (19), eabb3925.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins**, "Double/debiased machine learning for treatment and structural parameters," *The Econometrics Journal*, 2018, 21 (1), C1–C68.
- Clarke, Damian, Joseph P Romano, and Michael Wolf**, "The Romano–Wolf multiple-hypothesis correction in Stata," *The Stata Journal*, 2020, 20 (4), 812–843.
- Cohn, Alain, Michel André Maréchal, David Tannenbaum, and Christian Lukas Zünd**, "Civic honesty around the globe," *Science*, 2019, 365 (6448), 70–73.
- Currie, Janet and Firouz Gahvari**, "Transfers in cash and in-kind: Theory meets the data," *Journal of Economic Literature*, 2008, 46 (2), 333–383.
- Dana, Jason, Roberto A Weber, and Jason Xi Kuang**, "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, 2007, 33, 67–80.
- Danz, David, Lise Vesterlund, and Alistair J Wilson**, "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 2022, 112 (9), 2851–2883.
- d'Aspremont, Claude and Louis Gevers**, "Social welfare functionals and interpersonal comparability," *Handbook of social choice and welfare*, 2002, 1, 459–541.
- Doris, John, Stephen Stich, Jonathan Phillips, and Lachlan Walmsley**, "Moral psychology: Empirical approaches," 2006.
- Druckman, James N and Arthur Lupia**, "Preference formation," *Annual Review of Political Science*, 2000, 3 (1), 1–24.
- Enke, Benjamin**, "Moral values and voting," *Journal of Political Economy*, 2020, 128 (10), 3679–3729.
- , "Moral boundaries," *Annual Review of Economics*, 2024, 16 (1), 133–157.
- Exley, Christine L and Judd B Kessler**, "Equity concerns are narrowly framed," *American Economic Journal: Microeconomics*, 2024, 16 (2), 147–179.
- Fehr, Ernst and Klaus M Schmidt**, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, 114 (3), 817–868.
- and —, "The economics of fairness, reciprocity and altruism—experimental evidence and new theories," *Handbook of the economics of giving, altruism and reciprocity*, 2006, 1, 615–691.
- , **Holger Herz, and Tom Wilkening**, "The lure of authority: Motivation and incentive effects of power," *American Economic Review*, 2013, 103 (4), 1325–1359.
- Ferrario, Beatrice and Stefanie Stantcheva**, "Eliciting people's first-order concerns: Text analysis of open-ended survey questions," *AEA Papers and Proceedings*, 2022, 112, 163–169.
- Fischbacher, Urs and Franziska Föllmi-Heusi**, "Lies in disguise—an experimental study on

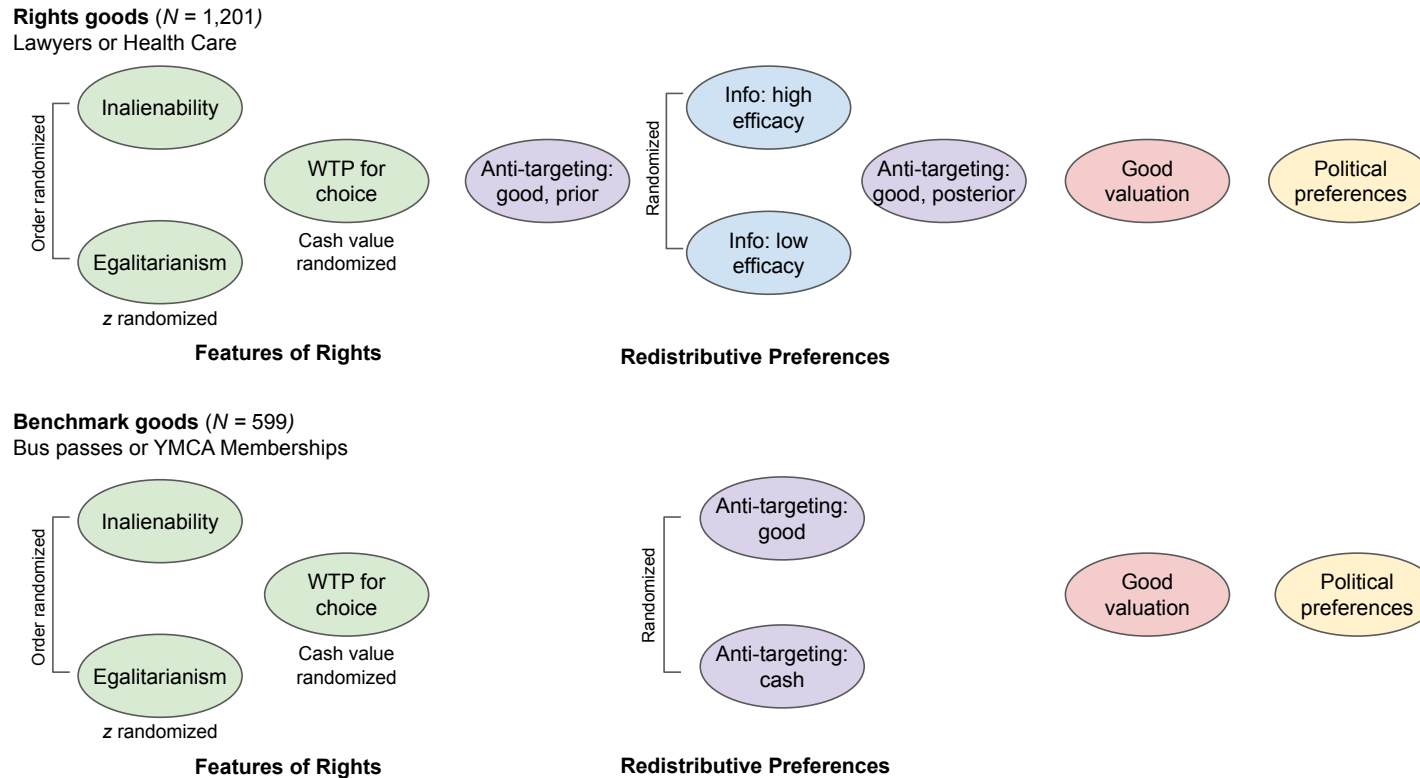
- cheating," *Journal of the European Economic Association*, 2013, 11 (3), 525–547.
- Fisman, Raymond, Shachar Kariv, and Daniel Markovits**, "Individual Preferences for Giving," *American Economic Review*, December 2007, 97 (5), 1858–1876.
- Fleurbaey, Marc**, "Economic theories of justice," *Annual Review of Economics*, 2019, 11 (1), 665–684.
- **and François Maniquet**, "Optimal income taxation theory and principles of fairness," *Journal of Economic Literature*, 2018, 56 (3), 1029–1079.
- **, Bertil Tungodden, and Howard F Chang**, "Any non-welfarist method of policy assessment violates the Pareto principle: A comment," *Journal of Political Economy*, 2003, 111 (6), 1382–1385.
- Foot, Philippa**, *The problem of abortion and the doctrine of double effect*, Vol. 5, Oxford, 1967.
- Freundt, Jana, Holger Herz, and Leander Kopp**, "Intrinsic preferences for choice autonomy," CESifo Working Paper no. 10342, 2023.
- Gneezy, Uri, Agne Kajackaite, and Joel Sobel**, "Lying aversion and the size of the lie," *American Economic Review*, 2018, 108 (2), 419–453.
- **, Stephan Meier, and Pedro Rey-Biel**, "When and why incentives (don't) work to modify behavior," *Journal of Economic Perspectives*, 2011, 25 (4), 191–210.
- Goodenough, Oliver R and Kristin Prehn**, "A neuroscientific approach to normative judgment in law and justice.," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2004, 359 (1451), 1709.
- Haaland, Ingar, Christopher Roth, and Johannes Wohlfart**, "Designing information provision experiments," *Journal of Economic Literature*, 2023, 61 (1), 3–40.
- **, —, Stefanie Stantcheva, and Johannes Wohlfart**, "Measuring what is top of mind," NBER Working Paper no. 32421, 2024.
- Hausman, Daniel M and Michael S McPherson**, "Taking ethics seriously: economics and contemporary moral philosophy," *Journal of Economic Literature*, 1993, 31 (2), 671–731.
- Hayek, Friedrich A**, "The constitution of liberty," *Chicago UP*, 1960.
- Hendren, Nathaniel**, "Measuring economic efficiency using inverse-optimum weights," *Journal of Public Economics*, 2020, 187, 104198.
- Holmes, Stephen and Cass R Sunstein**, *The cost of rights: why liberty depends on taxes*, WW Norton & Company, 2000.
- Hvidberg, Kristoffer B, Claus T Kreiner, and Stefanie Stantcheva**, "Social positions and fairness views on inequality," *Review of Economic Studies*, 2023, 90 (6), 3083–3118.
- Kaplow, Louis**, "Optimal income taxation," *Journal of Economic Literature*, 2024, 62 (2), 637–738.
- **and Steven Shavell**, "Any non-welfarist method of policy assessment violates the Pareto principle," *Journal of Political Economy*, 2001, 109 (2), 281–286.
- Knobe, Joshua, Wesley Buckwalter, Shaun Nichols, Philip Robbins, Hagop Sarkissian, and Tamler Sommers**, "Experimental philosophy," *Annual review of psychology*, 2012, 63, 81–99.
- Konow, James**, "Fair and square: the four sides of distributive justice," *Journal of Economic Behavior & Organization*, 2001, 46 (2), 137–164.
- Kőszegi, Botond and Matthew Rabin**, "A model of reference-dependent preferences," *The Quarterly Journal of Economics*, 2006, 121 (4), 1133–1165.
- Lenk, Alexandr**, "Do People Honor Others' Autonomy Out of Principle? An Experimental Investigation."
- Levitt, Steven D and John A List**, "What do laboratory experiments measuring social preferences reveal about the real world?," *Journal of Economic Perspectives*, 2007, 21 (2), 153–174.

- Liscow, Zachary and Abigail Pershing**, "Why is so much redistribution in-kind and not in cash? Evidence from a survey experiment," *National Tax Journal*, 2022, 75 (2), 313–354.
- List, John A.**, "On the interpretation of giving in dictator games," *Journal of Political Economy*, 2007, 115 (3), 482–493.
- Lockwood, Benjamin B and Matthew Weinzierl**, "Positive and normative judgments implicit in US tax policy, and the costs of unequal growth and recessions," *Journal of Monetary Economics*, 2016, 77, 30–47.
- Luttmer, Erzo F. P. and Monica Singhal**, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, February 2011, 3 (1), 157–79.
- Obama, Barack**, "Remarks by the President on the Affordable Care Act," *The White House*, 2013.
- Ottati, Victor C.**, "Determinants of political judgments: The joint influence of normative and heuristic rules of inference," *Political Behavior*, 1990, 12, 159–179.
- Polman, Evan**, "Self–other decision making and loss aversion," *Organizational Behavior and Human Decision Processes*, 2012, 119 (2), 141–150.
- Rabin, Matthew**, "Incorporating fairness into game theory and economics," *American Economic Review*, 1993, pp. 1281–1302.
- Rawls, John**, "A theory of justice," *Cambridge (Mass.)*, 1971.
- Roth, Christopher and Johannes Wohlfart**, "Experienced inequality and preferences for redistribution," *Journal of Public Economics*, 2018, 167, 251–262.
- Ruggles, Steven, Sarah Flood, Matthew Sobek, Danika Brockman, Grace Cooper, Stephanie Richards, and Megan Schouweiler**, "IPUMS USA: Version 13.0 [dataset]," 2023.
- Saccardo, Silvia and Marta Serra-Garcia**, "Enabling or limiting cognitive flexibility? Evidence of demand for moral commitment," *American Economic Review*, 2023, 113 (2), 396–429.
- Saez, Emmanuel and Stefanie Stantcheva**, "Generalized social marginal welfare weights for optimal tax theory," *American Economic Review*, 2016, 106 (01), 24–45.
- Sen, Amartya**, *Commodities and Capabilities*, North-Holland, 1985.
- Sheffrin, Steven M.**, *Tax fairness and folk justice*, Cambridge University Press, 2013.
- Sher, Itai**, "Generalized social marginal welfare weights imply inconsistent comparisons of tax policies," *American Economic Review*, 2024, 114 (11), 3551–3577.
- Simon, Herbert A.**, "A behavioral model of rational choice," *The Quarterly Journal of Economics*, 1955, pp. 99–118.
- Stantcheva, Stefanie**, "Understanding tax policy: How do people reason?," *The Quarterly Journal of Economics*, 2021, 136 (4), 2309–2369.
- , "Understanding of trade," NBER Working Paper no. 30040, 2022.
- Sullivan, Colin D.**, "Eliciting preferences over life and death: Experimental evidence from organ transplantation," 2023.
- Suzumura, Kotaro**, "Welfarism, individual rights, and procedural fairness," in "Handbook of social choice and welfare," Vol. 2, Elsevier, 2011, pp. 605–685.
- Thomson, Judith Jarvis**, "Killing, Letting Die, and the Trolley Problem," *The Monist*, 1976, 59 (2), 204–217.
- Tobin, James**, "On limiting the domain of inequality," *The Journal of Law and Economics*, 1970, 13 (2), 263–277.
- Walzer, Michael**, *Spheres of justice: A defense of pluralism and equality* 1983.
- Weinzierl, Matthew**, "The promise of positive optimal taxation: normative diversity and a role

- for equal sacrifice," *Journal of Public Economics*, 2014, 118, 128–142.
- , "A welfarist role for nonwelfarist rules: an example with envy," NBER Working Paper no. 23587, 2017.
- Yang, Jeffrey**, "On the Decision-Relevance of Subjective Beliefs," *Available at SSRN 4425080*, 2023.

7 Figures

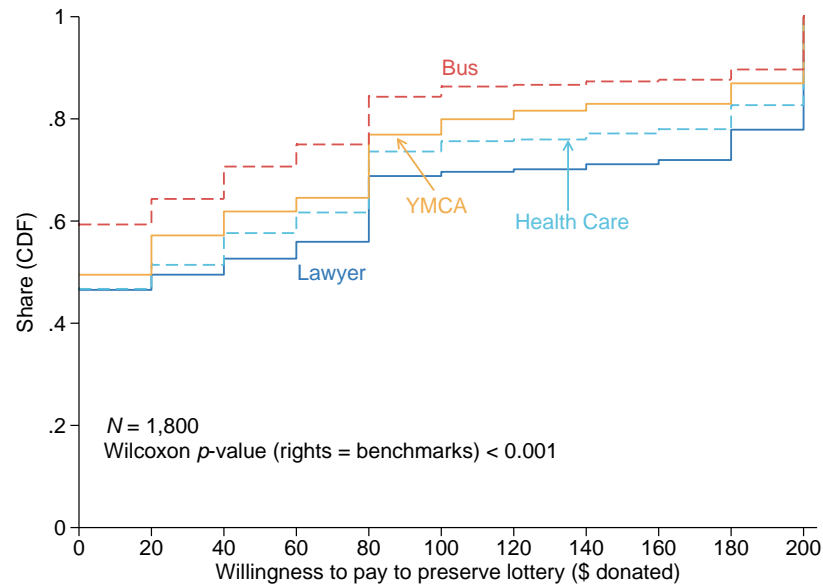
Figure 1: Survey Flow



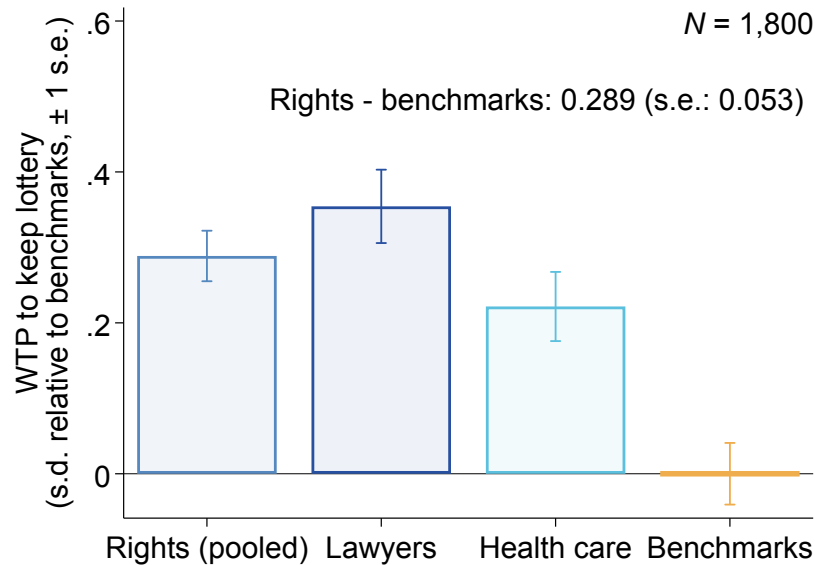
Note: This figure presents the flow of the survey, differentiating between those who saw rights goods (two-thirds of the sample) and those who saw benchmark goods (the remaining one-third). Within each arm, participants were randomly shown one of the two potential goods. After introductory information, participants first completed experiments related to features of rights. The order of *Inalienability* (Experiment 1) and *Egalitarianism* (Experiment 3) was randomized, followed by *Willingness to Pay for Choice* (Experiment 2). Next, participants answered questions about *Anti-Targeting*; those assigned to rights goods also saw an information treatment. Finally, we elicited participants WTP for the good they were assigned, demographic information, and political preferences. Participants were required to pass two out of three attention checks.

Figure 2: Experiment 1: Inalienability

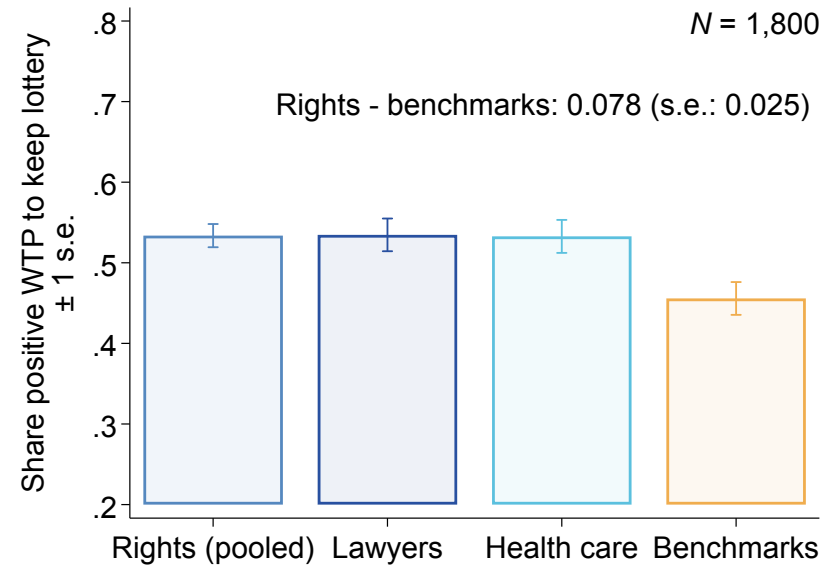
(a) CDF of WTP to keep lottery



(b) WTP to keep lottery

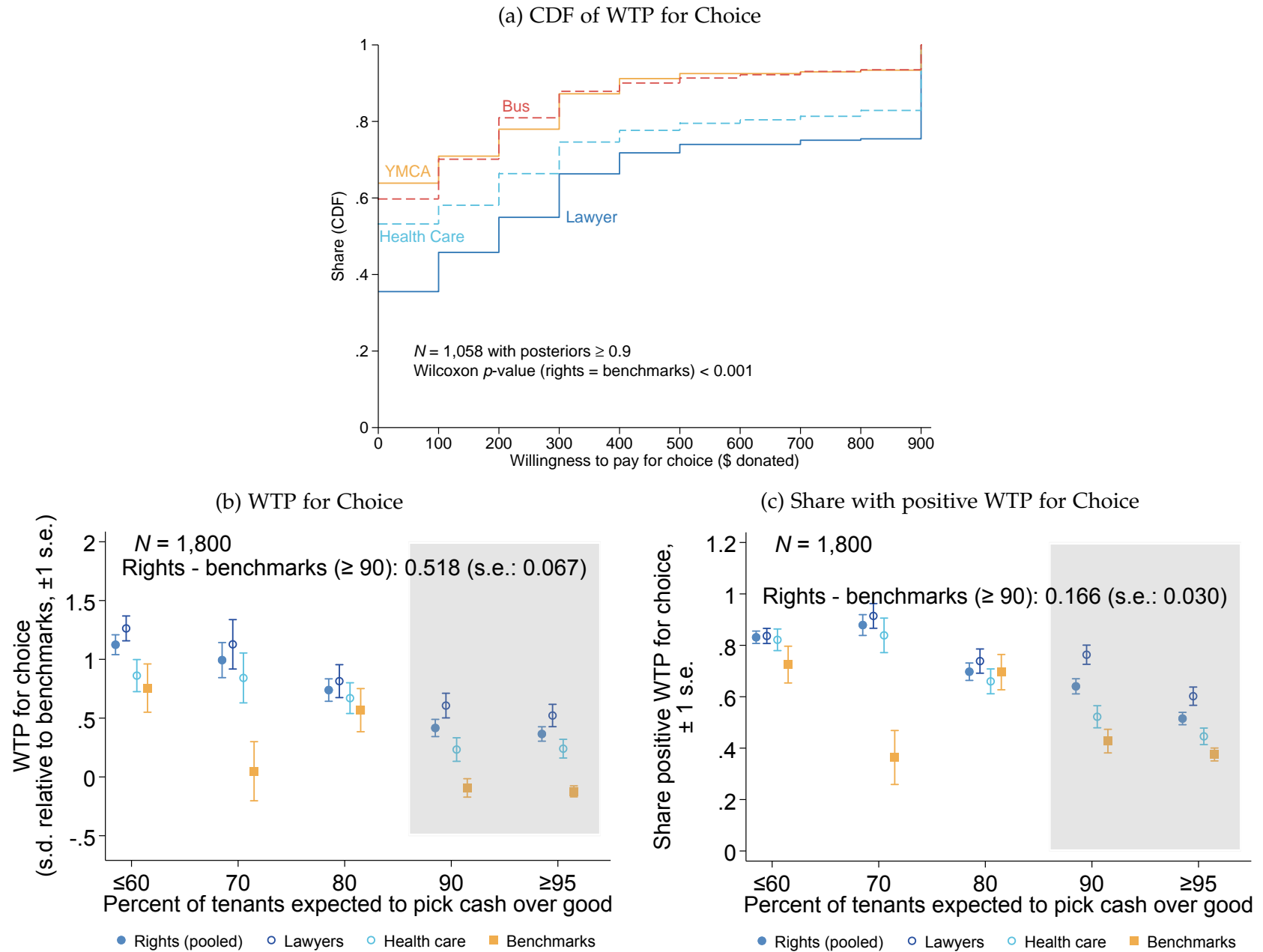


(c) Share with positive WTP to keep lottery



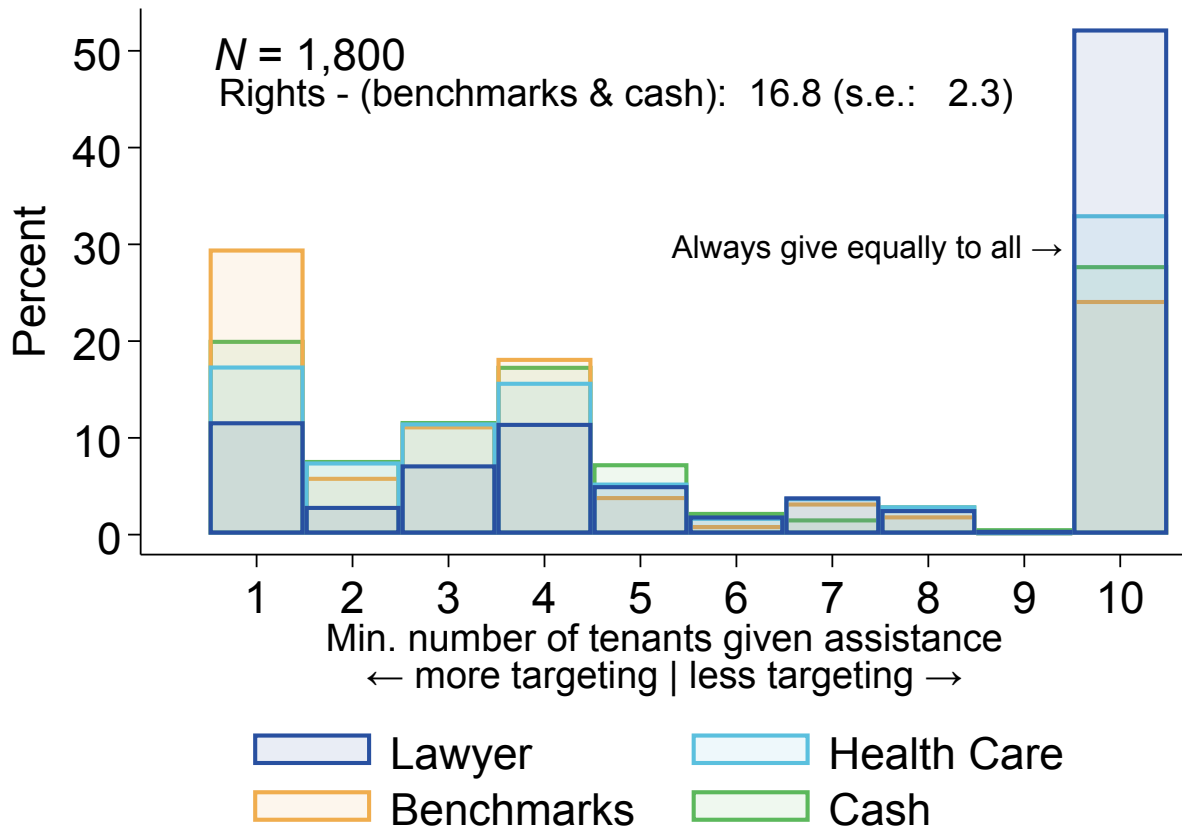
Note: This figure shows participants' decisions about keeping or rerunning a lottery in order to save money for future programs in Experiment 1 (Section 3.1). In panel A, we show the CDF of the WTP to keep the initial lottery result for each good. In panel B, we show the same WTP in standard deviations relative to the benchmark goods. The far left blue bar shows both rights goods pooled, while the next two bars show results for lawyers and health care disaggregated. Panel C shows the share of respondents with positive WTP to keep the lottery by good. Both panels include ± 1 standard errors. See Table A3 for detailed regression results.

Figure 3: Experiment 2 (Choice)



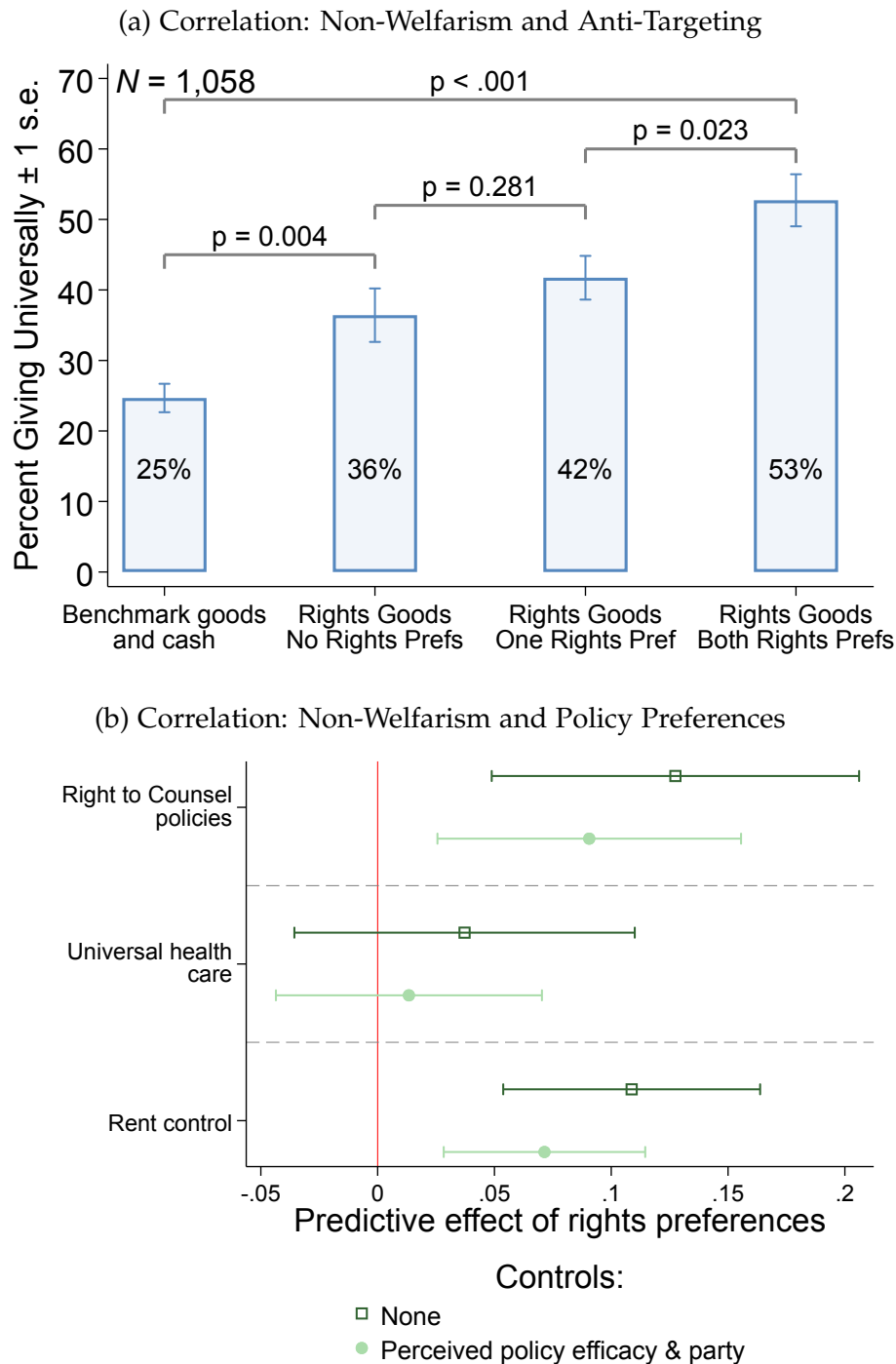
Note: This figure shows participants' decisions about giving tenants a choice between cash and a good, versus saving money for future programs in Experiment 2 (Section 3.2). Panel A shows the CDF of WTP to give tenants a choice for each good among participants whose posterior belief about the percent of tenants expected to pick cash over the good was at least 90%. Panel B plots participants' WTP in standard deviations relative to the benchmark goods (± 1 s.e.). We show this for different posterior beliefs about the percent of tenants expected to pick cash over the good. Belief bins are posterior beliefs $< 65, 65 - 74, 75 - 84, 85 - 94, 95 - 100$. The shaded gray area emphasizes those with high posteriors ($\geq 90\%$). Panel C plots the same for the share with positive WTP. See Table A7 for detailed regression results.

Figure 4: Experiment 4 (Anti-Targeting)



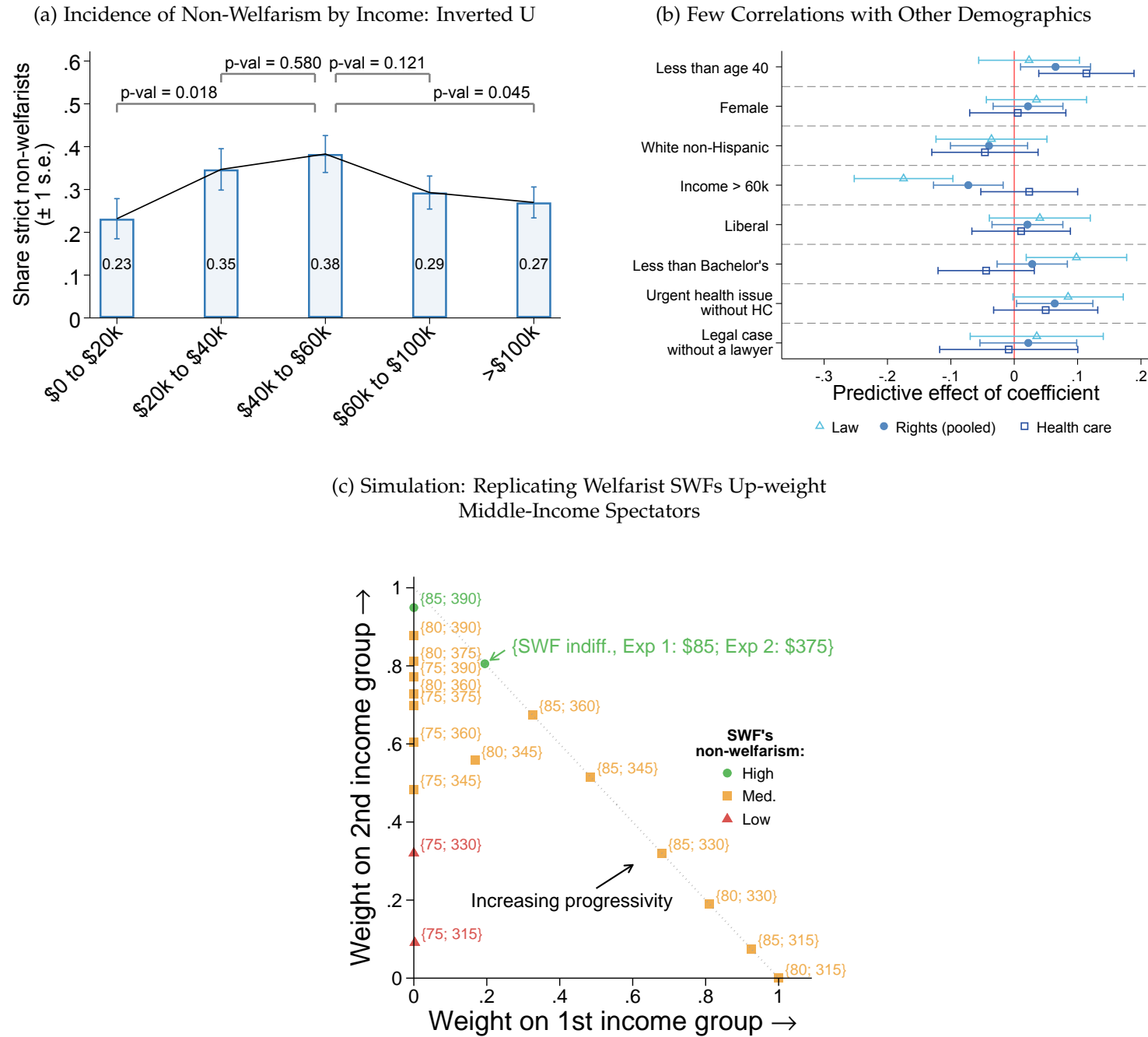
Note: This figure shows the distribution of the minimum number of tenants given assistance in Experiment 4 (Section 4.1). For example, if a participant chose to give two tenants the good and cash, but to give everyone a lawyer rather than giving one tenant the good and cash, the minimum number of tenants given assistance is two. If the minimum is 10, the participant always chose to give the good to everyone. The blue series show distributions for lawyers and health care, the orange series shows the distribution for benchmark goods, and the green series shows the distribution for cash; half of participants who saw benchmark goods throughout the survey were asked about distributing cash in this experiment. See Table A16 for detailed regression results.

Figure 5: Policy Implications: Non-Welfarism and Redistribution



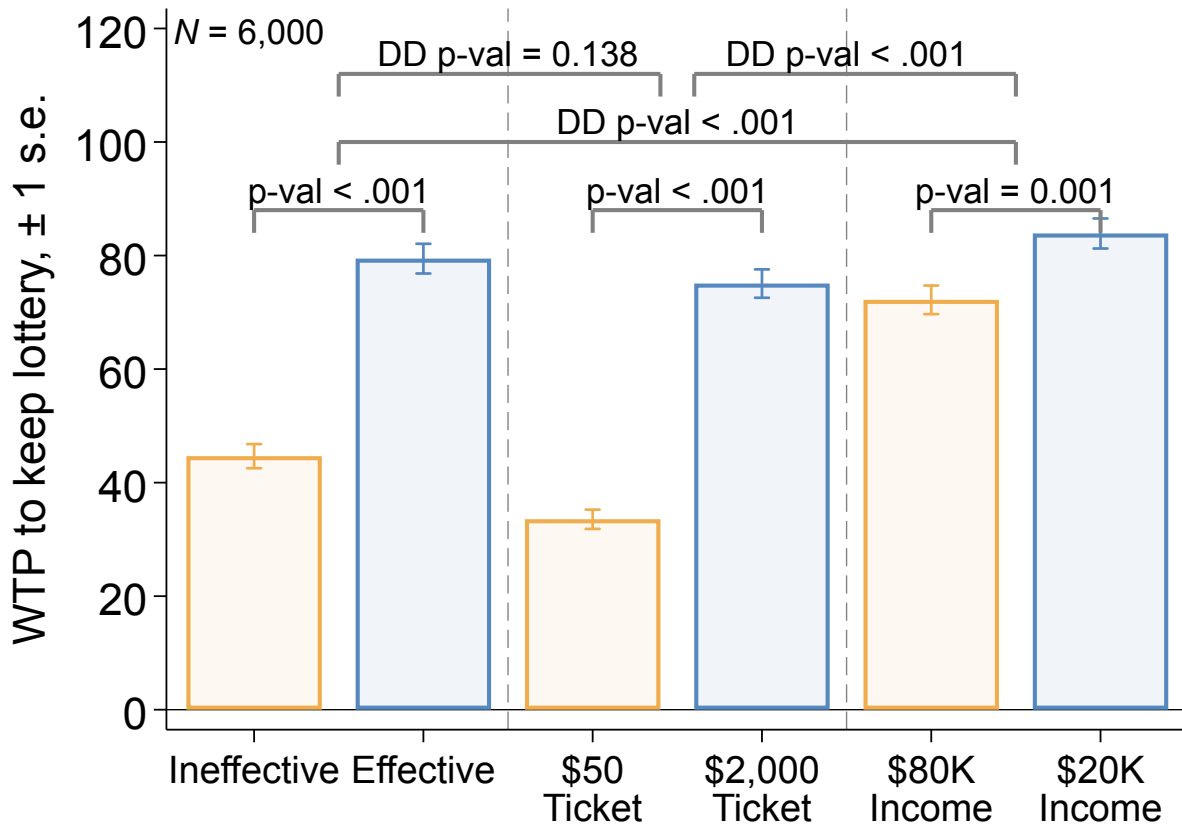
Note: Panel A, restricted to those with posteriors ≥ 0.9 , shows the share of participants who give universally in Experiment 4 for participants who saw benchmark goods and those who saw rights goods, broken down by whether they had positive WTP in Experiments 1 and 2. Spectators have “rights preferences” if they have positive WTP in a given experiment. Panel B shows the predictive effect of non-welfarist preferences—having positive WTP in both Experiment 1 and 2—on support for specific policies. We elicit beliefs about policy efficacy by asking whether people in the U.S. would be on average worse or better off with the policy. We show results for *Right to Counsel policies* among those who did the experiment with lawyers, *Universal health care* among those who did the experiment with health care, and *Rent control* for all who saw rights goods. Estimates shown with ± 1.96 standard errors.

Figure 6: Incidence of Non-Welfarism



Note: Panel A shows prevalence of non-welfarist preferences by income group (defined as having positive WTPs in both Experiments 1 and 2). See quantification in Table A21. In Panel B, we regress the non-welfarism outcome on demographics. Estimates are shown with with ± 1.96 standard errors. Panel C shows results from Equations (17)–(19). Each point represents the weights placed on the first or second tercile household, for a given “normatively correct” WTP in Experiments 1 or 2. The normatively correct WTP is shown in braces for other dots. The results show that, as the normatively correct WTP grows (meaning the SWF is more non-welfarist), the SWF necessarily places less value on the bottom-tercile household (as the weights move up and to the left). In this exercise, we use probabilities from the whole sample, even though C_i values come from the sample with posteriors about choosing cash that exceed 0.9. The Social Welfare Function’s non-welfarism is “high” if $I^* = 85$ and $C^* = 375$ or $C^* = 390$; “low” if $I^* = 65$ and $C^* = 315$ or $C^* = 330$; and “medium” otherwise.

Figure 7: Mechanisms: Stakes



Note: This figure shows participants' WTP to preserve lottery results in an extension of Experiment 1 (Section 5). In orange are low-stakes variations, while high-stakes variations are in blue. The left two bars compare WTP in a lottery allocating either an ineffective vs. effective lawyer to a tenant facing eviction. The center two bars compare WTP in a lottery allocating a lawyer for a speeding ticket case with either a \$50 or \$2,000 fine. Finally, the right two bars compare WTP in a lottery allocating a lawyer for a tenant facing eviction to a tenant with either \$80,000 or \$20,000 in annual income. See Table A27 for detailed regression results. Estimates shown with ± 1 standard error.

Table 1: Demographics and Balance

| | (1) U.S. | (2) Experimental sample | (3) Rights -Benchmarks |
|--------------------------------|----------------|-------------------------------|------------------------------|
| White non-Hispanic | 0.66 [0.47] | 0.71 [0.45] | -0.01 (0.02) |
| Income > 60k | 0.24 [0.43] | 0.48 [0.50] | -0.03 (0.02) |
| Less than Bachelor's | 0.67 [0.47] | 0.45 [0.50] | -0.00 (0.02) |
| Female | 0.51 [0.50] | 0.51 [0.50] | -0.02 (0.02) |
| Less than age 40 | 0.33 [0.47] | 0.56 [0.50] | 0.00 (0.02) |
| Liberal | | 0.57 [0.49] | 0.05 (0.02) |
| Legal case without a lawyer | | 0.16 [0.36] | 0.01 (0.03) |
| Urgent health issue without HC | | 0.31 [0.46] | -0.02 (0.03) |
| <i>F</i> -statistic | | | .996 |
| <i>p</i> -value | | | 0.437 |
| Observations | 2,624,206 | 1,800 | 1,800 |

Note: This table shows the composition of our experimental sample relative to all U.S. adults (18+, a requirement on Prolific) in the 2021 ACS (Ruggles et al., 2023). Column (3) shows differences in participants assigned to lawyers and health care compared to those assigned to benchmarks from an OLS regression. The *F*-statistic is from a joint test of significance for the listed demographic variables. Brackets show standard deviations. Parentheses show robust standard errors.

Table 2: Prevalence of Non-Welfarism Overall

| | (1) Non-zero WTP in Experiment 1 | (2) Non-zero WTP in Experiment 2 | (3) Non-zero WTP in ≥ 1 Experiment | (4) Non-zero WTP in Both Experiments |
|----------------------|--|--|---|--|
| Rights good (= 1) | 0.078 (0.025) [0.002] | 0.166 (0.030) [0.000] | 0.125 (0.029) [0.000] | 0.121 (0.026) [0.000] |
| Mean (benchmarks) | 0.456 | 0.382 | 0.605 | 0.186 |
| Posterior ≥ 0.9 | | ✓ | ✓ | ✓ |
| Observations | 1,800 | 1,058 | 1,058 | 1,058 |

Note: This table shows the effect of doing Experiments 1–2 with rights goods on the prevalence of non-welfarist choices. Each entry represents the difference in non-welfarist preferences when Spectators do the experiment with rights goods, compared with benchmark goods. One can obtain levels by adding the difference to the mean among benchmarks. In columns (1) and (2), the outcome is an indicator for having non-zero WTP in Experiment 1 (inalienability) and 2 (choice), respectively. In column (3), the outcome is an indicator for having non-zero WTP in either Experiment 1 or 2. In column (4), the outcome is an indicator for having non-zero WTP in both experiments. In columns (2)–(4), we restrict to participants with posterior beliefs in Experiment 2 that exceed 0.9, so that our test in that experiment is well-posed. Parentheses show robust standard errors. Brackets show p -values.

Table 3: Tests of Correlations between Experiments 1–2 and Experiment 4

| | Dep. Var.: Universal (= 1) | | | |
|--------------------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|
| | (1) | (2) | (3) | (4) |
| Panel A. Rights goods (pooled) | | | | |
| WTP for inalienability (s.d.) | 0.031 (0.012) [0.013] | | | 0.017 (0.018) [0.331] |
| WTP for choice (s.d.) | | 0.089 (0.015) [0.000] | | 0.085 (0.016) [0.000] |
| Mean WTP (s.d.) | | | 0.107 (0.020) [0.000] | |
| Panel B. Lawyers | | | | |
| WTP for inalienability (s.d.) | 0.031 (0.017) [0.067] | | | -0.005 (0.026) [0.852] |
| WTP for choice (s.d.) | | 0.109 (0.021) [0.000] | | 0.110 (0.024) [0.000] |
| Mean WTP (s.d.) | | | 0.113 (0.027) [0.000] | |
| Panel C. Health care | | | | |
| WTP for inalienability (s.d.) | 0.021 (0.018) [0.238] | | | 0.030 (0.024) [0.215] |
| WTP for choice (s.d.) | | 0.060 (0.022) [0.006] | | 0.056 (0.022) [0.012] |
| Mean WTP (s.d.) | | | 0.087 (0.029) [0.003] | |
| Posterior ≥ 0.9 | | ✓ | ✓ | ✓ |
| Mean | 0.428 | 0.437 | 0.437 | 0.437 |
| Observations | 1,201 | 600 | 600 | 600 |

Note: This table shows the effect of WTP for inalienability (Experiment 1), WTP for choice (Experiment 2), and universal provision in the anti-targeting experiment (Experiment 4). All WTPs are in terms of s.d. relative to the benchmark goods. Mean WTP (s.d.) is an average of the two standardized WTPs. The outcome is an indicator for whether the participant distributed the good universally. The sample is those who did the experiment with rights goods: Panel A shows pooled results, while Panels B and C show lawyers and health care, respectively. Columns (2)-(4), which include measures of WTP from Experiment 2, further restrict to participants with posteriors ≥ 0.9 . Parentheses show robust standard errors. Brackets show p -values.

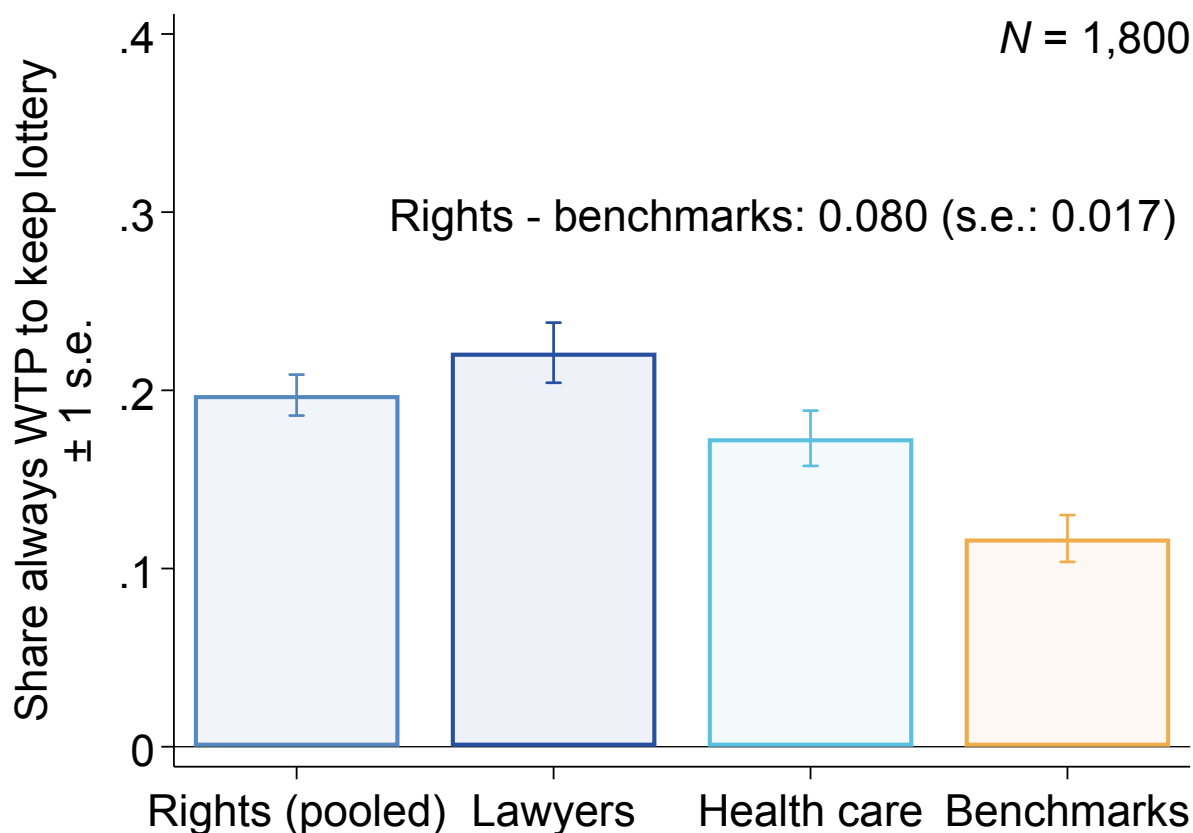
Appendices for Online Publication

Contents

| | | |
|----------|----------------------------|------------|
| A | Additional Figures | 52 |
| B | Additional Tables | 68 |
| C | Proofs | 96 |
| D | Experiment Details | 98 |
| E | Additional Analyses | 103 |
| F | Pre-Registration | 111 |

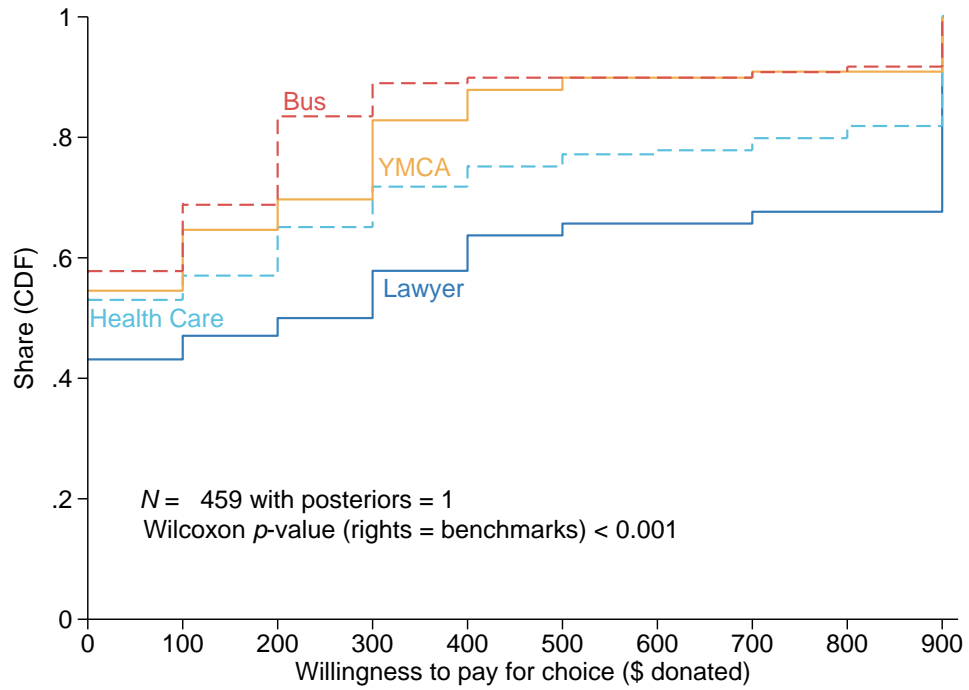
A Additional Figures

Figure A1: Share Always Willing to Pay to Keep Lottery



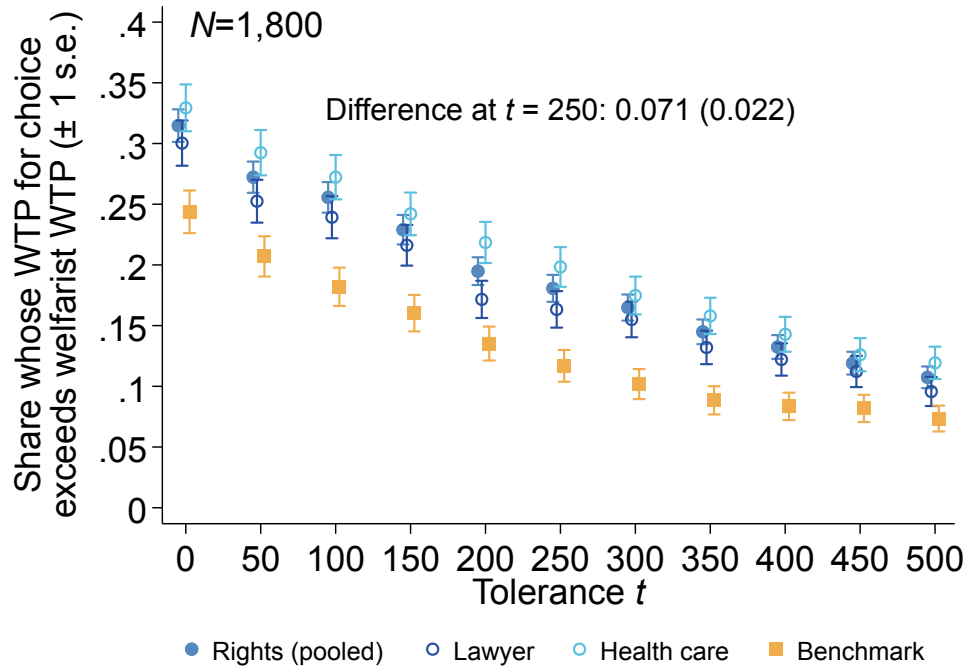
Note: This figure shows participants' decisions about keeping or rerunning a lottery in order to save money for future programs in Experiment 1 (Section 3.1). The y-axis is the share of participants who are willing to pay the maximum value we elicit to keep the lottery. The far left blue bar shows both rights goods pooled, while the next two bars show results for lawyers and health care disaggregated.

Figure A2: CDF of WTP for Choice, Posterior = 1



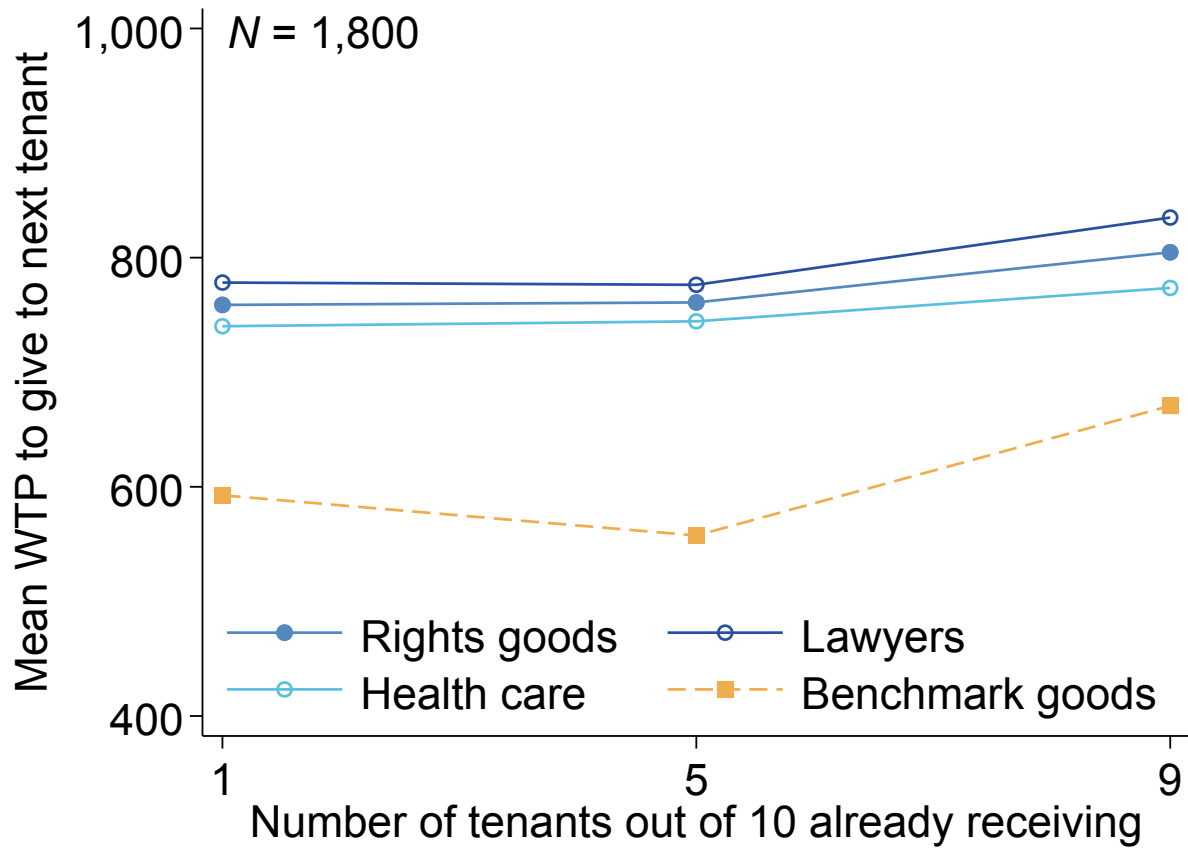
Note: This replicates Panel A of Figure 3, instead restricting to participants with a posterior belief that 100% of tenants would pick cash over the good.

Figure A3: Share with WTP for Choice Exceeding Welfarist WTP



Note: This figure shows the share of participants whose WTP for choice exceeds the secondary measure of non-welfarist value to provide choices in Experiment 2 (Equation 37) as a function of the tolerance t .

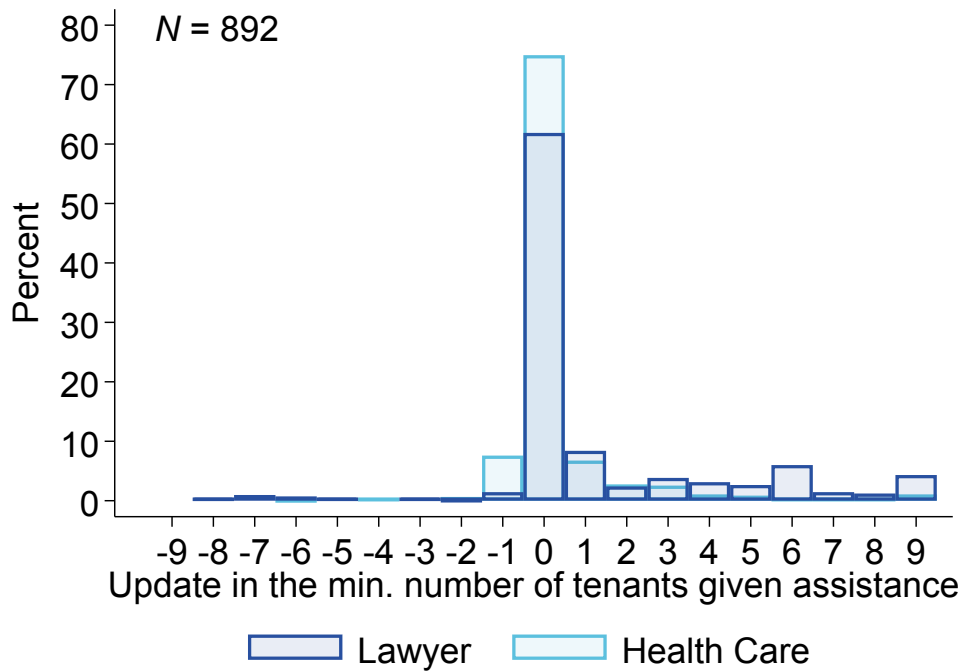
Figure A4: Egalitarianism



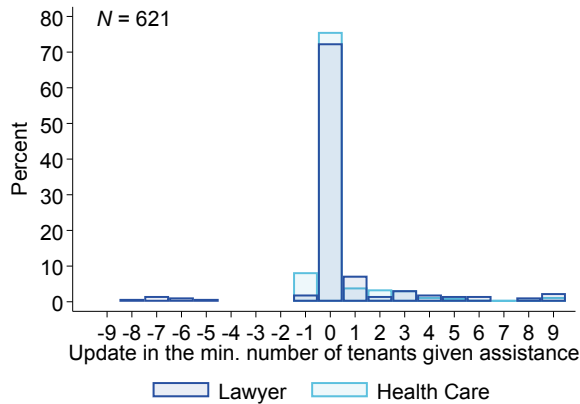
Note: This figure plots mean WTP to give the good to the next tenant in Experiment 3 (Section 3.3). Along the x-axis is the number of tenants out of 10 already assigned to receive the good. Rights goods are presented both pooled and disaggregated by lawyers health care.

Figure A5: Anti-Targeting: Updates After Information Treatment

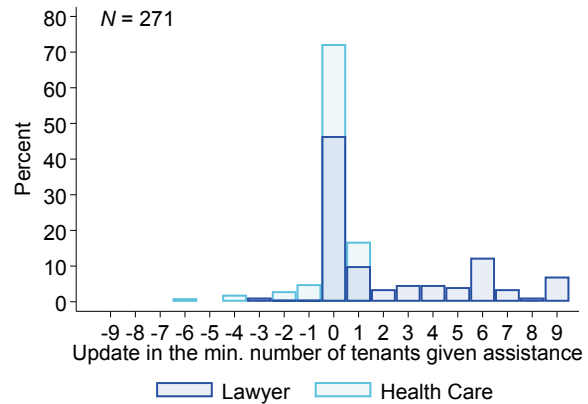
(a) Rights good participants



(b) Information below prior

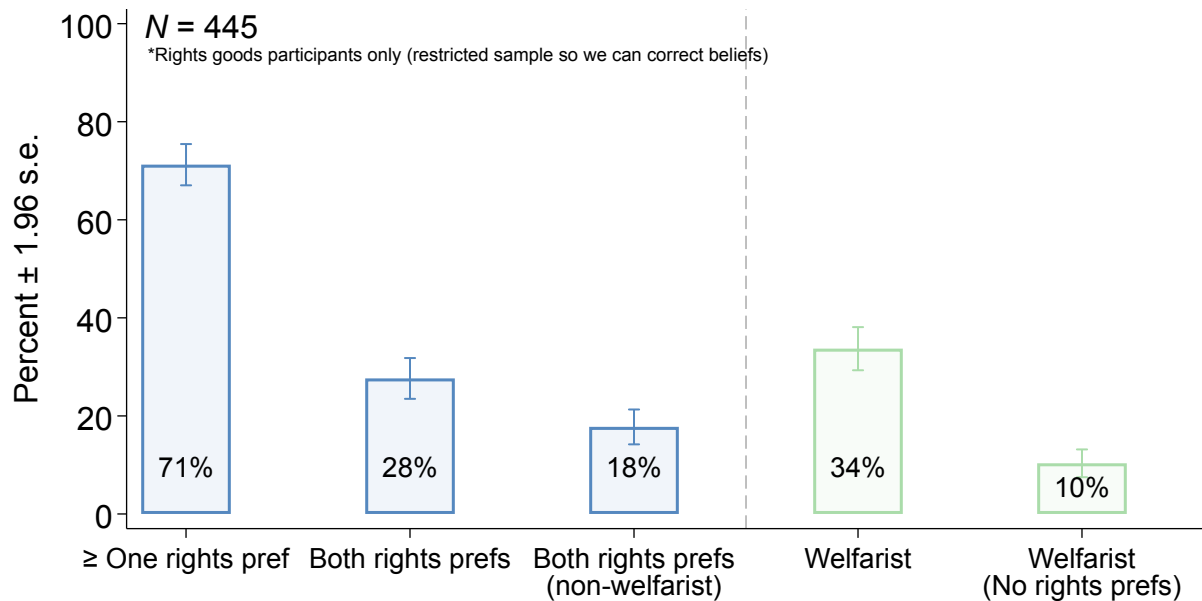


(c) Information above prior



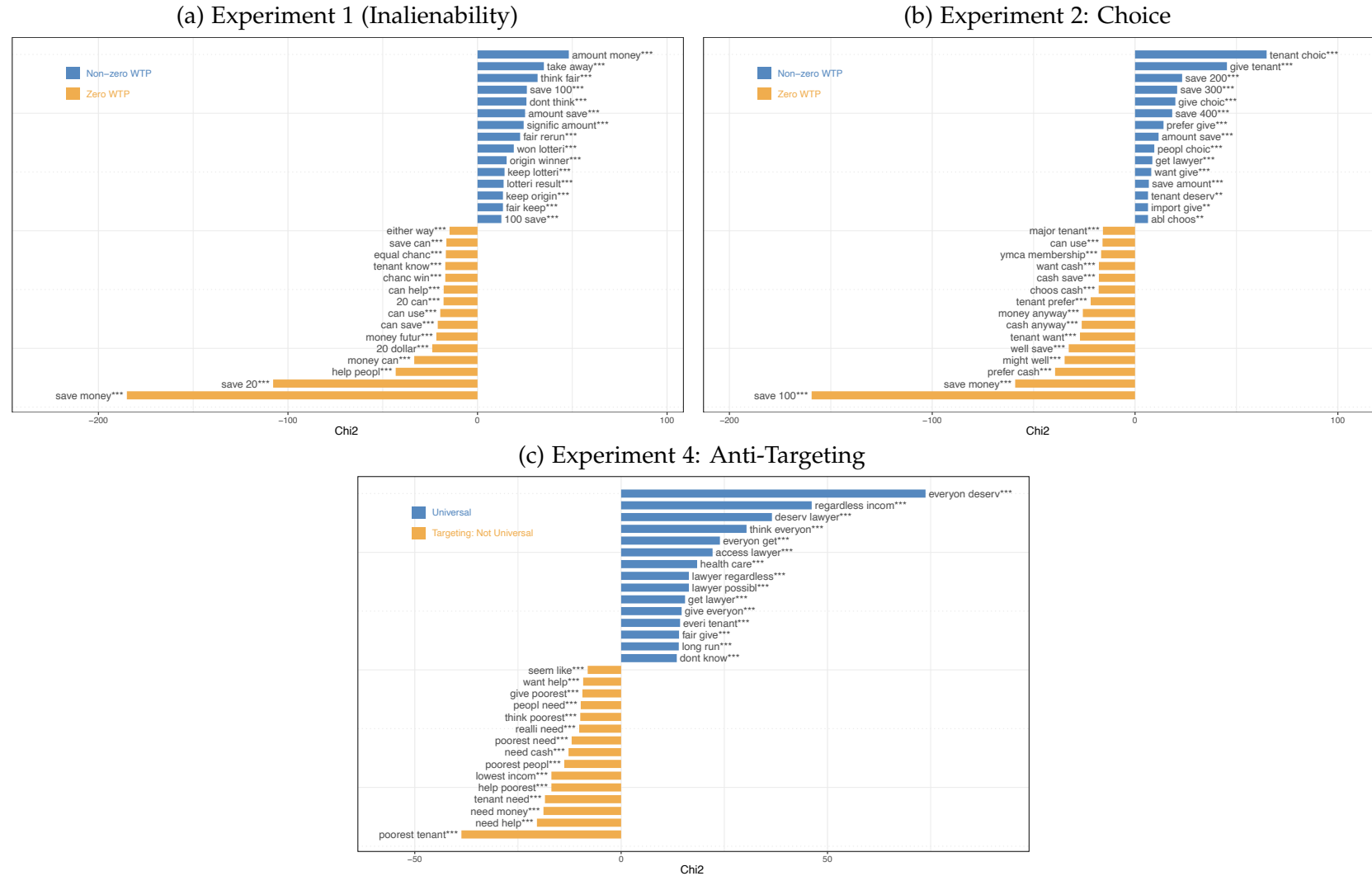
Note: This figure shows the distribution of updates in the minimum number of tenants given assistance (Figure 4) after being shown either a high or low information treatment (Appendix E.5). Only participants in lawyer or health care treatments saw information and were offered the chance to update. We restrict the sample to participants who saw rights goods, and further to those we can classify as welfarist or non-welfarist based on their prior beliefs. This means we exclude those that allocate universally in the original experiment and see information implying lawyers are more effective than their prior belief, and those that target only the poorest tenant and see information implying lawyers are less effective than their prior. Panel (a) includes all participants from this sample, while panels (b) and (c) split the sample by whether the information the participant saw was above or below their prior. Zero indicates the participant declined to update. Positive numbers indicate the participant chose to update in the direction of information shown, relative to their priors.

Figure A6: Prevalence of Non-Welfarism (from Information Experiment)



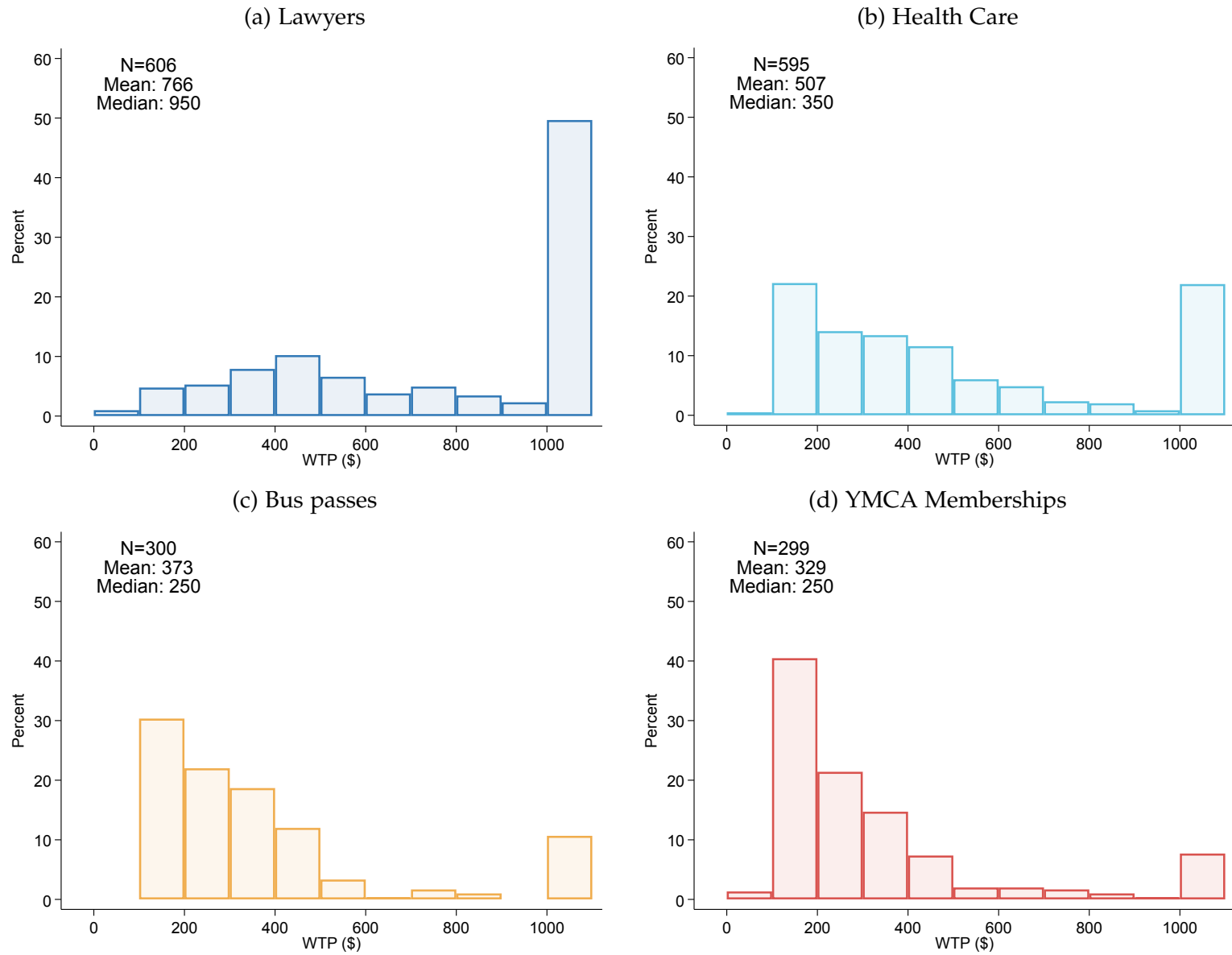
Note: This figure presents the prevalence of non-welfarist preferences (“rights preferences” in the figure label) and welfarism according to several definitions. First, we quantify the percent of participants with non-welfarist preferences as defined. Next, *Both rights prefs (Non-Welfarist)* requires also not updating in the information provision experiment. *Welfarist* participants are those who change their choices about targeting when information about efficacy of lawyers or health care disagrees with their priors. *Welfarist (No rights prefs)* includes the subset of welfarists who do not have positive WTP in either Experiment 1 or 2. The sample is participants who saw rights goods, those with posteriors ≥ 0.9 in Experiment 2, and those whom we could classify as welfarist based on their prior beliefs. Specifically, the last restriction is that we drop people who saw that lawyers/health care were more effective than their prior beliefs, and who gave goods universally; and we drop people who saw that lawyers/health care were less effective than their prior beliefs, and who gave the minimum number of tenants the good.

Figure A7: Keywords from Open-Ended Responses



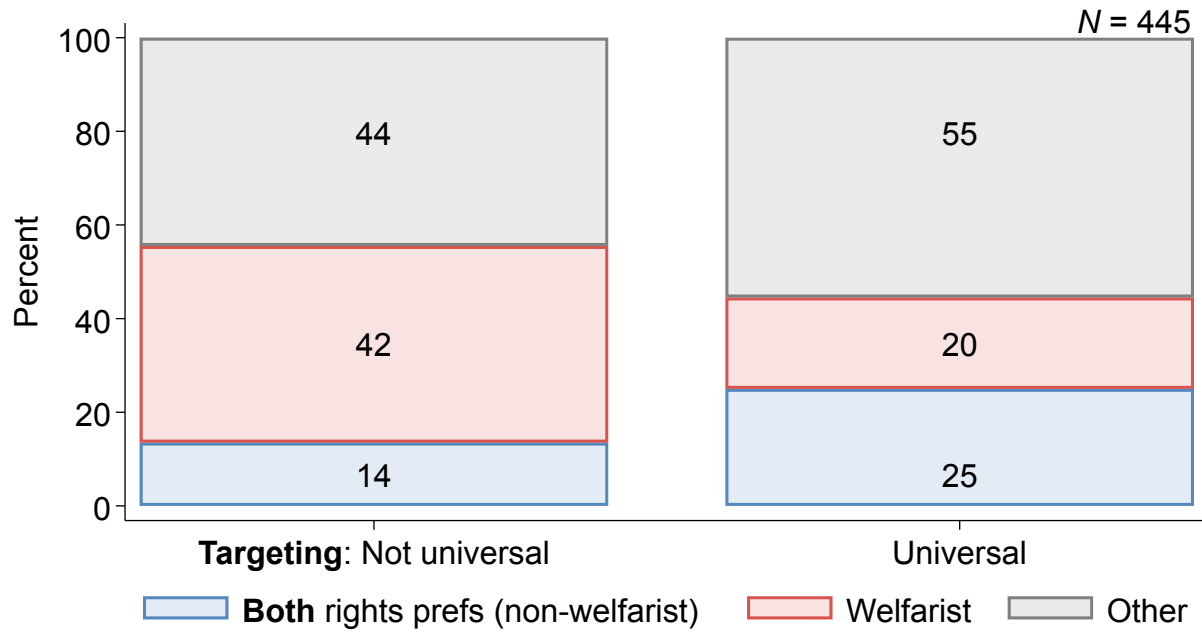
Note: This figure shows keywords from open-ended responses where we asked participants to explain the reasons for their choices. Keywords are operationalized as tokenized bi-grams. We test whether keywords appear with the same frequency between the two groups following Ferrario and Stantcheva (2022) and using implementation by Benoit et al. (2018). Each bar shows the χ^2 statistic. Significant at *10%, **5%, ***1%. Panels A and B correspond to Experiments 1 and 2, showing the top 15 keywords for two groups: those with positive WTP (blue) in the experiment and those with zero WTP (orange). Panel C shows the top keywords for those who allocate universally in Experiment 4 (blue) and for those who do not (orange).

Figure A8: Distribution of Valuations of Good



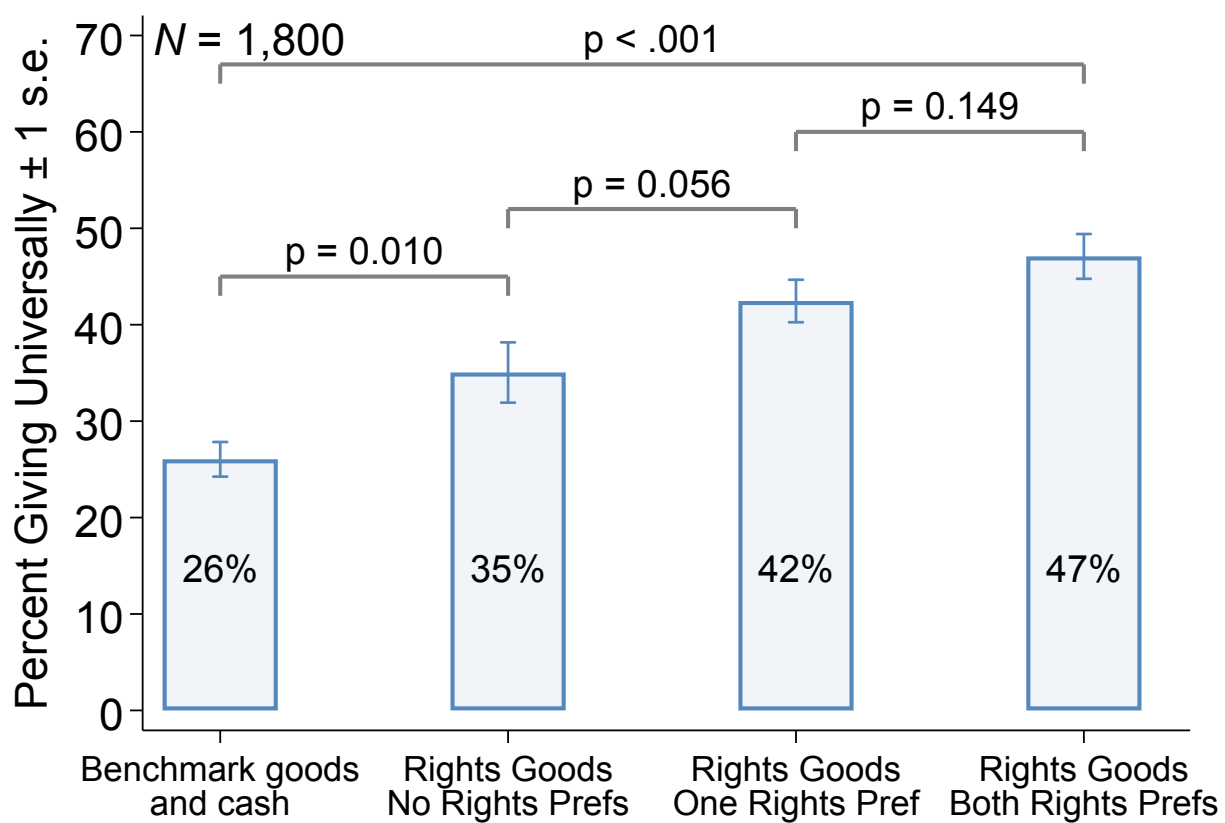
Note: This figure shows the distribution of elicited WTP for each good.

Figure A9: Non-Welfarist Preferences, Welfarism, and Universalism



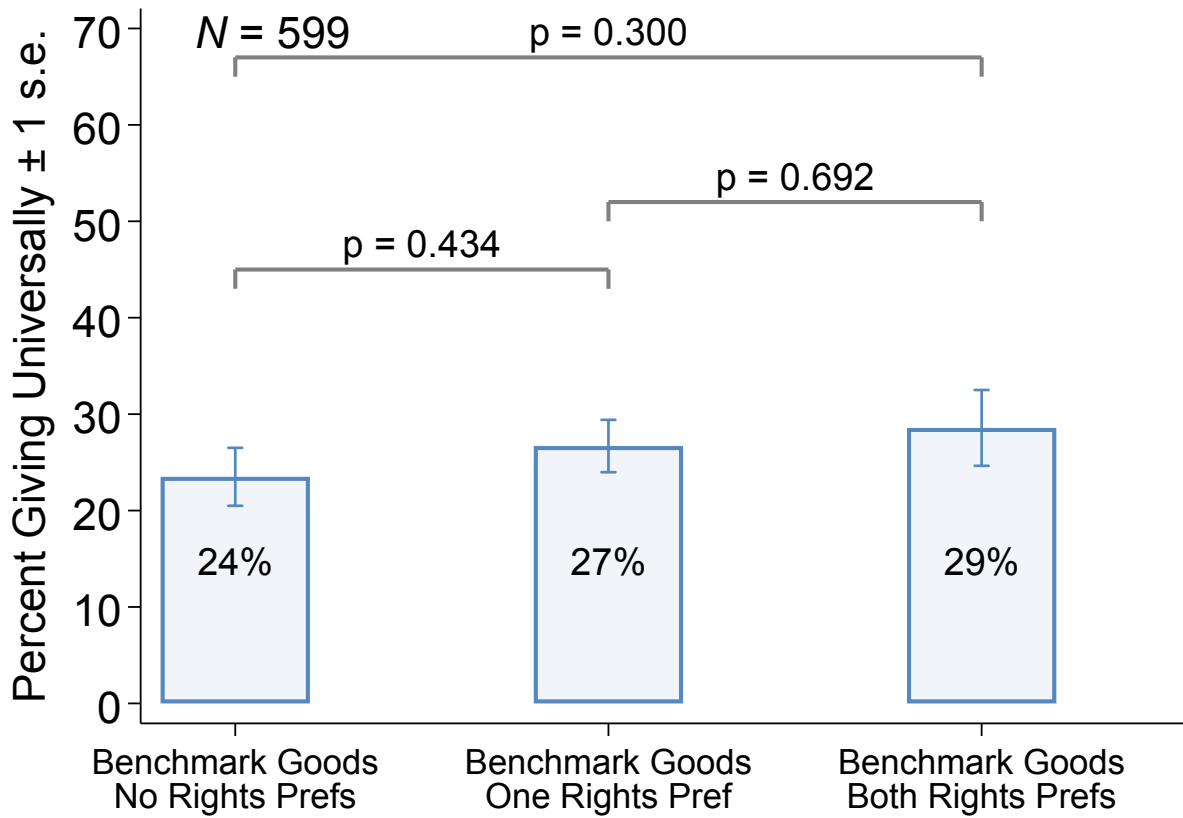
Note: This figure plots the share of *Both rights prefs (Non-Welfarist)*, *Welfarist*, and *Other* participants among those who did and did not distribute the good universally in Experiment 4. See Table A20 for detailed regression results. The figure includes participants who saw rights goods, those with posteriors ≥ 0.9 in Experiment 2, and those we can classify as welfarist or non-welfarist based on their prior beliefs.

Figure A10: Robustness: Correlations Between Universalism and Features of Rights



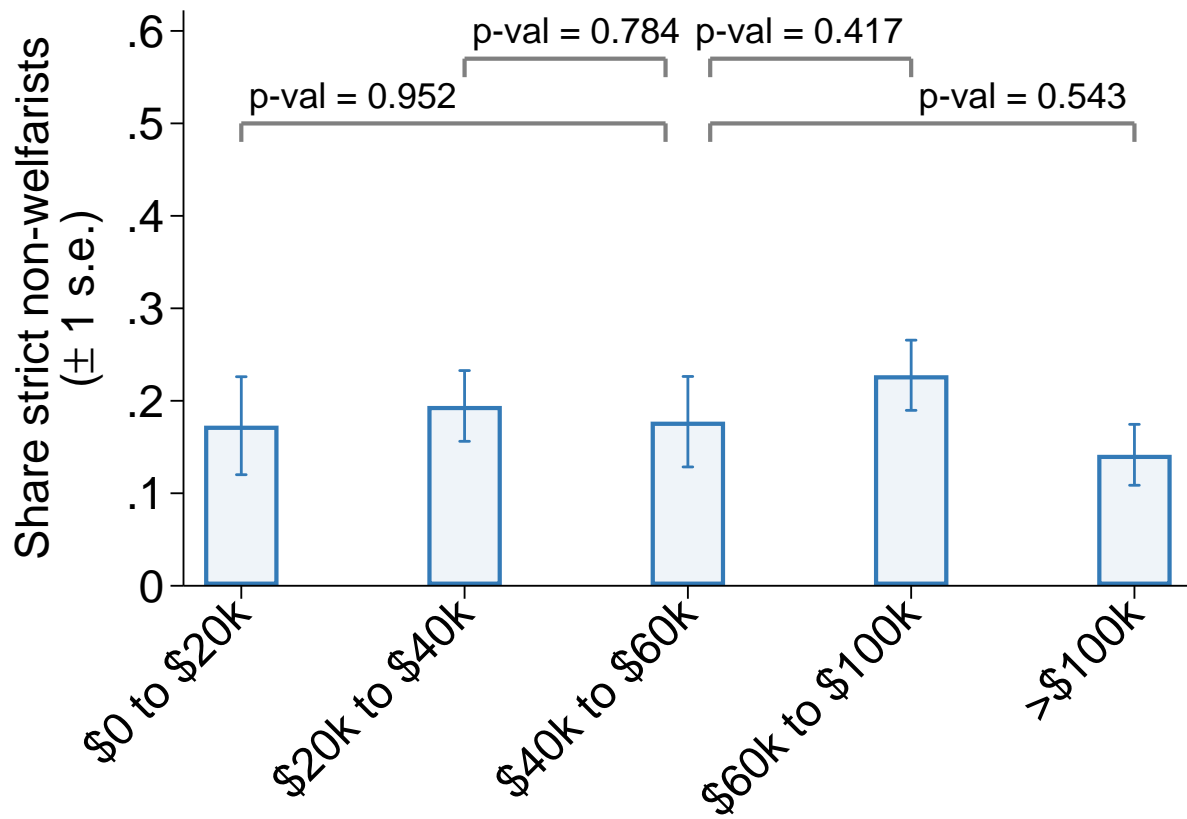
Note: This figure replicates Figure 5A, relaxing the restriction on Experiment 2 posteriors and including all participants who did experiments with rights goods.

Figure A11: Correlations Between Universalism and Features of Rights for Benchmark Goods



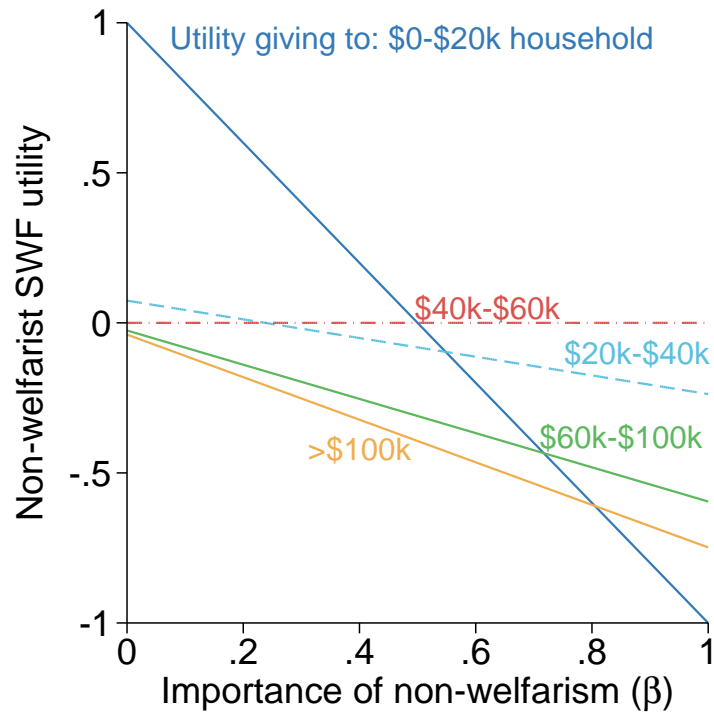
Note: This figure breaks down the leftmost bar of Figure 5. The sample is restricted to participants who saw benchmark goods, including those who saw cash in the targeting experiment. Each bar displays the share of participants with the described non-welfarist preferences who give the benchmark good (YMCA membership or bus pass) universally in Experiment 4.

Figure A12: Benchmark Goods: Incidence by Income



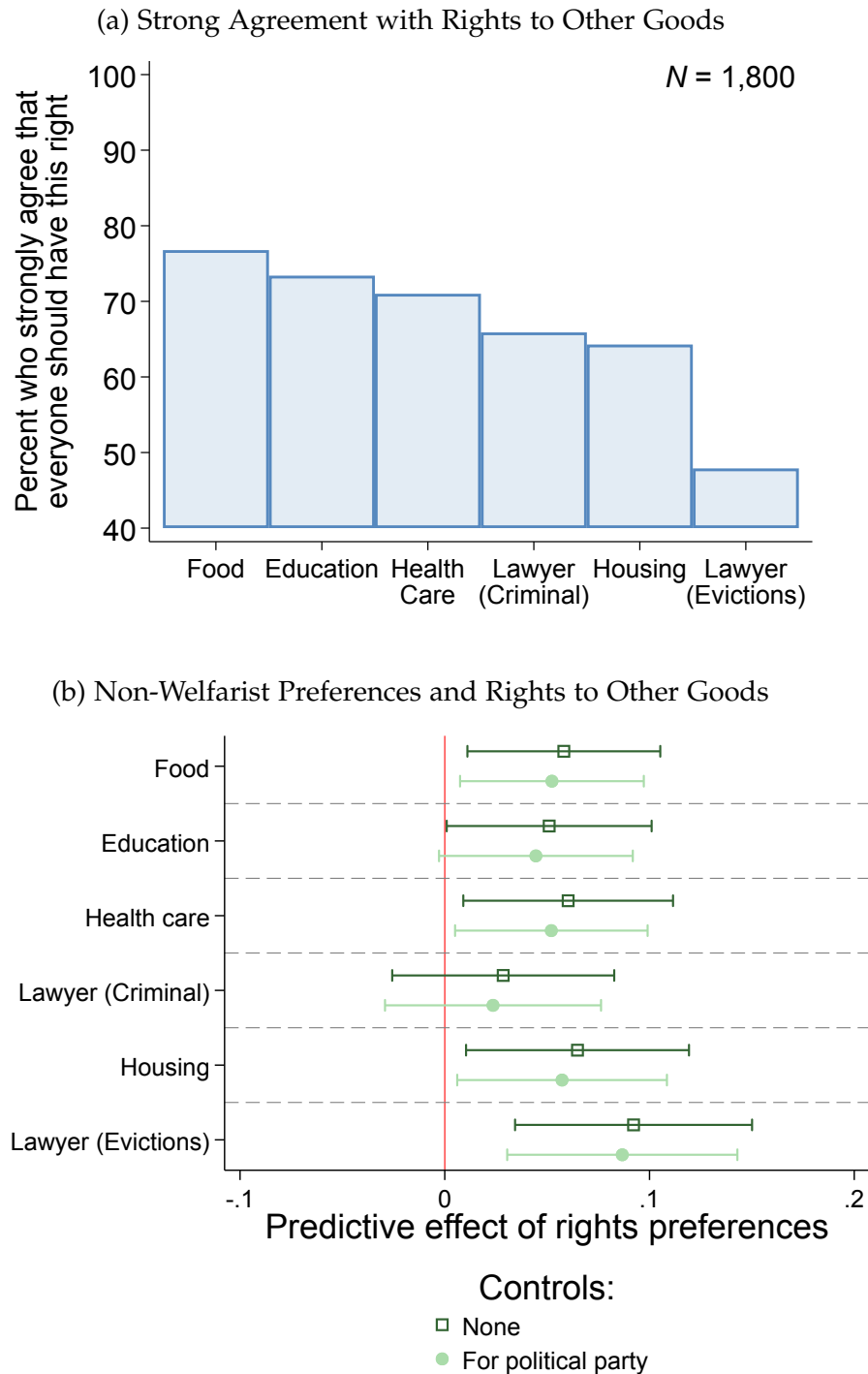
Note: This figure replicates Figure 6A using benchmark goods instead of rights goods.

Figure A13: Trade-Off between Non-Welfarism and Progressivity



Note: This figure complements Figure 6. We show results from the exercise in Equation (38). The parameter β represents the importance that the SWF places on non-welfarist preferences. Each line represents the SWF's utility from taking \$1 from the \$40k–60k household and giving to the indicated household. The results show that the SWF always prefers to give to the poorest household (low non-welfarist concerns) or keep with the middle-income household (high non-welfarist concerns).

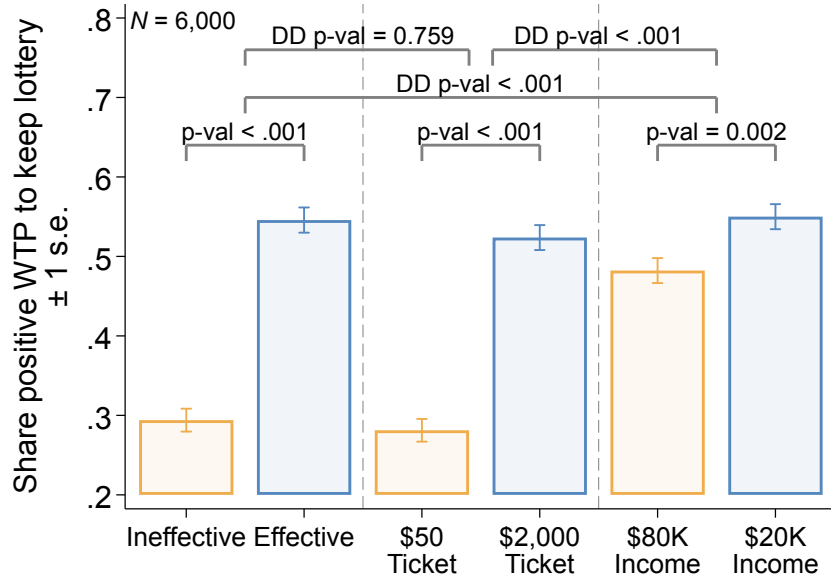
Figure A14: Rights to Other Goods



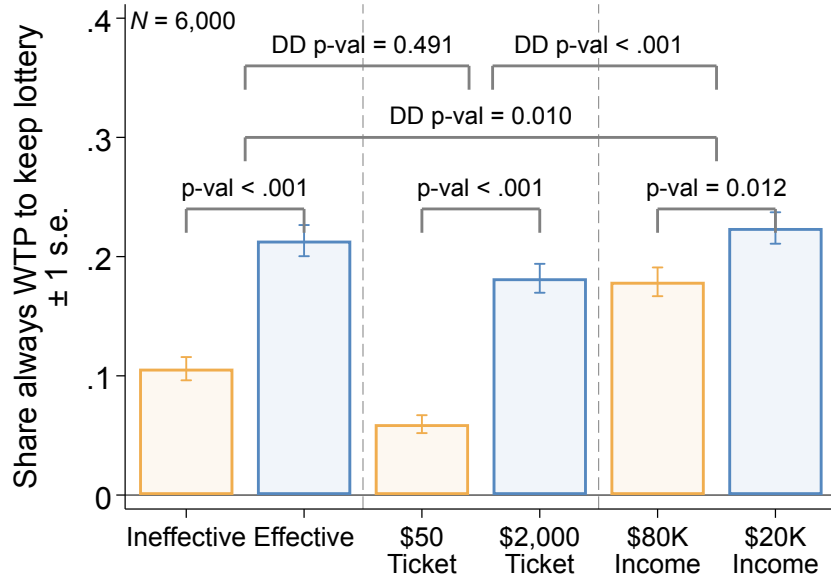
Note: This figure plots results from asking Spectators if they agree there is a right to several types of goods. Panel A plots the percent who “strongly agree” that each good should be a right. Panel B shows the predictive effect of having positive WTP in both Experiments 1 and 2 on strongly agreeing each good should be a right. Estimates shown with ± 1.96 standard errors.

Figure A15: Mechanisms: Alternate Outcomes

(a) Share with positive WTP to keep lottery



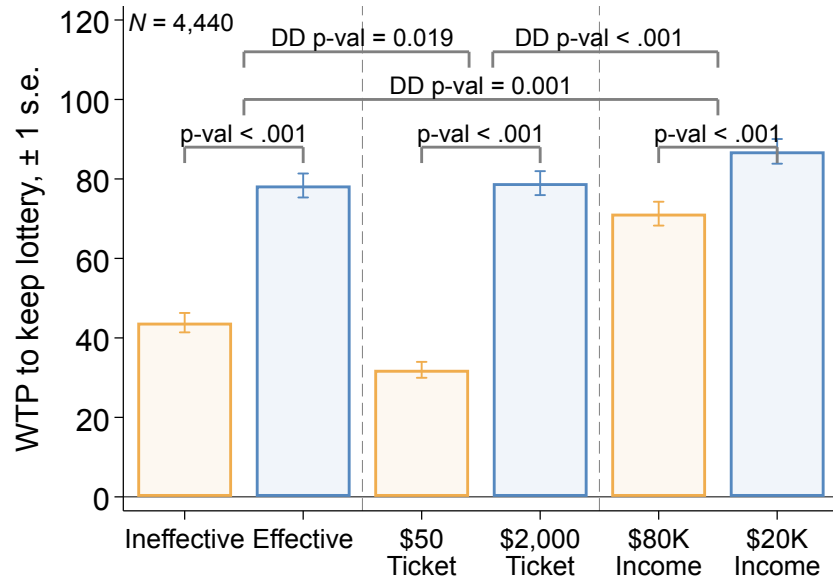
(b) Share always willing to pay to keep lottery



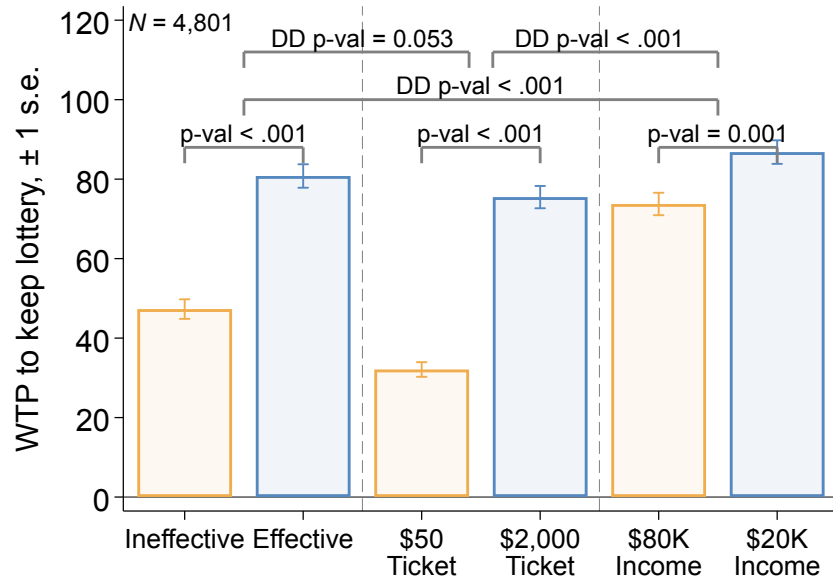
Note: These figures replicate Figure 7 for two alternate outcomes. Rather than WTP, Panel A plots the share of participants with a positive WTP to preserve the lottery in each arm. Panel B plots the share of participants who are always willing to pay to preserve the lottery for all values of \$X we show (up to \$200). See Table A27 for detailed regression results.

Figure A16: Mechanisms: Comprehension and Timing

(a) Passed all comprehension checks



(b) Response time between p10 and p90



Note: These figures replicate Figure 7, each with a sample restriction. Panel A restricts to participants who correctly answered all three comprehension checks. Panel B restricts to participants between the 10th and 90th percentiles of response times.

B Additional Tables

Table A1: Main Study: Balance

| | Rights goods | | | Benchmark goods | | | p-values | | |
|--------------------------------|----------------|----------------|----------------|-----------------|----------------|----------------|----------|-------|-------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | Pooled | Lawyers | Health care | Pooled | YMCA | Bus pass | 1=4 | 2=4 | 3=4 |
| White non-Hispanic | 0.71 [0.46] | 0.71 [0.45] | 0.70 [0.46] | 0.72 [0.45] | 0.72 [0.45] | 0.72 [0.45] | 0.72 | 0.85 | 0.71 |
| Liberal | 0.59 [0.49] | 0.57 [0.50] | 0.61 [0.49] | 0.54 [0.50] | 0.53 [0.50] | 0.55 [0.50] | 0.03 | 0.23 | 0.01 |
| Income > 60k | 0.47 [0.50] | 0.49 [0.50] | 0.45 [0.50] | 0.50 [0.50] | 0.53 [0.50] | 0.48 [0.50] | 0.15 | 0.59 | 0.05 |
| Less than Bachelor's | 0.45 [0.50] | 0.45 [0.50] | 0.45 [0.50] | 0.45 [0.50] | 0.40 [0.49] | 0.49 [0.50] | 0.96 | 0.79 | 0.76 |
| Legal case without a lawyer | 0.16 [0.36] | 0.17 [0.38] | 0.14 [0.35] | 0.16 [0.36] | 0.15 [0.36] | 0.16 [0.37] | 0.72 | 0.25 | 0.55 |
| Urgent health issue without HC | 0.31 [0.46] | 0.30 [0.46] | 0.32 [0.47] | 0.32 [0.47] | 0.30 [0.46] | 0.33 [0.47] | 0.43 | 0.20 | 0.96 |
| Female | 0.50 [0.50] | 0.51 [0.50] | 0.50 [0.50] | 0.52 [0.50] | 0.53 [0.50] | 0.51 [0.50] | 0.42 | 0.67 | 0.31 |
| Less than age 40 | 0.57 [0.50] | 0.55 [0.50] | 0.58 [0.49] | 0.56 [0.50] | 0.54 [0.50] | 0.57 [0.50] | 0.97 | 0.90 | 0.82 |
| <i>F</i> -statistic | | | | | | | 0.996 | 0.510 | 1.586 |
| <i>p</i> -value | | | | | | | 0.437 | 0.849 | 0.124 |
| Observations | 1,201 | 606 | 595 | 599 | 299 | 300 | 1,800 | 1,205 | 1,194 |

Note: This table expands on Table 1. Columns (1)-(6) show demographic characteristics for different goods, pooled and separately. Columns (7)-(9) show the *p*-values of the differences between pooled benchmarks and each rights goods column. The *F*-statistic is from a joint test of significance for the listed demographic variables. Brackets show standard deviations.

Table A2: Main Study: Attrition

| | Number of participants | | | |
|--|------------------------|--------------------|-------------|-----------------|
| | (1) Lawyers | (2) Health care | (3) YMCA | (4) Bus pass |
| Started | 627 | 623 | 309 | 316 |
| Egalitarianism★ | 615 | 605 | 306 | 310 |
| Passed at least 1 of 2 attention checks★ | 614 | 604 | 303 | 309 |
| Inalienability★ | 614 | 603 | 306 | 309 |
| Choice | 612 | 600 | 302 | 306 |
| Anti-targeting | 610 | 599 | 302 | 306 |
| Passed at least 2 of 3 attention checks | 607 | 596 | 300 | 300 |
| Good valuation | 607 | 596 | 300 | 300 |
| Demographics & political preferences | 606 | 595 | 299 | 300 |

Note: This table shows the attrition in our survey. Rows show the number of Spectators participating, for each good, at each stage of our survey flow. Rows marked with ★ were presented in a randomized order.

Table A3: Tests of Experiment 1 (Inalienability)

| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) WTP (s.d.) | (5) WTP (s.d.) |
|-----------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Rights good (= 1) | 0.078 (0.025) [0.002] | 0.080 (0.017) [0.000] | 0.289 (0.053) [0.000] | | |
| Lawyers (= 1) | | | | 0.354 (0.064) [0.000] | |
| Health care (= 1) | | | | | 0.222 (0.061) [0.000] |
| Raw mean (benchmarks) | | | 58.4 | 58.4 | 58.4 |
| Raw s.d. (benchmarks) | | | 69.9 | 69.9 | 69.9 |
| Mean (benchmarks) | 0.456 | 0.117 | 0.000 | 0.000 | 0.000 |
| Observations | 1,800 | 1,800 | 1,800 | 1,205 | 1,194 |

Note: This table shows the effects of being assigned to a rights good on three measures of WTP for inalienability (Equations 13 and 14). Columns (1)-(3) pool lawyers and health care, while column (4) shows results for lawyers and column (5) shows results for health care. Parentheses show robust standard errors. In columns (3)-(5), *Mean (benchmarks)* and *WTP* are reported in standard deviations relative to the benchmark goods in the full sample. Parentheses show robust standard errors. Brackets show *p*-values.

Table A4: Inalienability Robustness: Experiment Order

| | Inalienability first | | Inalienability second | |
|-----------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|
| | (1) Non-zero WTP (= 1) | (2) WTP (s.d.) | (3) Non-zero WTP (= 1) | (4) WTP (s.d.) |
| Rights good (= 1) | 0.074 (0.034) [0.030] | 0.294 (0.065) [0.000] | 0.080 (0.035) [0.025] | 0.278 (0.081) [0.001] |
| Raw mean (benchmarks) | | 43.2 | | 74.4 |
| Raw s.d. (benchmarks) | | 59.5 | | 76.2 |
| Mean (benchmarks) | 0.378 | -0.217 | 0.538 | 0.229 |
| Observations | 915 | 915 | 885 | 885 |

Note: This table replicates Columns 1 and 3 of Table A3, splitting by the random order in which the participant did Experiments 1 and 3 (Figure 1). Parentheses show robust standard errors. Brackets show p -values.

Table A5: Inalienability Robustness: Controlling for Valuation of Good

| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) WTP (s.d.) | (5) WTP (s.d.) |
|-----------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Rights good (= 1) | 0.065 (0.027) [0.017] | 0.064 (0.019) [0.001] | 0.248 (0.058) [0.000] | | |
| Lawyers (= 1) | | | | 0.255 (0.077) [0.001] | |
| Health care (= 1) | | | | | 0.235 (0.064) [0.000] |
| WTP for good FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Raw mean (benchmarks) | | | 58.4 | 58.4 | 58.4 |
| Raw s.d. (benchmarks) | | | 69.9 | 69.9 | 69.9 |
| Mean (benchmarks) | 0.456 | 0.117 | 0.000 | 0.000 | 0.000 |
| Observations | 1,800 | 1,800 | 1,800 | 1,205 | 1,194 |

Note: This table replicates Table A3 with added fixed effects for valuation of the good. WTP fixed effects are indicators for \$100 bins. Parentheses show robust standard errors. Brackets show p -values.

Table A6: Balance Among High Posterior Participants

| | (1) |
|--------------------------------|------------------------|
| | Rights-Benchmark goods |
| White non-Hispanic | -0.01 (0.03) |
| Liberal | 0.05 (0.03) |
| Income > 60k | -0.03 (0.03) |
| Less than Bachelor's | 0.02 (0.03) |
| Legal case without a lawyer | 0.01 (0.04) |
| Urgent health issue without HC | -0.03 (0.03) |
| Female | -0.01 (0.03) |
| Less than age 40 | 0.02 (0.03) |
| <i>F</i> -statistic | 0.750 |
| Observations | 1,198 |

Note: This table shows balance for respondents with high posteriors ($\geq 90\%$) about the percent of tenants expected to pick cash over the good. The *F*-statistic is from a joint test of significance for the listed demographic variables. Parentheses show robust standard errors.

Table A7: Tests of Experiment 2 (Willingness to Pay for Choice)

| | | | | Posterior ≥ 0.9 | | | Posterior = 1 | | |
|-----------------------|------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|-----------------------------|------------------------------|-----------------------------|-----------------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Non-zero WTP (= 1) | (5) Max WTP (= 1) | (6) WTP (s.d.) | (7) Non-zero WTP (= 1) | (8) Max WTP (= 1) | (9) WTP (s.d.) |
| Rights good (= 1) | 0.171 (0.025) [0.000] | 0.123 (0.018) [0.000] | 0.495 (0.057) [0.000] | 0.166 (0.030) [0.000] | 0.139 (0.020) [0.000] | 0.518 (0.067) [0.000] | 0.072 (0.047) [0.122] | 0.153 (0.033) [0.000] | 0.516 (0.109) [0.000] |
| Posterior | -0.442 (0.044) [0.000] | -0.244 (0.047) [0.000] | -1.046 (0.135) [0.000] | | | | | | |
| Raw mean (benchmarks) | | | 223.0 | | | 189.3 | | | 213.0 |
| Raw s.d. (benchmarks) | | | 277.4 | | | 247.5 | | | 268.3 |
| Mean (benchmarks) | 0.432 | 0.092 | 0.000 | 0.382 | 0.066 | -0.121 | 0.438 | 0.087 | -0.036 |
| Observations | 1,800 | 1,800 | 1,800 | 1,058 | 1,058 | 1,058 | 459 | 459 | 459 |

Note: This table shows the effects of being assigned to a rights good on three measures of WTP for choice (Equation 13). Columns (1)-(3) show results for the whole sample controlling for posterior beliefs about the percent of tenants who will choose cash over the good. Columns (4)-(6) restrict to those with posteriors beliefs greater than or equal to 90% and columns (7)-(9) restrict to those with posteriors of 100%. In columns (3), (6) and (9), *Mean (benchmarks)* and *WTP* are reported in standard deviations relative to the benchmark goods. Parentheses show robust standard errors. Brackets show *p*-values.

Table A8: Willingness to Pay for Choice Robustness: Lawyers and Health Care

| | | | | Posterior ≥ 0.9 | | | Posterior = 1 | | |
|---------------------------|------------------------------|-------------------------|----------------------|------------------------------|-------------------------|----------------------|------------------------------|-------------------------|----------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Non-zero WTP (= 1) | (5) Max WTP (= 1) | (6) WTP (s.d.) | (7) Non-zero WTP (= 1) | (8) Max WTP (= 1) | (9) WTP (s.d.) |
| Panel A. Lawyers | | | | | | | | | |
| Lawyers (= 1) | 0.250 | 0.168 | 0.642 | 0.263 | 0.180 | 0.677 | 0.131 | 0.237 | 0.752 |
| | (0.029) | (0.024) | (0.072) | (0.037) | (0.029) | (0.089) | (0.060) | (0.050) | (0.157) |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.029] | [0.000] | [0.000] |
| Posterior | -0.351 | -0.278 | -1.085 | | | | | | |
| | (0.055) | (0.059) | (0.170) | | | | | | |
| | [0.000] | [0.000] | [0.000] | | | | | | |
| Panel B. Health Care | | | | | | | | | |
| Health Care (= 1) | 0.099 | 0.079 | 0.349 | 0.086 | 0.106 | 0.384 | 0.032 | 0.095 | 0.354 |
| | (0.029) | (0.020) | (0.067) | (0.036) | (0.024) | (0.080) | (0.054) | (0.037) | (0.124) |
| | [0.001] | [0.000] | [0.000] | [0.016] | [0.000] | [0.000] | [0.547] | [0.011] | [0.004] |
| Posterior | -0.517 | -0.166 | -0.932 | | | | | | |
| | (0.065) | (0.059) | (0.179) | | | | | | |
| | [0.000] | [0.005] | [0.000] | | | | | | |
| Raw mean (benchmarks) | | | 223.0 | | | 189.3 | | | 213.0 |
| Raw s.d. (benchmarks) | | | 277.4 | | | 247.5 | | | 268.3 |
| Mean (benchmarks) | 0.432 | 0.092 | 0.000 | 0.382 | 0.066 | -0.121 | 0.438 | 0.087 | -0.036 |
| Observations: lawyers | 606 | 606 | 606 | 273 | 273 | 273 | 102 | 102 | 102 |
| Observations: health care | 595 | 595 | 595 | 327 | 327 | 327 | 149 | 149 | 149 |
| Observations: benchmarks | 599 | 599 | 599 | 458 | 458 | 458 | 208 | 208 | 208 |

Note: This table replicates Table A7 separately for lawyers (Panel A) and health care (Panel B). Parentheses show robust standard errors. Brackets show p -values.

Table A9: Willingness to Pay for Choice Robustness: High Posteriors

| | Posterior = 1 | | | | Posterior ≥ 0.9 | | | |
|-----------------------|------------------------------|-----------------------------|-----------------------------|------------------------------------|------------------------------|-----------------------------|-----------------------------|------------------------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Entitled to Choice (= 1) | (5) Non-zero WTP (= 1) | (6) Max WTP (= 1) | (7) WTP (s.d.) | (8) Entitled to Choice (= 1) |
| Rights good (= 1) | 0.080 (0.048) [0.095] | 0.152 (0.034) [0.000] | 0.515 (0.111) [0.000] | 0.115 (0.045) [0.011] | 0.151 (0.031) [0.000] | 0.134 (0.020) [0.000] | 0.492 (0.067) [0.000] | 0.161 (0.029) [0.000] |
| Raw mean (benchmarks) | | | 213.0 | | | | 189.3 | |
| Raw s.d. (benchmarks) | | | 268.3 | | | | 247.5 | |
| Mean (benchmarks) | 0.438 | 0.087 | -0.036 | 0.308 | 0.382 | 0.066 | -0.121 | 0.264 |
| DDML | ✓ | ✓ | ✓ | | | | | |
| Posterior FE | | | | | ✓ | ✓ | ✓ | |
| Observations | 459 | 459 | 459 | 459 | 1,058 | 1,058 | 1,058 | 1,058 |

Note: This table shows robustness checks for Willingness to Pay for choice using the sub-sample of respondents who believe 100 out of 100 tenants would choose cash over the good. Columns (1)-(3) replicate columns (7)-(9) of Table A7 with double/de-biased machine learning (Chernozhukov et al., 2018). The model selects from the demographic controls reported in Table 1: race, income, education, gender, age, political beliefs, having experienced a legal case without a lawyer, and having had an urgent health issue without access to health care. The outcome in columns (4) and (8) is a dummy that indicates a respondent selected “All tenants should be entitled to the choice of a [good]” as one reason for their decisions about giving the tenant a choice. Columns (5)-(7) replicate columns (4)-(6) of Table A7, adding fixed effects for posterior beliefs within the sample. Parentheses show robust standard errors. Brackets show p -values.

Table A10: Willingness to Pay for Choice Robustness: Controlling for Valuation of Good

| | | | | Posterior ≥ 0.9 | | | Posterior = 1 | | |
|-----------------------|------------------------------|------------------------------|------------------------------|------------------------------|-----------------------------|-----------------------------|------------------------------|-----------------------------|-----------------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Non-zero WTP (= 1) | (5) Max WTP (= 1) | (6) WTP (s.d.) | (7) Non-zero WTP (= 1) | (8) Max WTP (= 1) | (9) WTP (s.d.) |
| Rights good (= 1) | 0.082 (0.027) [0.002] | 0.066 (0.019) [0.000] | 0.256 (0.061) [0.000] | 0.070 (0.034) [0.038] | 0.082 (0.021) [0.000] | 0.279 (0.070) [0.000] | -0.021 (0.049) [0.665] | 0.098 (0.033) [0.003] | 0.281 (0.110) [0.011] |
| Posterior | -0.355 (0.044) [0.000] | -0.179 (0.045) [0.000] | -0.790 (0.130) [0.000] | | | | | | |
| WTP for good FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Raw mean (benchmarks) | | | 223.0 | | | 223.0 | | | 223.0 |
| Raw s.d. (benchmarks) | | | 277.4 | | | 277.4 | | | 277.4 |
| Mean (benchmarks) | 0.432 | 0.092 | 0.000 | 0.432 | 0.092 | 0.000 | 0.432 | 0.092 | 0.000 |
| Observations | 1,800 | 1,800 | 1,800 | 1,058 | 1,058 | 1,058 | 459 | 459 | 459 |

Note: This table replicates Table A7, but adds to all columns a fixed effect for participant willingness to pay for the good (see Figure A8). WTP fixed effects are indicators for \$100 bins. Parentheses show robust standard errors. Brackets show p -values.

Table A11: Willingness to Pay for Choice Robustness: Reasoning

| | All | | Prioritize Saving | | Entitled to Choice | | No One Would Choose | | Would Really Want | |
|-------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|
| | (1) Non-zero WTP (= 1) | (2) WTP (s.d.) | (3) Non-zero WTP (= 1) | (4) WTP (s.d.) | (5) Non-zero WTP (= 1) | (6) WTP (s.d.) | (7) Non-zero WTP (= 1) | (8) WTP (s.d.) | (9) Non-zero WTP (= 1) | (10) WTP (s.d.) |
| Rights good (= 1) | 0.166 (0.030) [0.000] | 0.518 (0.067) [0.000] | 0.068 (0.048) [0.151] | 0.098 (0.053) [0.067] | 0.063 (0.038) [0.098] | 0.715 (0.130) [0.000] | 0.052 (0.037) [0.164] | 0.099 (0.056) [0.077] | 0.088 (0.057) [0.124] | 0.564 (0.147) [0.000] |
| Mean (benchmarks) | 0.382 | -0.121 | 0.246 | -0.425 | 0.843 | 0.631 | 0.172 | -0.456 | 0.724 | 0.342 |
| Observations | 1,058 | 1,058 | 358 | 358 | 376 | 376 | 453 | 453 | 252 | 252 |

Note: This table shows robustness checks for the choice experiment, conditioning on answers to a question following Experiment 2: “Which of the following reasons motivated your choice(s)? Select all that apply.” Columns (1)-(2) show results for all participants. Columns (3)-(4) show result for participants who selected the reason “Saving is my priority.” Columns (5)-(6) show result for participants who selected the reason “All tenants should be entitled to the choice of a [good].” Columns (7)-(8) show result for participants who selected the reason “I did not think anyone would choose the [good] in reality.” Columns (9)-(10) show result for participants who selected the reason “I thought anyone who would choose the [good] would really want it.” 4.5% of participants answered “None of the above.” Parentheses show robust standard errors. Brackets show p -values.

Table A12: Willingness to Pay for Choice Robustness: WTP for Choice by Cash Value

| | | Posterior ≥ 0.9 | Posterior = 1 |
|--|----------|----------------------|---------------|
| | (1) | (2) | (3) |
| | WTP | WTP | WTP |
| Rights good=1 | 225.102 | 130.789 | 289.695 |
| | (78.690) | (94.574) | (154.849) |
| | [0.004] | [0.167] | [0.062] |
| Cash value (\$) | -0.375 | -0.412 | -0.041 |
| | (0.226) | (0.231) | (0.373) |
| | [0.097] | [0.075] | [0.913] |
| Rights good=1 \times Cash value (\$) | -0.210 | 0.048 | -0.580 |
| | (0.307) | (0.369) | (0.603) |
| | [0.494] | [0.895] | [0.336] |
| Mean (benchmarks) | 223.0 | 223.0 | 223.0 |
| Observations | 1,800 | 1,058 | 459 |

Note: This table shows robustness checks for the choice experiment (Experiment 2), controlling for the interaction of rights goods and the value of cash in the choice. The outcome is participants' WTP in dollars to give tenants a choice between the good and cash. Cash values are either \$200 or \$300. Column (1) includes all participants, while columns (2) and (3) condition on posteriors about the share of tenants who would pick cash. Parentheses show robust standard errors. Brackets show p -values.

Table A13: Willingness to Pay for Choice Robustness: Beliefs by Cash Value

| | Rights Goods | | Benchmark Goods | |
|----------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|
| | (1) P(choose cash) | (2) P(choose cash) | (3) P(choose cash) | (4) P(choose cash) |
| Cash value (\$) | 0.021 (0.015) [0.162] | 0.005 (0.003) [0.121] | 0.006 (0.013) [0.661] | -0.003 (0.003) [0.410] |
| Mean (Cash= 200) | 77.114 | 95.931 | 89.947 | 97.199 |
| Posterior ≥ 0.9 | | ✓ | | ✓ |
| Observations | 1,201 | 600 | 599 | 458 |

Note: This table shows the effect of the randomized cash value in Experiment 2 (either \$200 or \$300) on posterior beliefs about the probability that tenants will choose cash over the good. Parentheses show robust standard errors. Posterior beliefs are between 0 to 100. Brackets show p -values.

Table A14: Tests of Egalitarianism

| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Non-zero WTP (= 1) | (5) Max WTP (= 1) | (6) WTP (s.d.) |
|---|------------------------------|-----------------------------|------------------------------|------------------------------|-----------------------------|------------------------------|
| Rights good=1 | 0.082 (0.020) [0.000] | 0.193 (0.033) [0.000] | 0.562 (0.066) [0.000] | 0.074 (0.019) [0.000] | 0.144 (0.034) [0.000] | 0.483 (0.068) [0.000] |
| $\mathbb{1}(z = 9)$ | 0.043 (0.023) [0.068] | 0.124 (0.038) [0.001] | 0.291 (0.079) [0.000] | 0.055 (0.023) [0.017] | 0.122 (0.038) [0.001] | 0.307 (0.079) [0.000] |
| Rights good=1 \times $\mathbb{1}(z = 9)$ | -0.046 (0.026) [0.075] | 0.011 (0.047) [0.809] | -0.154 (0.090) [0.087] | -0.055 (0.025) [0.027] | 0.011 (0.047) [0.815] | -0.170 (0.090) [0.058] |
| WTP for good FE | | | | ✓ | ✓ | ✓ |
| Raw mean (benchmarks, $\mathbb{1}(z < 9)$) | | | 575.3 | | | 575.3 |
| Raw s.d. (benchmarks, $\mathbb{1}(z < 9)$) | | | 328.4 | | | 328.4 |
| Mean (benchmarks, $\mathbb{1}(z < 9)$) | 0.888 | 0.258 | 0.000 | 0.888 | 0.258 | 0.000 |
| Observations | 1,800 | 1,800 | 1,800 | 1,800 | 1,800 | 1,800 |

Note: This table reports estimates from the differences-in-differences specification from Equation 15. $\mathbb{1}(z < 9)$ indicates the Spectator saw 9 out of 10 tenants had already received the good, rather than 1 or 5 tenants. Outcomes are the extensive margin (*Non-zero WTP*), an indicator for having the maximum WTP that we elicit (*Max WTP*), and the intensive margin (*WTP*). WTP is reported in terms of standard deviations relative to benchmark goods. Columns (4)-(6) add fixed effects for WTP for the good directly. WTP fixed effects are indicators for \$100 bins. Parentheses show robust standard errors. Brackets show *p*-values.

Table A15: Robustness: Multiple Hypothesis Testing Corrections

| | Inalienability | | Choice | | Egalitarianism | |
|-----------------------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|------------------------------|------------------------------|
| | (1) Non-zero WTP (= 1) | (2) WTP (s.d.) | (3) Non-zero WTP (= 1) | (4) WTP (s.d.) | (5) Non-zero WTP (= 1) | (6) WTP (s.d.) |
| Rights good | 0.078 (0.025) [0.002] | 0.289 (0.053) [0.000] | 0.166 (0.030) [0.000] | 0.518 (0.067) [0.000] | 0.082 (0.020) [0.000] | 0.562 (0.066) [0.000] |
| Rights good & $\mathbb{1}(z = 9)$ | | | | | -0.046 (0.026) [0.075] | -0.154 (0.090) [0.087] |
| Romano-Wolf p -value | 0.008 | < 0.001 | < 0.001 | < 0.001 | 0.082 | 0.096 |
| Posterior ≥ 0.9 | | | ✓ | ✓ | | |
| Observations | 1,800 | 1,800 | 1,058 | 1,058 | 1,800 | 1,800 |

Note: This table shows robustness to multiple hypothesis testing corrections for Experiments 1–3 using Romano-Wolf p -values (Clarke et al., 2020) with 1,000 iterations. We perform the correction separately for extensive and intensive margin outcomes: columns (1), (3), and (5) are tested together and columns (2), (4), and (6) are tested together. Parentheses show robust standard errors. Brackets show p -values.

Table A16: Tests of Experiment 4 (Anti-Targeting)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|-------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| | Universal (= 1) | No. Tenants | Universal (= 1) | No. Tenants | Universal (= 1) | No. Tenants |
| Rights good (= 1) | 0.168 (0.023) [0.000] | 1.419 (0.176) [0.000] | | | | |
| Lawyers (= 1) | | | 0.263 (0.027) [0.000] | 2.140 (0.201) [0.000] | | |
| Health care (= 1) | | | | | 0.071 (0.026) [0.007] | 0.685 (0.204) [0.001] |
| Mean (benchmarks) | 0.260 | 4.855 | 0.260 | 4.855 | 0.260 | 4.855 |
| Observations | 1,800 | 1,800 | 1,205 | 1,205 | 1,194 | 1,194 |

Note: This table shows the effects of being assigned to a rights good on choices in the anti-targeting experiment (Equations 13 and 14). The outcome columns (1), (3) and (5) is an indicator for whether the participant provided the good universally. The outcome columns (2), (4) and (6) is the minimum number of tenants the participant distributed to, which ranges between 1 (preferring to giving the poorest tenant the good and cash) and 10 (always preferring universal provision). Columns (1)-(2) pool lawyers and health care, while columns (3)-(4) shows results for lawyers and column (5)-(6) shows results for health care. Parentheses show robust standard errors. Brackets show p -values.

Table A17: Anti-Targeting Robustness: Controlling for Valuation of Good

| | (1) | (2) | (3) | (4) | (5) | (6) |
|-------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| | Universal (= 1) | No. Tenants | Universal (= 1) | No. Tenants | Universal (= 1) | No. Tenants |
| Rights good (= 1) | 0.044 (0.024) [0.067] | 0.458 (0.187) [0.015] | | | | |
| Lawyers (= 1) | | | 0.116 (0.034) [0.001] | 1.007 (0.252) [0.000] | | |
| Health care (= 1) | | | | | 0.013 (0.026) [0.621] | 0.215 (0.204) [0.293] |
| WTP for good FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Mean (benchmarks) | 0.260 | 4.855 | 0.260 | 4.855 | 0.260 | 4.855 |
| Observations | 1,800 | 1,800 | 1,205 | 1,205 | 1,194 | 1,194 |

Note: This table replicates Table A16 with added fixed effects for valuation of the good. WTP fixed effects are indicators for \$100 bins. Parentheses show robust standard errors. Brackets show p -values.

Table A18: Correlations with Anti-Targeting Robustness: Controlling for Valuation of Good

| | Dep. Var.: Universal (= 1) | | | |
|--------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| | (1) | (2) | (3) | (4) |
| Panel A. Rights goods (pooled) | | | | |
| WTP for inalienability (s.d.) | 0.022 (0.012) [0.058] | | | 0.019 (0.017) [0.263] |
| WTP for choice (s.d.) | | 0.052 (0.016) [0.001] | | 0.048 (0.016) [0.004] |
| Mean WTP (s.d.) | | | 0.068 (0.020) [0.001] | |
| Panel B. Lawyers | | | | |
| WTP for inalienability (s.d.) | 0.022 (0.016) [0.184] | | | 0.000 (0.026) [0.992] |
| WTP for choice (s.d.) | | 0.085 (0.024) [0.001] | | 0.085 (0.027) [0.001] |
| Mean WTP (s.d.) | | | 0.086 (0.029) [0.003] | |
| Panel C. Health care | | | | |
| WTP for inalienability (s.d.) | 0.017 (0.017) [0.314] | | | 0.032 (0.023) [0.163] |
| WTP for choice (s.d.) | | 0.036 (0.021) [0.093] | | 0.031 (0.021) [0.148] |
| Mean WTP (s.d.) | | | 0.062 (0.028) [0.026] | |
| Posterior ≥ 0.9 | | ✓ | ✓ | ✓ |
| WTP for good FE | ✓ | ✓ | ✓ | ✓ |
| Mean | 0.428 | 0.437 | 0.437 | 0.437 |
| Observations | 1,201 | 600 | 600 | 600 |

Note: This table replicates Table 3 with added fixed effects for valuation of the good. WTP fixed effects are indicators for \$100 bins. Parentheses show robust standard errors. Brackets show p -values.

Table A19: Information Treatment and Anti-Targeting Updating

| | Dep. var: Update in min. no tenants | | | |
|------------------|-------------------------------------|-----------------------------|-----------------------------|------------------------------|
| | (1) | (2) | (3) | (4) |
| | All | Info. Below Prior | Info. Above Prior | Lawyers Only |
| High info (= 1) | 1.558 (0.158) [0.000] | 0.730 (0.135) [0.000] | 1.555 (0.332) [0.000] | |
| Posterior Belief | | | | 0.088 (0.010) [0.000] |
| Prior Belief | | | | -0.026 (0.006) [0.000] |
| Mean (low info) | -0.58 | -0.68 | 0.22 | -0.37 |
| IV | | | | ✓ |
| Observations | 892 | 621 | 271 | 417 |

Note: This table shows the effect of providing information about the effectiveness of lawyers or health care (either high or low) on participants' decisions to update their responses in Experiment 4. We use the same sample as in Figure A5: participants who saw rights goods and whom we can classify as welfarist or non-welfarist based on their prior beliefs. This means we exclude those that allocate universally in the original experiment and see information implying lawyers are more effective than their prior belief, and those that target only the poorest tenant and see information implying lawyers are less effective than their prior. The outcomes are the magnitude of updates in the minimum number of tenants given assistance, where positive numbers indicate allocating lawyers or health care to more tenants. Column (1) includes the full sample and shows the effect of being shown the high information treatment (vs. the low information treatment). Columns (2)-(3) show the same analysis, splitting the sample by prior beliefs relative to the treatment. Column (4) uses an instrumental variables specification, restricting to only participants who saw lawyers (because the treatment for health care was not numeric but directional). Beliefs are in percentage points. We use the treatment assignment to instrument for posterior beliefs. However, because we did not collect posterior beliefs, we substitute treatment (either 80% or 20%) for posterior beliefs. Prior beliefs are winsorized at the 5th and 95th percentiles.

Table A20: Tests of Universalism and Welfarism

| | (1) | (2) | (3) | (4) |
|-------------------------------|-----------------------------|------------------------------|--------------------------------------|------------------------------|
| | Any rights prefs | Both rights prefs | Both rights prefs (non-welfarist) | Welfarist |
| Panel A. All rights goods | | | | |
| Universal (=1) | 0.068 (0.028) [0.014] | 0.026 (0.034) [0.447] | 0.088 (0.032) [0.006] | -0.190 (0.030) [0.000] |
| Constant | 0.765 (0.018) [0.000] | 0.354 (0.020) [0.000] | 0.233 (0.017) [0.000] | 0.376 (0.020) [0.000] |
| Observations | 892 | 892 | 892 | 892 |
| Panel B. Posteriors ≥ 90 | | | | |
| Universal (=1) | 0.079 (0.043) [0.069] | 0.055 (0.045) [0.220] | 0.115 (0.040) [0.004] | -0.225 (0.043) [0.000] |
| Constant | 0.683 (0.028) [0.000] | 0.256 (0.026) [0.000] | 0.135 (0.020) [0.000] | 0.420 (0.030) [0.000] |
| Observations | 445 | 445 | 445 | 445 |
| Panel C. Lawyers | | | | |
| Universal (=1) | 0.049 (0.037) [0.187] | 0.047 (0.051) [0.352] | 0.133 (0.048) [0.005] | -0.255 (0.046) [0.000] |
| Constant | 0.812 (0.024) [0.000] | 0.410 (0.030) [0.000] | 0.244 (0.026) [0.000] | 0.474 (0.031) [0.000] |
| Observations | 417 | 417 | 417 | 417 |
| Panel D. Health care | | | | |
| Universal (=1) | 0.079 (0.041) [0.050] | -0.003 (0.045) [0.943] | 0.042 (0.043) [0.325] | -0.140 (0.039) [0.000] |
| Constant | 0.726 (0.025) [0.000] | 0.308 (0.026) [0.000] | 0.224 (0.023) [0.000] | 0.296 (0.026) [0.000] |
| Observations | 475 | 475 | 475 | 475 |

Note: This table offers formal tests for Figure A9. Each panel shows estimates of an indicator for universal provision in Experiment 4 on four outcomes. The outcome in column (1) is an indicator for having positive WTP in either Experiment 1 or 2, in column (2) is an indicator for positive WTP in *both* experiments, and in Column (3) is an indicator for positive WTP in *both* experiments and non-welfarism. Finally, the outcome in column (4) is an indicator for welfarist preferences. The sample is Spectators who saw rights goods and who can be classified as welfarist or non-welfarist based on their prior beliefs. Panel A pools both rights goods, Panel B restricts the sample in Panel A to those with posteriors $\geq 90\%$, Panel C shows results for lawyers, and Panel D shows results for health care. Parentheses show robust standard errors. Brackets show p -values.

Table A21: Incidence by Income

| | Experiment 1 | | | Experiment 2 | | | Both | (8) Joint test: (1)&(4) |
|----------------------|------------------------------|-------------------------|----------------------|------------------------------|-------------------------|----------------------|------------------------------|-------------------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (s.d.) | (4) Non-zero WTP (= 1) | (5) Max WTP (= 1) | (6) WTP (s.d.) | (7) Non-zero WTP (= 1) | |
| Relative to \$40-60k | | | | | | | | |
| \$0-20k | -0.044 | -0.040 | -0.119 | -0.187 | -0.087 | -0.477 | -0.151 | |
| | (0.051) | (0.040) | (0.118) | (0.070) | (0.057) | (0.180) | (0.064) | |
| | [0.380] | [0.321] | [0.312] | [0.008] | [0.127] | [0.008] | [0.018] | [0.022] |
| \$20-40k | -0.048 | -0.009 | -0.136 | 0.021 | -0.033 | -0.099 | -0.036 | |
| | (0.045) | (0.037) | (0.105) | (0.066) | (0.057) | (0.176) | (0.065) | |
| | [0.286] | [0.809] | [0.196] | [0.750] | [0.562] | [0.574] | [0.580] | [0.511] |
| \$60-100k | -0.088 | -0.029 | -0.186 | -0.023 | -0.079 | -0.272 | -0.090 | |
| | (0.042) | (0.034) | (0.098) | (0.060) | (0.051) | (0.157) | (0.058) | |
| | [0.037] | [0.388] | [0.058] | [0.703] | [0.118] | [0.083] | [0.121] | [0.110] |
| \$100k+ | -0.142 | -0.014 | -0.185 | -0.102 | -0.067 | -0.340 | -0.113 | |
| | (0.043) | (0.035) | (0.101) | (0.060) | (0.050) | (0.156) | (0.056) | |
| | [0.001] | [0.684] | [0.069] | [0.089] | [0.183] | [0.030] | [0.045] | [0.002] |
| Mean (\$40-60k) | 0.602 | 0.214 | 0.416 | 0.602 | 0.259 | 0.710 | 0.383 | 0.383 |
| N | 1,201 | 1,201 | 1,201 | 600 | 600 | 600 | 600 | 600 |

Note: This table offers formal tests for Figure 6A, showing the incidence of different measures of non-welfarism by income group. All columns only include those who did experiments with rights goods. Columns (4)-(7) further restrict to those with posteriors ≥ 0.9 in Experiment 2. Column (8) shows the p-values jointly testing column (1) and column (4). Parentheses show robust standard errors. Brackets show p -values.

Table A22: Robustness: Benchmark Incentives and Features of Rights

| | Inalienability | | | | Choice | | | | Anti-targeting | | | |
|---------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|
| | (1) WTP (s.d.) | (2) WTP (s.d.) | (3) WTP (s.d.) | (4) WTP (s.d.) | (5) WTP (s.d.) | (6) WTP (s.d.) | (7) WTP (s.d.) | (8) WTP (s.d.) | (9) No. Tenants | (10) No. Tenants | (11) No. Tenants | (12) No. Tenants |
| Unincentivized (= 1) | -0.002 (0.096) [0.980] | | | | -0.052 (0.095) [0.586] | | | | -0.422 (0.326) [0.195] | | | |
| Rights good (= 1) | | 0.289 (0.053) [0.000] | 0.288 (0.058) [0.000] | 0.290 (0.090) [0.001] | | 0.518 (0.067) [0.000] | 0.506 (0.071) [0.000] | 0.557 (0.096) [0.000] | | 1.419 (0.176) [0.000] | 1.312 (0.195) [0.000] | 1.734 (0.299) [0.000] |
| Benchmarks Only | ✓ | | | | ✓ | | | | ✓ | | | |
| Incentivized Benchmarks | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | |
| Unincentivized Benchmarks | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | ✓ |
| Observations | 599 | 1,800 | 1,648 | 1,353 | 458 | 1,058 | 952 | 706 | 599 | 1,800 | 1,648 | 1,353 |

Note: This table shows the results of incentivization of benchmark goods on Spectators' decisions in Experiments 1, 2 and 4. Column (2) replicates column (3) of Table A3, column (6) replicates column (6) of Table A7, and column (10) replicates column (2) of Table A16, including all benchmark good observations. Columns (3), (7), and (11) restrict benchmark observations to only those who were incentivized, while column (4), (8), and (12) restrict benchmark observations to only those who were unincentivized. Columns (5)-(8) restrict to those with posteriors ≥ 0.9 . Among those who saw incentivized goods (lawyers, YMCA memberships, bus passes), half randomly had their WTP for the good incentivized while the other half did not. Parentheses show robust standard errors. Brackets show p -values.

Table A23: Robustness: Incentives and Valuation of Good

| | Dep. Var.: WTP for good | |
|----------------|--------------------------------|--------------------------------|
| | (1) | (2) |
| | Lawyers | Benchmarks |
| Unincentivized | -11.448 (26.478) [0.666] | -7.180 (28.301) [0.800] |
| Intercept | 771.382 (18.891) [0.000] | 341.391 (20.087) [0.000] |
| Observations | 606 | 303 |

Note: This table shows the results of differential incentivization on Spectators' WTP for the good. Column (1) presents the effect of incentivization for lawyers and column (2) presents the same for benchmark goods. Among those who saw incentivized goods (lawyers, YMCA memberships, bus passes), half randomly had their WTP for the good incentivized while the other half did not. Parentheses show robust standard errors. Brackets show p -values.

Table A24: Robustness: Re-weighting Participants to U.S. Population

| | Inalienability | | Choice | | Anti-targeting | |
|----------------------|------------------------------|-----------------------------|------------------------------|-----------------------------|-----------------------------|-----------------------------|
| | (1) Non-zero WTP (= 1) | (2) WTP (s.d.) | (3) Non-zero WTP (= 1) | (4) WTP (s.d.) | (5) Universal (= 1) | (6) No. Tenants |
| Rights good (= 1) | 0.066 (0.035) [0.056] | 0.235 (0.072) [0.001] | 0.133 (0.043) [0.002] | 0.428 (0.090) [0.000] | 0.151 (0.033) [0.000] | 1.342 (0.247) [0.000] |
| Posterior ≥ 0.9 | | | ✓ | ✓ | | |
| Observations | 1,800 | 1,800 | 1,058 | 1,058 | 1,800 | 1,800 |

Note: This table shows our main results, re-weighting to match all U.S. adults (18+, a requirement on Prolific) in the 2021 ACS (Ruggles et al., 2023). We use the 5 binary characteristics present in both our survey and ACS from Column (1) of Table 1: indicators for being White non-Hispanic, income greater than \$60,000, having less than a bachelor's degree, being female, and being less than age 40. We reweight by forming 25 cells of these characteristics. Parentheses show robust standard errors. Brackets show p -values.

Table A25: Robustness: Attention Checks

| | Inalienability | | Choice | | Anti-targeting | |
|----------------------------|------------------------------|-----------------------------|------------------------------|------------------------------|-----------------------------|------------------------------|
| | (1) Non-zero WTP (= 1) | (2) WTP (s.d.) | (3) Non-zero WTP (= 1) | (4) WTP (s.d.) | (5) Universal (= 1) | (6) No. Tenants |
| Failed 1 Attn. Check (= 1) | 0.054 (0.045) [0.233] | 0.162 (0.107) [0.131] | 0.025 (0.063) [0.687] | -0.051 (0.141) [0.715] | 0.033 (0.045) [0.458] | -0.062 (0.347) [0.859] |
| Posterior ≥ 0.9 | | | ✓ | ✓ | | |
| Observations | 1,800 | 1,800 | 1,058 | 1,058 | 1,800 | 1,800 |

Note: This table shows the effect of failing one attention check on outcomes in Experiments 1, 2, and 4. We included three attention checks in the survey: participants were dropped if they failed two, but remained in the survey if they only failed one. 7.1% (7.4%) of participants in the final sample who saw rights goods (benchmarks) failed one attention check. Parentheses show robust standard errors. Brackets show p -values.

Table A26: Mechanisms Experiment: Balance and Attrition

| | Effectiveness | | | Stakes | | | Identity | | | p-values | | |
|----------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------|-------|-------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | Pooled | High Stakes | Low Stakes | Pooled | High Stakes | Low Stakes | Pooled | High Stakes | Low Stakes | 2=3 | 5=6 | 8=9 |
| White non-Hispanic | 0.63 [0.48] | 0.62 [0.48] | 0.63 [0.48] | 0.62 [0.49] | 0.62 [0.49] | 0.62 [0.49] | 0.62 [0.49] | 0.62 [0.49] | 0.61 [0.49] | 0.90 | 0.95 | 0.53 |
| Liberal | 0.51 [0.50] | 0.53 [0.50] | 0.49 [0.50] | 0.52 [0.50] | 0.52 [0.50] | 0.51 [0.50] | 0.52 [0.50] | 0.51 [0.50] | 0.52 [0.50] | 0.08 | 0.57 | 0.79 |
| Income > 60k | 0.56 [0.50] | 0.57 [0.50] | 0.56 [0.50] | 0.55 [0.50] | 0.54 [0.50] | 0.57 [0.50] | 0.55 [0.50] | 0.56 [0.50] | 0.55 [0.50] | 0.83 | 0.13 | 0.82 |
| Less than Bachelor's | 0.43 [0.49] | 0.42 [0.49] | 0.43 [0.50] | 0.44 [0.50] | 0.44 [0.50] | 0.44 [0.50] | 0.44 [0.50] | 0.43 [0.50] | 0.45 [0.50] | 0.68 | 0.66 | 0.50 |
| Female | 0.57 [0.49] | 0.57 [0.50] | 0.58 [0.49] | 0.56 [0.50] | 0.56 [0.50] | 0.57 [0.50] | 0.57 [0.50] | 0.56 [0.50] | 0.58 [0.49] | 0.52 | 0.76 | 0.55 |
| Less than age 40 | 0.57 [0.49] | 0.57 [0.49] | 0.57 [0.50] | 0.57 [0.50] | 0.58 [0.49] | 0.55 [0.50] | 0.56 [0.50] | 0.56 [0.50] | 0.55 [0.50] | 0.91 | 0.27 | 0.67 |
| Attrition Rate (%) | 0.03 | 0.04 | 0.03 | 0.03 | 0.04 | 0.02 | 0.03 | 0.03 | 0.03 | 0.947 | 0.103 | 0.355 |
| <i>F</i> -statistic | | | | | | | | | | 0.643 | 0.739 | 0.279 |
| <i>p</i> -value | | | | | | | | | | 0.696 | 0.618 | 0.947 |
| Observations | 1,984 | 984 | 1,000 | 2,004 | 1,012 | 992 | 2,012 | 1,000 | 1,012 | 1,984 | 2,004 | 2,012 |

Note: Columns (1)-(9) show demographic characteristics for by arms in the mechanisms experiment, pooled and separately. Columns (10)-(12) show the *p*-values of the differences between high and low stakes for each comparison. The *F*-statistic is from a joint test of significance for the listed demographic variables. Brackets show standard deviations.

Table A27: Tests of Mechanisms

| | Within Comparison | | | | | | Between Comparison | | |
|------------------------------------|------------------------------|-----------------------------|------------------------------|------------------------------|------------------------------|------------------------------|-------------------------------|-------------------------------|------------------------------|
| | (1) Non-zero WTP (= 1) | (2) Max WTP (= 1) | (3) WTP (\$) | (4) WTP (\$) | (5) WTP (\$) | (6) WTP (\$) | (7) WTP (\$) | (8) WTP (\$) | (9) WTP (\$) |
| High Stakes (= 1) | 0.187 (0.013) [0.000] | 0.091 (0.009) [0.000] | 29.180 (1.964) [0.000] | | | | 34.791 (3.370) [0.000] | 41.511 (3.021) [0.000] | 41.511 (3.021) [0.000] |
| High Stakes \times Effectiveness | | | | 34.791 (3.370) [0.000] | | | | | -6.720 (4.526) [0.138] |
| High Stakes \times Magnitude | | | | | 41.511 (3.021) [0.000] | | | | |
| High Stakes \times Identity | | | | | | 11.686 (3.655) [0.001] | -23.105 (4.972) [0.000] | -29.825 (4.742) [0.000] | |
| Identity (vs. Effectiveness) | | | | | | | 27.534 (3.297) [0.000] | | |
| Identity (vs. Magnitude) | | | | | | | | 38.645 (3.039) [0.000] | |
| Effectiveness (vs. Magnitude) | | | | | | | | | 11.112 (2.718) [0.000] |
| Mean (low stakes) | 0.353 | 0.115 | 50.266 | 44.660 | 33.548 | 72.194 | 58.509 | 53.064 | 39.127 |
| Observations | 6,000 | 6,000 | 6,000 | 1,984 | 2,004 | 2,012 | 3,996 | 4,016 | 3,988 |

Note: This table offers formal tests for the results of the mechanisms experiment shown in Figure 7 (Section 5). Columns (1)-(3) pool all three comparisons into high and low stakes and show the effect of seeing the high stakes arm on 3 measures of WTP. Columns (4)-(6) display results for each comparison, and show the effect of seeing the high stakes arm on WTP. Columns (7)-(9) show pairwise (between comparison) difference-in-difference results. Parentheses show robust standard errors. Brackets show p -values.

Table A28: Attention in Mechanisms Experiment

| | (1) Time on Q1 (seconds) | (2) Time on Q1 (seconds) | (3) Time on Q1 (seconds) | (4) Time on Q1 (seconds) |
|------------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| High Stakes (= 1) | -0.315 (0.720) [0.662] | | | |
| High Stakes \times Effectiveness | | -0.353 (1.343) [0.793] | | |
| High Stakes \times Magnitude | | | -0.205 (1.093) [0.851] | |
| High Stakes \times Identity | | | | -0.349 (1.297) [0.788] |
| Mean (low stakes) | 22.874 | 23.929 | 21.789 | 22.895 |
| Observations | 6,000 | 1,984 | 2,004 | 2,012 |

Note: This table replicates columns (3)-(6) of Table [A27](#) with a new dependent variable: the number of seconds spent on the first choice in the WTP elicitation (to re-run the lottery and save \$20, or to preserve the lottery results). Parentheses show robust standard errors. Brackets show p -values.

C Proofs

C.1 Proof of Proposition 1

Proof. (\Leftarrow): If the distinguishing condition is satisfied, there exists an inverse g such that $g(v_1(x, y), \dots, v_I(x, y)) = \rho(x, y)$. Then put

$$\tilde{F}(v_1, \dots, v_I, \rho(x, y)) = \hat{F}(v_1, \dots, v_I, x, y) \quad (21)$$

$$= \tilde{F}(v_1, \dots, v_I, g(v_1(x, y), \dots, v_I(x, y))) \quad (22)$$

$$\equiv F(v_1, \dots, v_I) \quad (23)$$

where the first line follows by the other hypothesis, and the last line gives the desired replicating welfarist SWF.

(\Rightarrow): Suppose, by hypothesis, there exists a replicating SWF F , so

$$F(v_1(x, y), \dots, v_I(x, y)) = \hat{F}(v_1, \dots, v_I, x, y). \quad (24)$$

Then, there always exists *some* $\rho(x, y)$, which preserves the replicating relationship

$$F(v_1(x, y), \dots, v_I(x, y)) = \tilde{F}(v_1, \dots, v_I, \rho(x, y)) \quad (25)$$

but $\rho(x, y)$ does not necessarily distinguish the utility vector. For instance $\rho(x, y)$ could always be the identity function. We now show how to start from $\rho(x, y)$ and construct $\hat{\rho}(x, y)$ which distinguishes the utility vector and preserves the replicating relationship.

First, choose $\rho(x, y)$ as the identity function. If $\rho(x, y) = (x, y)$ distinguishes the utility vector, we are done.

Otherwise, suppose that $\rho(x, y) = (x, y)$ does not distinguish the utility vector. Then, there exists $\rho(x, y) \neq \rho(x', y')$, yet $v_i(x, y) = v_i(x', y')$ for all i and for some x, x', y, y' . In this case, we have by the definition of a replicating SWF that

$$\tilde{F}(v_1, \dots, v_I, \rho(x, y)) = F(v_1(x, y), \dots, v_I(x, y)) \quad (26)$$

$$= F(v_1(x', y'), \dots, v_I(x', y')) \quad (27)$$

$$= \tilde{F}(v_1, \dots, v_I, \rho(x', y')). \quad (28)$$

To each distinct utility vector $(v_1(x, y), \dots, v_I(x, y)) = \vec{v}_0$, associate a representative (x_0, y_0) such that $(v_i(x_0, y_0))_{i \in \mathcal{I}} \equiv \vec{v}_0$.³⁰ Then, the function $h(x, y) = (x_0, y_0)$ is well-defined. That is,

³⁰At points (x, y) that uniquely produce \vec{v}_0 , there is only one such choice. But at other vectors \vec{v}_0 such that many (x, y) produce the same utility vector, there are many possible selections of (x_0, y_0) .

$h(x, y)$ maps each (x, y) to a reference point (x_0, y_0) such that $(v_i(x, y))_{i \in \mathcal{I}} = (v_i(x_0, y_0))_{i \in \mathcal{I}} = \vec{v}_0$.³¹

Proceed to define $\hat{\rho}(x, y) = \rho(h(x, y))$. The function $\hat{\rho}$ moves any possible points $\rho(x, y)$ which would map to the same utility vector, and instead maps those points (x, y) to the same value $\rho(x_0, y_0) = \rho_0$.

The idea behind the above step is that by reducing the size of the image $\hat{\rho}(x, y)$, we produce a function that distinguishes the utility function. Observe that $\hat{\rho}(x, y)$ indeed distinguishes the utility vector, as $(v_i(x, y)) = (v_i(x', y'))$ implies that $h(x, y) = h(x', y')$ so $\hat{\rho}(x, y) = \hat{\rho}(x', y')$.

To complete the proof, we claim that $\hat{\rho}(x, y)$ preserves the relationship

$$\tilde{F}(v_1, \dots, v_I, \hat{\rho}(x, y)) = F(v_1, \dots, v_I), \quad (29)$$

since:

$$\tilde{F}(v_1, \dots, v_I, \hat{\rho}(x, y)) = \tilde{F}(v_1, \dots, v_I, \rho(h(x, y))) \quad (30)$$

$$= \tilde{F}(v_1, \dots, v_I, \rho(x, y)). \quad (31)$$

For if $\rho(h(x, y)) = \rho(x, y)$, then Equation (31) follows immediately. Otherwise if $\rho(h(x, y)) \neq \rho(x, y)$, then $(v_i(h(x, y))) = (v_i(x, y))$, so

$$\tilde{F}(v_1, \dots, v_I, \rho(h(x, y))) = \tilde{F}(v_1, \dots, v_I, \rho(x, y)) \quad (32)$$

follows from Equation (28). □

C.2 Proof of Proposition 2

Proof. Consider an allocation $\tilde{Y} \equiv (\tilde{y}_1, \dots, \tilde{y}_J)$ that maximizes welfarist social welfare for a given set of welfare weights, utility functions, and incomes. Thus \tilde{Y} solves

$$\max \sum_{j=1}^J \gamma_j u(x_j, y_j), \text{ such that } \sum_j y_j \leq m. \quad (33)$$

Call $\Delta u(y_j = k) \equiv u(x_j, k) - u(x_j, k-1)$, which is the difference in private utility generated from providing the k th good to individual j . Call $\Delta v(y_j = k) \equiv \gamma_j \Delta u(y_j = k)$, which is the difference in welfarist social welfare from providing the k th good to individual j .

Now consider the $Y^* \equiv (y_1^*, \dots, y_J^*)$ which solves

$$\max \sum_{j=1}^J \gamma_j (u(x_j, y_j) + \eta_y \mu_i(y_j, r_{ij})) , \text{ such that } \sum_j y_j \leq m, \quad (34)$$

³¹To understand $h(x, y)$, observe that each (x, y) pair belongs to an equivalence class, defined by pairs that generate the same utility vector \vec{v} . The function h maps (x, y) to a reference point (x_0, y_0) within the equivalence class.

where $\mu_i(\cdot)$ is as in Equation (12).

As in the text, let $\tilde{\mathcal{J}} \equiv \{j \leq J : \tilde{y}_j > 0\}$ and $\mathcal{J}^* \equiv \{j \leq J : y_j^* > 0\}$. We want to show that $\tilde{\mathcal{J}} \subseteq \mathcal{J}^*$.

As is standard, we assume $\eta > 0, \lambda > 1$, and that $u(\cdot)$ is weakly concave in its inputs (thus, $\Delta u(y_j = 1) \geq \Delta u(y_j = k)$ for all $k > 0$). We also assume $r_j = 1$ for all j , and that y is always positive.

Comparing the maximand in Equation (34) to Equation (33), the marginal welfare generated by the first y provided to any individual $j \notin \tilde{\mathcal{J}}$ increased by $\eta\lambda\Delta v(y_j = 1)$ and the marginal welfare generated by providing $k > 1$ to any individual $j \in \tilde{\mathcal{J}}$ increased by $\eta\Delta v(y_j = k)$.

There are two possible cases. First, \tilde{Y} is already “flat,” that is, $\max_{j \leq J}(\tilde{y}_j) = 1$. we show Y^* must also be flat. Second, \tilde{Y} is not flat. Then, either (i) there exists a reallocation such that $\tilde{\mathcal{J}} \subsetneq \mathcal{J}^*$, or (ii) there is no reallocation and $\tilde{\mathcal{J}} = \mathcal{J}^*$.

Case 1. Consider reallocating $y = 1$ from j' to j , for $j', j \in \tilde{\mathcal{J}}$. Notice that the relative welfare value of providing a k th y to any individual $j \in \tilde{\mathcal{J}}$ relative to a first y to j' has decreased by $\eta(\lambda\Delta v(y_{j'} = 1) - \Delta v(y_j = k))$, comparing the maximand in Equation (34) to Equation (33). Thus, if \tilde{Y} is flat, then Y^* is also flat.

Case 2. Comparing the maximand in Equation (34) to Equation (33), the marginal welfare generated by providing the first good to each j' for whom $\tilde{y}_{j'} = 0$ has increased by $\eta\lambda\Delta v(y_{j'} = 1)$. Let $\underline{j} \equiv \arg \min_j \{\Delta v(\tilde{y}_j = k)\}$ for $k > 0$. That is, \underline{j} is the individual to whom providing the marginal unit of y gives the least amount of social welfare, yet received at least one unit of y under \tilde{Y} . The marginal social welfare generated by providing the last good to \underline{j} has increased by only $\eta\Delta v(y_{\underline{j}} = \tilde{y}_{\underline{j}})$. Thus, the welfare gain from flattening provision by allocating toward any $j' \notin \tilde{\mathcal{J}}$ has increased by $\eta(\lambda\Delta v(y_{j'} = 1) - \Delta v(y_{\underline{j}} = \tilde{y}_{\underline{j}}))$. If that increase is sufficiently large relative to the initial gap in marginal social welfare, then Y^* will involve reallocating toward individuals in the loss domain and \mathcal{J}^* will be strictly flatter than $\tilde{\mathcal{J}}$. Otherwise, the set of people receiving the good does not change and is weakly flatter. \square

D Experiment Details

D.1 Survey Recruitment

We recruited participants on Prolific. We advertised for an 18-minute survey entitled “Research Study,” with \$6 compensation. We restricted potential participants to fluent English speakers in the U.S. The median survey completion time was 19 minutes and 12 seconds.

D.2 WTP Elicitation Details

Our main WTP elicitation measures in Experiments 1, 2, and 3 are multiple price lists. We show one question at a time and impose monotonicity. Experiment 1 starts at \$20 and moves in \$20 increments. Experiment 2 starts at \$100 and moves in \$100 increments. Experiment 3 starts at \$500 and moves in \$100 increments. See Figures 12, 15, and 22 in the Experiment Instructions online for examples of each.

D.3 Incentives

Rights Good Incentivization and Framing. We incentivize choices for rights goods in two ways. First, for Spectators who are assigned to do the experiments with lawyers, we use the strategy method. Second, we inform Spectators who do the experiments with both lawyers and health care that the nonprofit will be informed about participants' choices, which may impact their policies. Specifically, we inform Spectators doing experiments with lawyers:

"Sometimes we will ask you what type of assistance to provide to tenants facing eviction. Please take these questions very seriously. Some participants will be randomly chosen to have their answers made in real life. **If you are chosen, your answers here will have significant impacts on the lives of real people**, so please take your time and respond truthfully."

We implement choices via a nonprofit partner in Memphis that works with tenants. Health care is not incentivized because it was not possible to implement. Instead, we tell these Spectators:

"We will present you with a series of hypothetical scenarios and ask you what type of assistance to provide to tenants facing eviction. Please take these questions very seriously. A Memphis nonprofit who helps tenants facing eviction will be informed of participants' opinions on resource allocation."

Spectators then see contextual information about either lawyers³² or health care³³, and then do a comprehension check about that fact. We then introduce the types of assistance the nonprofit provides: either lawyers *or* health care depending on which good they see, bus passes, YMCA memberships, and cash.

³²We tell these Spectators: "Evictions often end up in court. In court, **tenants usually do not have lawyers**, because they are usually low-income and cannot afford them. **Landlords usually do have lawyers**." The government guarantees attorneys to anyone charged with a crime, but usually does not provide lawyers in civil settings, including eviction cases."

³³We tell these Spectators: "Most low-income households in the United States say that they or a family member in their household delayed or went without some type of medical or dental care in the past year because they had difficulty affording the cost."

Benchmark Incentivization. Participants doing experiments with YMCA memberships or bus passes are randomized into seeing lawyers or health care as an alternate assistance option that the nonprofit provides. Based on which framing they see, their incentives follow that of lawyers or health care. Half of Spectators who see the lawyers framing are randomized into hypothetical good valuation.

Implementation. We implement one choice from each of three categories: features of rights (combining Experiments 1–3), anti-targeting (Experiment 4), and WTP to directly give the good. Only choices over lawyers, or choices over bus passes and YMCA with incentivized framing, are eligible. We randomly select a choice from each category from a dataset of all choices from the study as well as incentivized pilots. This dataset contains approximately 51,000 choices over features, 15,000 choices over anti-targeting, and 8,000 choices over WTP. These include both actual and implied choices from each multiple price list—we only display enough questions to identify a participant’s indifference point, but include implied choices so that a participant’s probability of being selected does not depend on their choices. If we selected a choice from a pilot, only choices from the main study were eligible to be selected in the remaining categories (in practice, the randomization never selected a choice from a pilot). Once a choice over a specific good was selected, we drop all other choices over that good. This is to ensure we implement choices for each of the different types of goods in our study.

Incentivization Tests. We provide more details on the tests in Section 4.5. First, we test the effect of incentivization on WTP in our 3 main experiments among benchmark Spectators. There is no effect of incentivization on outcomes in inalienability, choice, or anti-targeting experiments (Table A22). Moreover, our main results looks similar using either only incentivized or unincentivized Spectators as the control group.

Our second test of incentivization is within incentivized participants. Half of Spectators seeing lawyers or benchmark goods are disincentivized in their choices about their WTP for the good. We inform selected participants:

“At the beginning of the survey, we told you some choices will be randomly selected to be implemented for real tenants. In this section, all choices will be purely hypothetical. However, your choices are still important and the nonprofit will be informed of the results.”

This is the last incentive-eligible section of the survey. We compare Spectators’ valuations and find no difference by incentivization (Table A23).

Belief Elicitations. We further incentivize both belief elicitation questions that are in the survey. At the start of the survey, we inform only incentivized participants:

“Sometimes we will ask you to predict what choices tenants have made or the impacts

programs affecting tenants have had. Please take these questions very seriously. Some participants will be randomly chosen to be paid bonuses if the answers they give are close enough to the truth. **If you are chosen, your answers could increase your participation bonus**, so please take your time and respond truthfully.”

First, in Experiment 2, we elicit prior beliefs and posterior beliefs after information about the percent of tenants expected to choose \$y in cash over the good. All incentivized participants (seeing lawyers, or seeing YMCA or bus passes with lawyers framing) have this belief incentivized. Second, in Experiment 4, we elicit prior beliefs participants are given information about the efficacy of lawyers or health care. Here, incentivization is only for Spectators who see experiments about lawyers. In both modules, participants are informed: “Choose your responses carefully. You can earn bonuses for correct answers! You may request more details if you are curious about how the payment works.” Interested participants saw that they would be enrolled in a lottery with a 10% chance of winning; for one up their upcoming predictions, they are compensated \$1 in Spring 2024 if their answer is within 4 percentage points of the correct answer.³⁴ In practice, all bonuses for selected participants are issued based on accuracy of prior beliefs, before participants were shown information.

D.4 Attention Checks

We include three attention checks. The first attention check asks participants to select a specific multiple choice option. The second two checks provide a list of cities and their populations, and ask participants to rank them from most to least populous.

Participants exit the survey if they fail two attention checks. 12 participants began our survey and were dropped because of this restriction (Table A2). 7% of our final sample of 1,800 participants failed only one attention check. 41 failed the first check, 52 failed the second check, and 36 failed the third check.

We have several additional comprehension checks throughout the survey, which cannot be the basis for dropping participants per Prolific policy. Participants are always informed of the correct answer after completing a comprehension check.

D.5 Experiment 1 Details

To elicit participants’ WTP for inalienability, we set up a scenario where two eligible, comparably needy tenants are entered in a lottery for one available good. We explain that: “After the lottery takes place, but before the tenants are informed of the outcome, the nonprofit reserves the ability

³⁴Regarding the Spring 2024 payment, we were not able to issue bonus payments at the time of the survey itself because the accuracy of correct answers was, in some cases, determined by future events, so we made the payment once the events realized.

to rerun the lottery in some cases. Sometimes prices change and money can be saved by rerunning the lottery and assigning the [good] to whoever wins the second lottery.” This sets up the choice between either (1) leaving the lottery results as they are or (2) taking the good away from the original winner, giving it to whoever wins the second lottery, and saving some amount of money for future programs. We then reinforce this setup by asking confirmation questions emphasizing, first, money is saved when the lottery is rerun, and second, the tenants will not know the original allocation if the lottery is rerun.

We first ask Spectators if they would prefer to keep the lottery results or rerun the lottery and save \$20. We repeatedly ask Spectators the same question in increments of \$20, until either they elect to rerun the lottery or they prefer to keep the results over rerunning and saving \$200.

D.6 Experiment 2 Details

Information Treatment. We begin by eliciting beliefs about the percent of tenants expected to choose cash y over the good. The cash value is randomized $y \in \{\$200, \$300\}$. We ask participants to guess how many tenants, among 100 tenants who apply for assistance, would choose y over the good. We then truthfully inform participants: “Researchers who work with the nonprofit asked 10 tenants whether they would choose a [good] or y in cash. All of them chose to receive y over a [good].” We ensured the truthfulness of this information treatment using additional Prolific experiments, where we screened for tenants and asked about their preferences for all combinations of goods and cash y values. At least 10 tenants for each combination preferred cash. After sharing this information, we elicit posterior beliefs.

Elicitation. Following the information treatment and posterior elicitation, we proceed with the elicitation of WTP for choice (Experiment 2). We begin by informing tenants that the good typically costs \$350, and that the current budget of the nonprofit allocates y in cash to the tenant and saves the rest for future programs. We first ask Spectators if they would prefer to give the tenant y and save \$100, or give the tenant the choice between y and the good. We repeatedly ask Spectators the same question in increments of \$100, until either they elect to give y and save or they prefer to give the tenant the choice over giving y and saving \$900.

D.7 Direct WTP Elicitation Details

We elicit WTP to provide the good directly, which we use in several analyses and robustness checks. To address the potential for anchoring, we randomize the initial choice participants see when eliciting their WTP directly for their assigned good. We ask if they would prefer to give one tenant the good or cash, randomizing the initial cash value from $\{\$300, \$500, \$700\}$. We ask the same question in increments of \$100 until we identify their indifference point.

D.8 Mechanisms Study

Recruitment. We conduct one additional study, described in Section 5. We again recruit participants on Prolific. We advertised for a 5-minute survey entitled “Short Research Study,” with \$1.33 compensation. We restricted potential participants to fluent English speakers in the U.S. who did not participate in our main study. The median time to complete the survey was 5 minutes and 1 second.

Comprehension Checks. We include the two comprehensions checks from Experiment 1, confirming that money can be saved by re-running the lottery (77% respond correctly) and that neither tenant knows who won the original lottery (94% respond correctly). We add an additional comprehension check about what good is being allocated in the lottery (99% respond correctly). 74% of participants respond correctly to all three checks. Per Prolific policy, these checks cannot be the basis for dropping participants. Participants are always informed of the correct answer after completing a comprehension check. The study did not include any general attention checks. Results are unchanged when restricting only to participants who pass all comprehension checks or participants with survey times between the 10th and 90th percentiles (Figure A16).

Experiment Details. We extend the design of Experiment 1 (Section D.5) where two eligible, similarly needy tenants are entered in a lottery for one available good. We restrict the set of goods to lawyers, but change the characteristics of the lawyer or Recipients. We have six arms and three comparisons. Each participant only sees one of the six arms. In the first arm, we vary the effectiveness of the lawyer to be allocated to tenant facing eviction. We either say “The lawyer that will be assigned in the lottery is effective and almost always wins their cases” or “The lawyer that will be assigned in the lottery is ineffective and almost always loses their cases.” In the second arm, we vary the magnitude of the fines in a speeding ticket case. We either specify that *both* people “have ticket fines of \$2,000, which they cannot afford to pay” or “have ticket fines of \$50, which they can afford to pay.” In the third arm, we return to the eviction setting and vary Recipients’ income. We say that either both tenants “have annual incomes of \$20,000 and cannot afford a lawyer” or “have annual incomes of \$80,000 and can afford a lawyer.” The rest of the designs follows as in Experiment 1.

E Additional Analyses

E.1 Second Measure: Value of Choice (Section 3.2)

Explanation. The main measure is valid if beliefs are well-measured, but they may be noisy due to standard elicitation issues. To address this concern, we also form a second measure of a non-welfarist willingness to pay for choice. Welfarists value providing the choice of good g and

cash \$y versus cash as:

$$C_i^w = p(-i \text{ chooses } g) \times E[f(u_{-i}(g)) - f(u_{-i}(y)) \mid -i \text{ chooses } g]. \quad (35)$$

This expression says that welfarists value choice at their WTP to provide the good, times the probability of exercising the choice. We obtain the value $E[f(u_{-i}(g)) - f(u_{-i}(y)) \mid -i \text{ chooses } g]$ by eliciting Spectators' willingness to provide g versus cash to a Recipient. We assume small selection on gains, such that

$$E[f(u_{-i}(g)) - f(u_{-i}(y))] \approx E[f(u_{-i}(g)) - f(u_{-i}(y)) \mid -i \text{ chooses } g]. \quad (36)$$

We study the effect of rights goods on Δ_i , the difference between actual WTP for choice C_i and welfarist implied WTP for choice C_i^w :

$$\Delta_i(t) := \mathbb{1}(C_i - C_i^w - t > 0). \quad (37)$$

Setting tolerance $t = 0$ lets us examine whether the elicited WTP for choice is exactly equal to the welfarist WTP for choice. We focus on $t \gg 0$, to conservatively account for trembles (e.g., imperfect ability to scale WTP by beliefs) and selection on gains.³⁵

The advantage of the second measure relative to the first measure is that it does not require us to condition on having high beliefs. The disadvantage is that we lack a principled way to choose t . If t is too small and perceived selection on gains is large, the test is invalid. We use $t = \$250$ and show robustness to this decision.

Results. The second measure of WTP for choice corroborates the primary measure (Figure A3). We find that rights goods have larger $\Delta_i(t)$ for all tolerances between \$0 and \$500. For instance, focusing on a tolerance of \$250, we find that Spectators exposed to rights goods are 6.4 pp more likely to have a WTP for choice that exceeds their instrumentalist WTP by \$250 or more (s.e.: 1.7). The value $\Delta_i(t)$ is guaranteed to decrease in t . But the difference in $\Delta_i(t)$ for rights versus benchmark goods remains large as a share of benchmark goods' $\Delta_i(t)$.

E.2 Intrinsic Value of Choice (Section 3.2)

Two further pieces of evidence suggest that intrinsic value of choice does not explain our results. First, we use our qualitative question about why participants made their choice to do an informal

³⁵Setting $t > 0$ also accounts for minor elicitation differences between C_i and C_i^w . In particular, C_i is WTP in units of dollars of future programming for the nonprofit. C_i^w is WTP in units of dollars of money provided directly to that tenant. Crucially, no matter how large we set t , we find differences in $\Delta_i(t)$ across goods. Relatedly, both C_i and C_i^w are subject to top-coding in the multiple-price list. To be conservative and push against finding a large $\Delta_i(t)$, we top code the maximum direct WTP at \$1,500, whereas we top code C_i at \$950. Note that top coding across does not introduce bias unless differential by good.

mediation analysis. Among participants who say that “saving is their priority,” the treatment–control gap in willingness to pay to provide choice is highly attenuated (to 0.1 s.d. from 0.5 s.d., Table A11). This result suggests that the impact of rights goods operate primarily through changing people’s priorities to become less consequentialist and thus place less emphasis on savings — indicative of non-welfarism.

Second, if the intrinsic value of choice is non-additive, it is reasonable to expect it is correlated with the WTP to pay for the good directly. When we condition the rights–benchmark difference in WTP for choice on fixed effects for the valuation to give the good directly, the results drop but remain meaningful (0.26 s.d., Table A10). We expect some attenuation, as WTP to give the good directly may be correlated with non-welfarist concerns (a “bad control,” per Angrist and Pischke, 2009). Thus, the intrinsic value of choice must be both non-additive (else it is netted out with rights versus benchmark differences) and uncorrelated with direct WTP, which seems implausible. Most models of non-additive WTP would have value of choice scale with direct WTP (e.g., if the intrinsic value of choice comes from rejecting goods that are valuable).

The magnitudes that we find also cast doubt on intrinsic value of choice being the complete explanation. Freundt et al. (2023) estimate that choice is intrinsically valued at a modest 5% of the utility from the decision on average. If that 5% number holds in our setting, Spectators would need to overestimate Recipients’ intrinsic value of choice: across all treatments, they are willing to pay hundreds of dollars for choice, much more than 5% of their willingness to pay to give the good directly (see Figure A8 for average WTPs). In fact, the intrinsic value of choice could push against our results, as Spectators could want to choose for Recipients if Spectators themselves value being in control, like in Fehr et al. (2013).

What if our results do partially reflect intrinsic value of choice? Then, the inherent circularity between SWFs, Spectators, and Recipients becomes blurry. Such behavior could be viewed as welfarist, in the sense that Spectators aggregate Recipients’ utilities. However, such views are arguably non-welfarist on the part of Recipients, as they destroy surplus for their own moral concerns.³⁶ Then, welfarist SWFs, when aggregating Spectators’ preferences, would include such non-classical motivations or moral concerns. Ultimately, while we think this explanation is unlikely to explain the results, it would lead to a similar ultimate conclusion for welfare economics.³⁷

E.3 Sub-Experiment: Selection on Gains (Section 3.2)

As noted in the text, selection on gains is only a concern if beliefs p are mismeasured. Moreover, if mismeasurement is additively separable and equal across goods, then many possible stories are netted out with our comparison between rights versus benchmarks. But one can create stories

³⁶We allude to moral concerns because Recipients would need to value making choices over rights good differentially from benchmarks. This would not be the case for Recipients who hedonically valued making choices.

³⁷A caveat pertains to whether Spectators’ accurately perceive Recipients’ utilities.

where issues such as inattention or beliefs mismeasurement vary with the good.

We randomize the value of the cash option $y \in \{\$200, \$300\}$. Intuitively, randomizing the cash option traces a WTP curve to provide choice. Specifically, for higher values of cash, if selection on gains is meaningful, then Spectators should be much more likely to choose cash when the cash value rises.

We find that posteriors are only modestly associated with the cash value (Table A13), suggesting against important selection on gains. Focusing on rights goods (Column 1), a \$100 increase in cash is associated with a 2.1 pp increase in posterior beliefs that the Recipient chooses cash (s.e.: 1.5). And focusing on people who already express a posterior belief ≥ 0.9 , results attenuate even further (Column 2). While some selection on gains is evident from the downward sloping curves in Figure 3, it is not top-of-mind for Spectators and is unlikely to explain the large WTP for choice that we find.

E.4 Incidence Quantification 2 (Section 4.3)

We posit that social planner's utility from giving \$1 to Spectator i is:

$$W_{\text{non-welfarist}} = (1 - \beta)\lambda_1[v'_i(c) - v'_j(c)] + \beta\lambda_2 [\Pr(i \text{ is non-welfarist}) - \Pr(j \text{ is non-welfarist})], \quad (38)$$

where $v'_i(c)$ is i 's marginal utility of a dollar. The parameter $\beta \in [0, 1]$ represents the importance the planner places on non-welfarist versus welfarist concerns, and the λ 's scale each type of utility.

Equation (38) suggests the possible trade-off between progressivity in marginal utility and promoting non-welfarist preferences. If i is poorer than j , the first term is positive. But if i is less likely to be non-welfarist, the second term is negative. This Social Welfare Function is microfounded if one thinks of the non-welfarist SWF as wanting to redistribute money to someone who is themselves non-welfarist, in the spirit of Rabin (1993). But one can also interpret it as a reduced-form SWF, which lets us extract the price of non-welfarism in units of redistribution.

We simulate $W_{\text{non-welfarist}}$ using several assumptions. We imagine five people in society, with average incomes at the mid-points of each bin in Figure 6A. We imagine taking from person 3 and giving to one of the others. We posit v_i has CRRA utility in income with $\gamma = 2$. We normalize the λ 's so that the social planner's utility of giving to the lowest-income person is 1 if $\beta = 0$ (since $v'_1(c) > v'_3(c)$), and -1 if $\beta = 1$ (since group 3 is more non-welfarist than group 1). Then, for any β , $W_{\text{non-welfarist}}$ is completely determined from Figure 6A.

We find a trade-off between redistribution and prioritizing non-welfarist concerns (Figure A13). For low weight on non-welfarist concerns β , the social planner prefers to give to the lowest-income person. But for $\beta > 0.5$, she prefers to give to the middle-income person. Redistribution to the second, fourth, or richest person is always dominated by giving to group 1 or 3.

E.5 Information Provision Sub-Experiment

We have documented that non-welfarist preferences are common. But how do they compare to welfarist preferences? To assess the importance of welfarism, a sub-experiment provides information about the goods.

Design. Our goal is to see how Spectators change redistributive choices based on information. This question is intrinsically important, as support for transfer programs could partially stem from beliefs about their efficacy. Ultimately, we use this exercise to label Spectators as “welfarist” if they do revise redistributive choices when this information conflicts with their prior beliefs.

The simple idea is to give truthful information about the efficacy of lawyers and health care. Then we ask Spectators if they want to change redistributive choices based on the information.

We implement this design as follows. Spectators assigned to lawyers or health care are randomized into seeing information that the good is effective or ineffective. There is no equivalent experiment for the benchmark goods.

We tried to select information that is relevant to welfarist Spectators who maximize a SWF that only depends on Recipients’ utilities. Before giving information, we elicit prior beliefs about the efficacy of lawyers and health care.

For lawyers, we elicit priors in two parts: first asking how many out of 100 tenants without a lawyer receive an eviction judgement, and second asking the same for tenants *with* lawyers. We then tell Spectators in the [high/low] treatments: “Researchers studied a program that is providing lawyers to tenants facing eviction in Memphis. Among 100 of the tenants, having a lawyer led to a [80/20]% reduction in eviction rates. About 55% of tenants who did not receive a lawyer from the program were evicted in court, but only about [15/45]% of tenants who did receive one were.” These estimates were based on a concurrent RCT of providing lawyers to tenants facing eviction in Memphis, TN (Caspi and Rafkin, 2023). Both pieces of information are truthful because we redrew samples of 100 tenants that matched the reported numbers.

For health care, we ask Spectators how many of 100 tenants without health care vouchers will have improved health outcomes 1 year later. We tell Spectators positive and null results from the Oregon Health Insurance Experiment (Baicker et al., 2013; Allen et al., 2013). In the high information treatment, Spectators see: “Researchers studied Medicaid expansion in Oregon and found that among people who newly gained access to Medicaid, rates of depression fell by 9 percentage points and increased the likelihood of self-reporting health as good, very good, or excellent by 13 percentage points.” In the low information treatment, Spectators see: “Researchers studied Medicaid expansion in Oregon and found that among people who newly gained access to Medicaid, it did not have a significant effect on measured blood pressure or cholesterol.”

After providing information, we let Spectators choose whether to revise their targeting choice in Experiment 4. (Results in Section 4.1 report Spectators’ *initial* choices.) In particular, we ask

Spectators assigned to lawyers: “Previously, you made choices distributing a limited budget across hiring lawyers and giving tenants cash. Given this information, would you like to revise any of your choices?” Spectators who say they want to revise their choice then do the same targeting elicitation from Experiment 4. The setup is similar for health care.

Results. Upon receiving information, it is fairly rare to change the allocation of lawyers or health care in the targeting task (Figure A5). Information causes about 40% to change their allocation in the targeting task if they receive information about lawyers. Less than 30% change if they receive information about health care. Conditional on changing, Spectators tend to update in the direction of the information shown (e.g., they provide more lawyers if they get positive information about lawyers). As most Spectators do not change their allocation even when we surprise them with relevant information about the good’s consequences on recipients, some are probably not maximizing welfarist objectives.

We extract the causal relationship between beliefs and allocations by instrumenting for beliefs with the information treatment (Table A19). A 10 pp increase in beliefs about lawyers’ efficacy (half the median absolute difference between the information treatment and prior beliefs) causes Spectators to allocate to 0.88 more tenants (s.e.: .10). The impact of information is therefore economically meaningful. Even so, welfarist concerns cannot fully explain why so many Spectators provide goods universally.

Classification. We now classify Spectators as welfarist or non-welfarist. We focus on Spectators who had beliefs about lawyers/health care efficacy that disagree with information we provided them in the treatment. Among these Spectators, we label them as *welfarist* if they revise their targeting decision.³⁸ Spectators are *non-welfarist* if they exhibit positive WTP in Experiments 1–2. We then show how these classifications change if we adjust assumptions about how to label Spectators who express both non-welfarism and welfarism. As we use just three decisions to classify Spectators, we prefer a simpler approach rather than machine-learning clustering methods.

We restrict to a constant sample of Spectators who do experiments with rights goods and whom we could have observed as welfarist. That is, we only include Spectators for whom information goes against priors and the initial choice in Experiment 4. For instance, we exclude Spectators who make a universal choice and also see lawyers as surprisingly effective. In this case, we could not observe them make a welfarist decision to allocate more lawyers.

Even under conservative assumptions, welfarism is not much more common than non-welfarism (Figure A6). In the restricted sample, 71% exhibit non-welfarist preferences in at least one of Experiments 1–2, and 28% exhibit non-welfarist preferences in both experiments.

³⁸We elicited beliefs for lawyers that exactly correspond to the information treatment. Beliefs for health care were not identical to the information we provided. We label these Spectators as having priors that exceed the information if priors about how health care vouchers increase the percent of tenants with improved health outcomes 1 year later are above the analogous percent increase for lawyers (80% for high information, 20% for low information).

Meanwhile, 34% are welfarist, revising targeting choices after receiving information.

As a lower bound on the importance of welfarism, we label people as welfarists only if they revise after information, and do not exhibit non-welfarist preferences in *either* Experiments 1–2. Then, 10% are welfarist.

As a lower bound on the importance of non-welfarism, we label people as non-welfarists only if they *do not* revise after information, and make non-welfarist choices in *both* Experiments 1–2. Then, 18% are non-welfarist.

Overall, the experiments suggest that non-welfarism is quantitatively significant. All the data from Experiments 1, 2, and 4 suggest at least 18% have non-welfarist preferences, and less stringent interpretations of the data can push this number to 71%. Even to a skeptic, then, non-welfarism is more than half as prevalent as welfarism (18% versus 34%).

We caution against overly literal interpretation of the classification exercise. A particular limitation is that we assume that we measure all relevant beliefs about lawyers or health care’s welfarist impacts. For instance, we understate the amount of welfarism if Spectators care about lawyers beyond their impacts in court. This concern is mitigated if beliefs across outcomes are correlated. Similarly, we only measure a few of many possible non-welfarist views, so we probably misclassify some true non-welfarists.

Relation to the Literature. The sub-experiment replicates the general finding in the literature on information-provision experiments that information has a moderate impact on policy support (see references and discussion in Haaland et al., 2023 and Yang, 2023). Here, Spectators often do not update their allocations even upon seeing information that would likely be relevant if all choices were made for welfarist reasons.³⁹

Non-welfarist preferences help explain the moderate impacts in the literature. Information treatments often focus on policy consequences. However, strong non-welfarist concerns would mean that people’s support is less likely to hinge on consequences.⁴⁰

Correlation: Information Sub-Experiment and Anti-Targeting. While Figure 5A suggests behaviors in Experiments 1–2 predict anti-targeting, it is not guaranteed that non-welfarist preferences, which also require not updating based on relevant information, are still predictive. Yet this concern is unfounded (Figure A9, see Table A20 for standard errors and hypothesis tests). 25% of anti-targeters are non-welfarist versus 14% of non-anti-targeters (p -value of difference = 0.002). Moreover, we also find a sharp drop in welfarist preferences among anti-targeters (20% versus 42%, p -value of difference < 0.001).⁴¹ These correlations cast doubt on random elicitation

³⁹Treatment effects above are somewhat larger than in the literature because we restrict to a sample of Spectators whose redistributive decisions are not mechanically unresponsive to information. In the whole sample, 27% update their choice from Experiment 4 and make a mean absolute change of 0.9 tenants.

⁴⁰This explanation complements others in the literature, including uncertainty about the mapping from beliefs to actions (Yang, 2023) or the importance of (possibly welfarist) redistributive preferences (Stantcheva, 2021).

⁴¹The result that non-welfarism is more prevalent among those who anti-target is mostly robust to alternative

errors as explaining the large share of non-welfarists, since noise would not predict other choices.

Inattention in Information Sub-Experiment. Measurement error from inattention is undoubtedly present, but unlikely to change our conclusions that non-welfarist preferences are prevalent. First, we designed the belief updating task to require an active choice to update or not. Participants must select either: “**Yes**, I would like to revise my choices and give more people lawyers”; “**Yes**, I would like to revise my choices and give more people cash”; or “**No**, I would not like to revise my choices.” It is not obvious that inattentive participants would choose not to update versus choose to update. In fact, inattention might work in the other direction, by leading true non-welfarists not to exhibit non-welfarist preferences in both Experiments 1 and 2.

Second, participants are attentive overall (Section 2). This concern thus requires attention to lapse at precisely this elicitation and essentially nowhere else.

Finally, as our results are large in magnitude, an implausible amount of measurement error is required to undo them entirely. Suppose a full 50% of the Spectators whom we label as non-welfarists are actually inattentive welfarists (who pass two other attention checks). Even then, the share of non-welfarists to welfarists would still be quantitatively meaningful (about 30%).

definitions and not dropping those with low posteriors in Experiment 2 (Table A20), and vice-versa with welfarism. Results are driven by lawyers, and are not significant for non-welfarist preferences if we focus on health care alone (Panels C–D). Some correlations between non-welfarism and targeting attenuate with different definitions of non-welfarism (Columns 1–2). This attenuation relative to Figure 5 is caused by: (1) focusing on the extensive margin (those with I_i or C_i larger than 0), as continuous measures are robustly correlated with anti-targeting (Table A17), and (2) the fact that the sample of people whom we can unambiguously label as welfarist or non-welfarist is smaller.

F Pre-Registration

We pre-registered our main study ([AEA Registry #0012065](#)) and mechanisms study ([AEA Registry #0014218](#)). Section [F.1](#) presents a key between our pre-registrations and where to find the analysis in the paper. Section [F.2](#) discusses the exhibits not included in the preceding text.

F.1 Key to Pre-Registered Outcomes

We specify where in the paper each of our pre-registered outcomes is presented.

F.1.1 Main Experiment

Primary Outcomes

1. Elicitation 1: Targeting [Section [4.1](#)]
 - (a) “Indicator for choosing to allocate the good to all 10 Recipients” [Table [A16](#)]
 - (b) “Number of Recipients whom the Spectator chooses to receive the good (indifference point)” [Table [A16](#), Figure [4](#)]
2. Elicitation 2: Inalienability [Section [3.1](#)]
 - (a) “Indicator for choosing the maximum price to avoid switching the lottery” [Table [A3](#), Appendix Figure [A1](#)]
 - (b) “Willingness to pay to avoid switching the lottery (indifference point)” [Table [A3](#), Figure [2](#)]
3. Elicitation 3: Option value (i.e., Choice) [Section [3.2](#)]
 - (a) “Indicator for having a positive willingness to pay to ensure a choice” [Table [A7](#), Figure [3C](#)]
 - (b) “Willingness to pay to ensure a choice (indifference point)” [Table [A7](#), Figure [3B](#)]
 - (c) “Primary heterogeneity: willingness to pay to ensure a choice, conditional on having a high posterior belief that the recipient would choose cash” [Table [A7](#), Appendix Table [A9](#)]
4. Elicitation 4: Universalism (i.e., Egalitarianism) [Section [3.3](#)]
 - (a) “Amount willing to forgo from outside option to provide the good (indifference point)” [Appendix Table [A14](#), Appendix Figure [A4](#)]

Secondary Outcomes

1. “Correlations between Elicitation 1 outcomes and Elicitation 2, 3, and 4 outcomes (pairwise and joint)” [Table 3, Table F1]
2. “Tagging people as non-welfarist. We will identify people as non-welfarist if they engage in non-welfarist behaviors in any, some, or all of Elicitations 2, 3, and 4. We will report these numbers.” [Figure A6, Figure F1]
3. “Incentives. In Elicitation 1a, we employ a secondary treatment to test whether incentives affect behaviors. In 1a, we randomize whether we incentivize lawyers and placebos. The purpose of this randomization is to test the extent to which incentives matter (insofar as it could generate differences between lawyers and health care).” [Appendix Table A23]
4. “In a second test for the effect of incentives, we leverage the following source of variation. Because the presentation for the health care and lawyers varies slightly (including discussion of incentives), we randomize placebo goods into seeing the exact framing as lawyers versus health care, so some placebo goods are not incentivized throughout (see Experiment Details). This variation jointly tests the effects of presentation and incentives. Since we expect the presentation alone is small, it is also a test of the effects of incentives.” [Appendix Table A22]
5. Elicitation 1a: Valuation
 - (a) “Willingness to pay for the good (indifference point)” [Appendix Figure A8]
6. Elicitation 1b: Information [Appendix E.5]
 - (a) “Share of people who revise choice about how to allocate good”
 - (b) “Magnitude of revision in choice about how to allocate good” [Appendix Figure A5]
 - (c) “Suggestive: instrumental-variables specification that instruments for the effect of beliefs on Elicitation 1-outcomes.” [Table A19]
7. Secondary Heterogeneity
 - (a) “The relationship between political party and exhibiting non-welfarist preferences (behaviors in Elicitations 1–4)” [Figure 6B, Figure F2]
 - (b) “The relationship between personal experience (with facing legal problems without a lawyer or not seeking health care due to cost) and exhibiting non-welfarist preferences (behaviors in Elicitations 1–4)” [Figure 6B, Figure F2]
 - (c) “The relationship between support for Right to Counsel and health care and exhibiting non-welfarist preferences (behaviors in Elicitations 1–4)” [Figure 5B, Figure F1]

- (d) “The relationship between income and exhibiting non-welfarist preferences (behaviors in Elicitation 1–4). This heterogeneity is useful to explore because non-welfarist preferences may only be prevalent among the rich and/or those with high levels of education, in which case welfarist social welfare functions that aggregate up such preferences and place significant weight on “rights” would be regressive.” [Figure 6, Figure F2]

F1.2 Mechanisms Experiment

Primary Outcomes [Section 5]

1. “Whether the person chooses to rerun the lottery (binary).” [Appendix Figure A15A]
2. “The amount that the person pays to rerun the lottery (continuous).” [Figure 7]
3. Whether the person chooses to pay the maximum amount to rerun the lottery (binary). [Appendix Figure A15B]

Secondary Outcomes

1. “We focus on the following secondary heterogeneity, treatment effect on outcomes by: Political party, Education” [Figure F3]
2. “As a secondary test, we pool the tests in comparison 1–2 (that is, we compare arms 1 and 3 together, vs arms 2 and 4 together).” [Figure 7, Appendix Figure A15]

F.2 Additional Pre-Registered Outcomes

We now provide a brief discussion of additional outcomes.

F.2.1 Main Study

Correlations Between Anti-Targeting and Egalitarianism. We pre-registered presenting correlations between anti-targeting and all 3 features experiments (pairwise and joint). In our main analysis, we focus on correlations in just Experiments 1 and 2 due to the non-differential effect between rights goods and benchmarks goods in Experiment 3.

In Table F1, we supplement Table 3 with these tests. We find that non-welfarist preferences in Experiment 3, as well as our composite measure of non-welfarist preferences in Experiments 1-3, predict redistributive decisions in Experiment 4. A 1 s.d. increase in WTP for egalitarianism is associated with a 5 pp increase in allocating universally. Considering average WTP across the inalienability, choice, and egalitarianism, a 1 s.d. increase correlates with a 12 pp increase, which only slightly improves the relationship (11 pp increase) when only averaging inalienability and choice. Additionally, this relationship is stronger for lawyers than for health care.

Tagging Non-Welfarists Using Egalitarianism. We pre-registered tagging non-welfarist using all 3 features experiments, but again report these numbers using just Experiments 1 and 2. In Experiment 3 (Egalitarianism), we find high levels of non-welfarism that are not differential between rights and benchmark goods. Therefore, it is still interesting to tag non-welfarists using choices from this experiment. We adapt the analysis in Figure A6 in Figure F1. We present this analysis in the interest of completeness, but note that having positive WTP does not imply rights preferences in the egalitarianism experiment in the same way it does for inalienability and choice. This is because there is an instrumental benefit to providing the additional lawyer. Only 4% of this sample does not have positive WTP in the egalitarianism experiment; as a result, we find 99% have at least one of the three rights preferences and the remaining results are similar to the original analysis.

Secondary Heterogeneity in Anti-Targeting and Egalitarianism. We present above secondary heterogeneity on several dimensions (e.g. political and policy preferences, demographic characteristics) using participants' choices in Experiments 1-2. Now, we present such heterogeneity for Experiments 1-4 (Figure F2). Few heterogeneities are significant or consistent across experiments.

F.2.2 Mechanisms Experiment

Secondary Heterogeneity. We pre-registered presenting heterogeneity in the treatment effect by political party and education for this experiment (Figure F3). We find that neither of these

splits produce statistically significant differences. Directionally, those who report being liberal or having a Bachelor's degree or more have higher WTP when stakes are high.

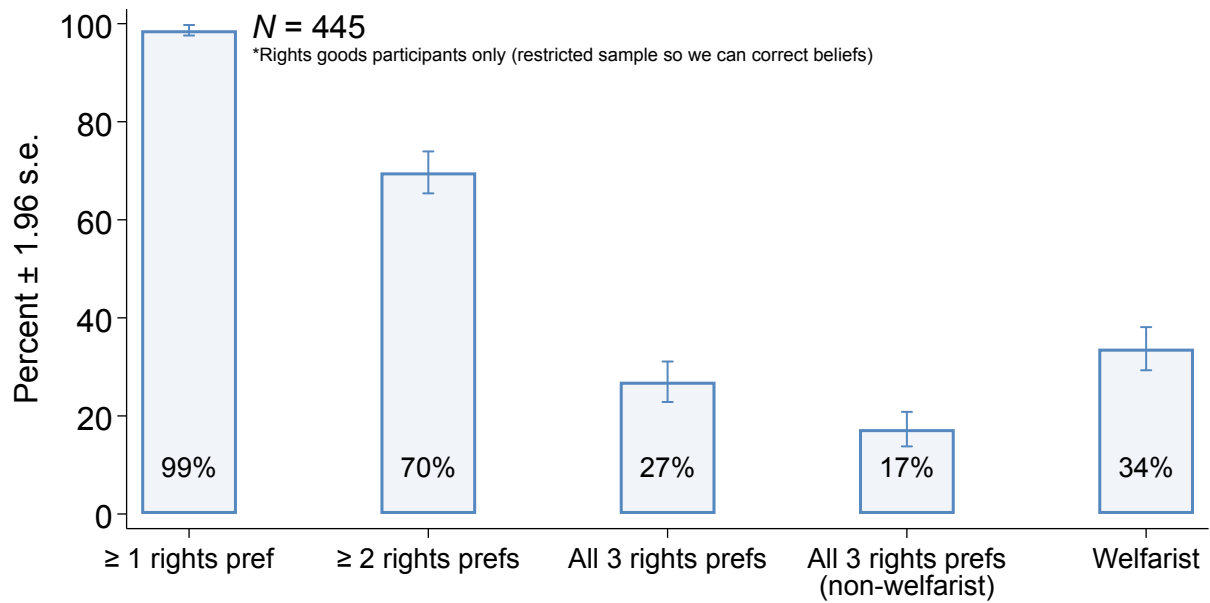
E.3 Tables and Figures for Additional Pre-Registered Outcomes

Table F1: Tests of Correlations of Egalitarianism with Anti-Targeting

| | Dep. Var.: Universal (= 1) | |
|--------------------------------|-----------------------------|-----------------------------|
| | (1) | (2) |
| Panel A. Rights goods (pooled) | | |
| WTP for egalitarianism (s.d.) | 0.048 (0.019) [0.012] | |
| Mean WTP (s.d.) | | 0.120 (0.027) [0.000] |
| Panel B. Lawyers | | |
| WTP for egalitarianism (s.d.) | 0.031 (0.017) [0.067] | |
| Mean WTP (s.d.) | | 0.129 (0.035) [0.000] |
| Panel C. Health care | | |
| WTP for egalitarianism (s.d.) | 0.021 (0.018) [0.238] | |
| Mean WTP (s.d.) | | 0.089 (0.040) [0.026] |
| Posterior ≥ 0.9 | | ✓ |
| Mean | 0.428 | 0.437 |
| Observations | 1,201 | 600 |

Note: This table extends Table 3 to analyze WTP for egalitarianism. It shows the effect of WTP for egalitarianism and mean WTP in Experiments 1-3. All WTPs are in terms of s.d. relative to the benchmark goods. The outcome is an indicator for whether the participant distributed the good universally. The sample is those who did the experiment with rights goods: Panel A shows pooled results, while Panels B and C show lawyers and health care, respectively. Columns (2), which includes measures of WTP from Experiment 2, further restricts to participants with posteriors ≥ 90 . Parentheses show robust standard errors. Brackets show p -values.

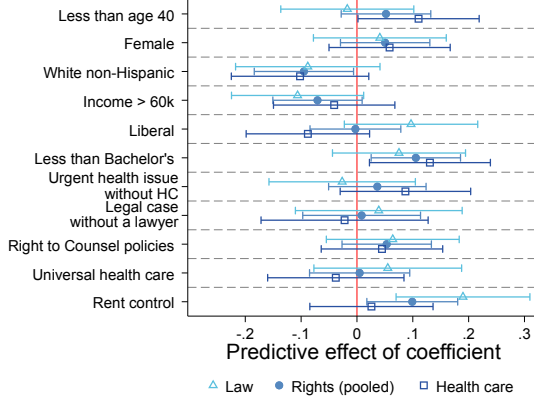
Figure F1: Prevalence, Including Egalitarianism



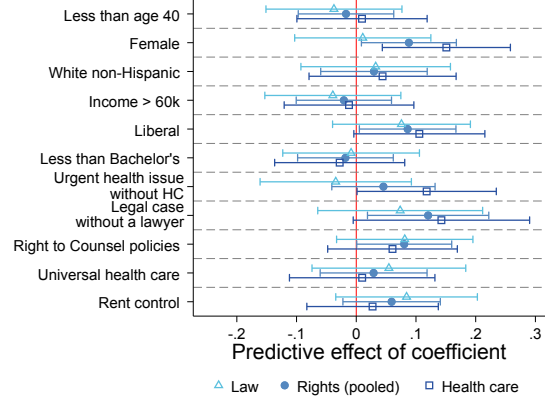
Note: This table extends Figure A6 to analyze prevalence of non-welfarist preference including egalitarianism (Experiment 3). We modify that analysis considering three potential rights preferences: positive WTP in inalienability, choice, and egalitarianism.

Figure F2: Heterogeneity in Experiments 1-4

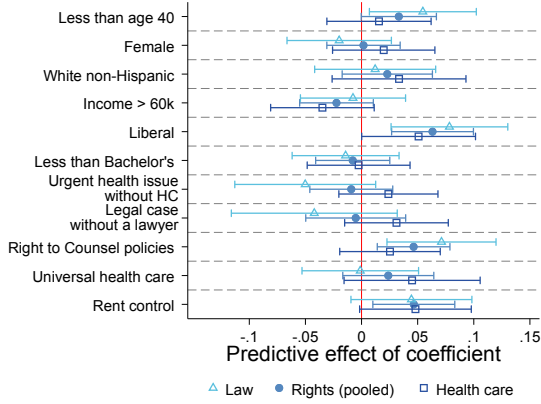
(a) Positive WTP for inalienability



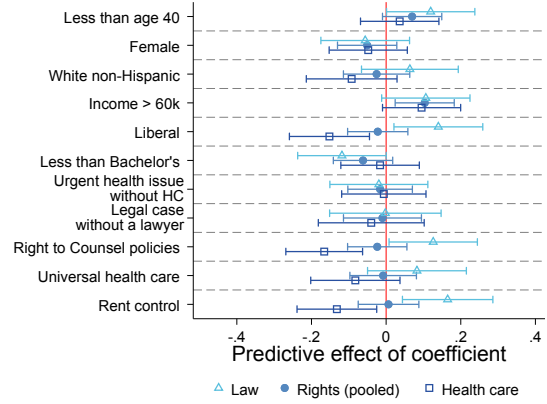
(b) Positive WTP for choice



(c) Positive WTP for egalitarianism

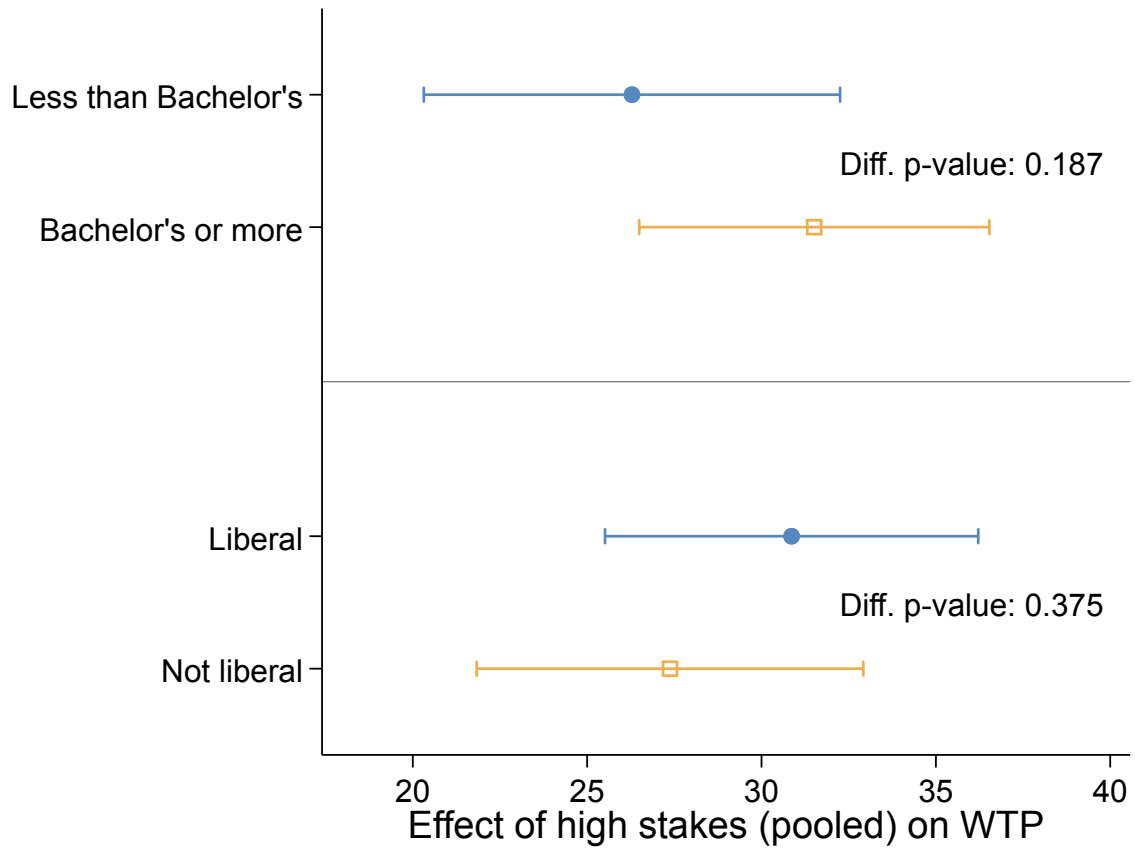


(d) Universal Allocation



Note: This figure shows the effect of different demographic characteristics and policy preferences on our main binary outcomes in each experiment. We present these effects on positive WTP for inalienability (panel A), positive WTP for choice (panel B), positive WTP for egalitarianism (panel C), and anti-targeting (panel D). All figures are restricted to participants who saw rights goods. Estimates are shown with with ± 1.96 standard errors.

Figure F3: Heterogeneity in Mechanisms Experiment



Note: This figure presents the effect of being in a high stakes arm on WTP to preserve the lottery in the mechanisms experiment for pre-registered subgroups. Participants were asked to report their level of education and political standing on a 5-point scale. Estimates are shown with with ± 1.96 standard errors.