Reference points and demand for redistribution: Experimental evidence

Jimmy Charité, Raymond Fisman, and Ilyana Kuziemko*

August 15, 2014

Abstract

If individuals—even the wealthy—are loss-averse relative to their reference point, then redistribution could reduce welfare even in the absence of moral hazard. We test whether this consideration affects the redistributive decisions of individuals acting as social planners. Subjects redistribute exogenous, unequal endowments between two strangers significantly less when the strangers *know* the initial endowments than when they do *not* (when in fact we observe near complete redistribution). Subjects who are themselves more loss-averse drive this effect. In a separate experiment, respondents choose a tax rate for someone who (due to luck) became rich five years (one year) ago. Respondents reward the more deeply embedded reference point in the five-year scenario with a lower tax rate. Our results offer a new explanation for why voters prefer lower levels of redistribution than standard models predict.

JEL Classification Numbers: C9, D63, H21, H23.
Key words: Redistributive preferences, optimal taxation, reference dependence

*We thank Angus Deaton, Larry Katz, David Moss, Howard Rudnick, Stefanie Stantcheva and Matt Weinzierl for helpful discussions, as well as seminar participants at Bocconi, Chicago, Columbia, Harvard, LSE, Northwestern, Princeton, and Stockholm. Financial support from the Tobin Project is gratefully acknowledged. Adith Srinivasamurthy provided invaluable assistance with JavaScript programming. Charité: Columbia University (email: jc4144@columbia.edu); Fisman: Columbia University (email: rf250@columbia.edu); Kuziemko: Princeton University (email: (kuziemko@princeton.edu).

1 Introduction

In a standard welfarist model of taxation with no moral hazard, a utilitarian social planner will aspire to equalize wealth across all members of society (Mirrlees, 1971). Moving from the normative to the positive, a democratic political process would similarly be expected to generate a high degree of redistribution: given a right-skewed ex-ante income distribution, the majority of voters have an incentive to expropriate the rich (Meltzer and Richard, 1981). However, many scholars have observed that neither survey-based expressions of policy preferences nor actual policies consistently reflect the more egalitarian predictions of standard models (see, e.g., Roemer, 1998 and citations therein). In the standard optimal tax model, moral hazard is the primary limiting factor, though standard estimates of labor-supply elasticity appear to predict much higher top income tax rates than are observed in most modern developed economies.[1]

Perhaps the clearest challenge to the standard model is individuals' reluctance to fully redistribute exogenous, initial wealth, a contradiction to utilitarianism recently studied by Saez and Stantcheva (2013) and Weinzierl (2012). Scholars have proposed a number of explanations for the limited demand for redistribution. Past work has focused on the prospect of upward mobility (Benabou and Ok, 2001), the effects of "policy-bundling" redistribution with other, cross-cutting issues (typically race in the U.S. context, see Lee and Roemer, 2006), and the public's misinformation about income inequality (see Ariely and Norton, 2011 on the level of misinformation, though see also Kuziemko *et al.* (2013) on the limited effect of information on policy preferences).[2]

This paper proposes an additional, heretofore unexplored explanation for the limited demand for redistribution, based on reference-dependent preferences. Whereas standard utility functions assume that only absolute levels of consumption enter into utility, decades of research support the view that in many important contexts individuals behave as though they evaluate options relative to a reference point. In particular, individuals are loss-averse relative to this reference point, in that the reduction in utility from a (relative) loss is greater than the increase in utility from a corresponding gain. As such, axing the rich may have large negative utility consequences that may not be offset by the utility benefits of increasing incomes of the poor. To the extent that voters recognize the reference-dependence of

[1] See Diamond and Saez (2011), who argue that a utilitarian social welfare function would yield top marginal tax rates over 70 percent given consensus estimates of labor supply elasticity.

[2] There is a related literature (see Harms and Zink, 2003 for a nice review) that examines why policy *outcomes* may be more regressive than voters' preferences even in a democracy (e.g., due to efforts by lobbying groups), but here we focus on the question of why voters' *preferences* might be more regressive than the predictions of standard models suggest they would be.

others (and incorporate others' welfare into their own utility), they may demand limited redistribution from rich to poor, even with exogenous, unearned initial endowments.

We explore the consequences of loss aversion on the demand for redistribution in an experimental setting where endowments are random and exogenous, thus removing concerns of deservedness or efficiency, two key factors that could limit redistribution. Our subjects (referred to for convenience in the paper as "redistributors," though such language is never used in the experiments) were recruited via Amazon's Mechanical Turk (mTurk), a rapidly growing online labor market, which we describe in more detail later in the paper. The experiment confronts subjects with a redistributive decision involving two other (randomly selected and anonymous) mTurk participants, who received (based on a coin flip) initial endowments, with one assigned to receive 5 dollars and the other 15. For most redistributors, the decision was presented as a hypothetical; for a subset we confirmed that the results hold when subjects are informed that with 10 percent probability their decisions will be implemented for real stakes. To vary redistributors' beliefs about the recipients' reference points, in the treatment arm redistributors were told that the recipients had already been informed of their initial allocation; in the control arm subjects were told that the recipients had *not* been informed of their initial endowments. Each redistributor was presented with both treatment and control conditions, with the ordering chosen at random. For most of our analysis we use between-subject estimates of the loss aversion treatment based on subjects' responses to the first condition they are presented with.

Our main finding is that subjects in the reference-point treatment (where subjects are told that recipients have been informed of their endowments) are significantly less likely to reduce inequality between recipients than subjects in the control condition. In our preferred specification, control group redistributors erase 94 percent of the initial 10 dollar difference between the recipients' endowments (indicating that a pure respect for "property rights" is limited), while only 76 percent of endowment differences are removed through reallocation in the treatment group. Our estimates hold across a large number of robustness checks—dropping those who finish the survey in a suspiciously short amount of time, dropping those who choose to increase inequality, changing presentational aspects of the experiment, and moving from hypothetical to "real stakes" scenarios—and remained stable as we collected data over several months.

To gauge the magnitude of these effects, in one session we asked respondents to redistribute unequal endowments between strangers whose endowments were assigned either by a coin toss, or earned by correctly answering Scholastic Assessment Test (SAT) questions, the main college admissions exam in the US. In both cases, the strangers would only know their final endowments, so the coin-toss scenario is identical to the control arm of the main exper-

iment. Past work has shown that redistributive decisions are strongly affected by whether income is earned by merit or by chance, and we find that to be the case in our setting: in the SAT treatment, redistributors close only 56% of the ex-ante gap. Our effect is thus just over half of the merit-versus-luck effect. We do not view our reference-point mechanism aand other explanations for limited redistribution as mutually exclusive and indeed take care to calibrate the size of our effect relative to what we ex ante believe to be among the most important reasons people do not seek full redistribution. The fact that our hypothesized mechanism appears to have an effect size that is over half that of the luck-versus-merit effect suggests it could be an important and heretofore overlooked explanation for limited redistribution.

To further connect our main results to the concept of reference points, we next investigate the extent to which the treatment effect is correlated with redistributors' *own* measured loss aversion. We use several distinct measures to capture loss aversion; our main approach is to present subjects with a sequence of hypothetical choices between a sure loss of \$5 dollars and a 50-50 gamble over losses of \$$L$ and \$0, where L takes on values of 8, 11, and 14 (Kahneman and Tversky, 1979). Consistent with earlier research, we observe that many subjects are willing to take actuarially unfair gambles in the loss domain, with 39 and 26 percent of subjects preferring the gamble with $L = 11$ and 14 respectively to a sure loss of \$5. Consistent with past work, subjects are far more risk-averse in the gain domain (for example, 73% choose a sure gain of \$5 to a 50-50 gamble over \$0 and \$11). We then test whether subjects' responsiveness to the treatment is correlated with the within-subject measure of loss aversion by including measures of loss-aversion and their interactions with the treatment indicator. We find that the treatment effect is driven almost entirely by loss-avers subjects: the interaction term between our loss-aversion measure and treatment status is highly significant, and its inclusion drives the main treatment effect close to zero.

Finally, we complement our lab-based results with survey evidence on actual redistributive policies, in particular income taxation of high-income individuals. Subjects were randomly assigned to one of two vignettes describing a person whose income had increased to \$250,000 owing to circumstances unrelated to skill or effort. The only difference between the treatment and control conditions was the length of time that the protagonist had been earning the higher income—in the treatment condition, he had been earning \$250,000 for five years and in the control condition for only one year. If respect for the reference points of the rich really deter voters from demanding more redistribution, then the effect should be stronger in situations where that reference point has become more deeply embedded (i.e., for the individual who had already been earning \$250,000 for five years).

Indeed, respondents choose a tax rate in the five-year scenario that is roughly 1.7 percent-

ages points lower than in the one-year scenario (for which they choose a rate of 28.3 percent). This 1.7 percentage point difference is over half the size of the gap between tax rates chosen by Obama and Romney supporters, and thus again appears of empirical importance.

Our work joins a large body of economics papers that attempts to incorporate insights from social psychology. While earlier researchers have fruitfully modeled the implications of "behavioral agents" for the taxation of *goods*, to our knowledge, little work has adopted the optimal *income* tax framework to consider the consequences of non-standard preferences like loss-aversion on the demand for redistribution.[3] This is surprising for at least two reasons. First, social preferences have played a central role in the field of behavioral economics and, in turn, social preferences are a central determinant of demand for redistribution. Second, at least in the American context, income taxes account for a far greater fraction of revenue than do taxes on specific goods, and thus better understanding voters' preferences in this area has important, practical policy implications.[4]

Second, we contribute to a relatively recent literature on reconciling differences between individuals' stated policy preferences and the prescriptions from standard utilitarian optimal tax models. Both Saez and Stantcheva (2013) and Weinzierl (2012) point to individuals' reluctance to fully redistribute income *even in the absence of moral hazard* as a serious challenge to the utilitarian framework. We offer some evidence that the full-redistribution result can be largely rehabilitated within standard utilitarianism, albeit with non-standard individual utility functions. Both these papers note other important departures from utilitarianism (e.g., limited use of tagging) that we do not address in this paper, and we certainly do not claim that these departures can also be reconciled with utilitarianism. (Saez and Stantcheva (2013) further discuss the possibility of a *social welfare function* that weights losers more heavily than winners. Such a SWF will, in a manner similar to loss aversion, limit transfers from the wealthy.)

Finally, to our knowledge, our study is one of the few to document choices that reflect the presumption of non-standard preferences *on behalf of others*. In this sense, our framework is related to models of guilt aversion (Battigalli and Dufwenberg, 2007), where individuals suffer disutility from a failure to meet the expectations of others, and as a result incorporate

[3]See Bernheim and Rangel (2004) on optimal taxation when agents can become addicted to a good, O'Donoghue and Rabin (2006) on optimal "sin taxes" when agents lack self-control, and Allcott *et al.* (2014) on optimal energy taxes and subsidies when agents are inattentive. The work on optimal income tax is theoretical, not experimental, and focuses on *relative utility*—when utility is a function of consumptions levels of peers—and not on loss-aversion. See Boskin and Sheshinski (1978) and Oswald (1983).

[4]Alesina and Passarelli (2014) analyzes how loss aversion affects policy formulation in general. Their finding of a status quo bias is driven by the strength of preferences by losers versus winners, rather than voters' regard for the losses of others.

others' utility into their own choices (see Battigalli *et al.* (2013) and Ellingsen *et al.* (2010) for experimental tests of guilt aversion). Our results are novel in their application to important questions in public finance and their contribution to our understanding of redistributive preferences and policies. It is also interesting to note that our findings stand in contrast to earlier research by Marshall *et al.* (1986), which looks at individuals' decisions when acting as *advisors*; in this setting, subjects do not exhibit loss aversion. Our findings indicate that a fruitful direction for future research may be examining the extent to which preference anomalies exist in decisions with direct consequences for others' (rather than one's own) payoffs.

The paper proceeds as follows. Section 2 presents a simple framework for thinking about optimal tax decisions with reference-dependent individuals. Section 3 describes the main lab experiment and Section 4 our mTurk sample. Section 5 reports the main results from the redistribution experiment. Section 6 introduces the tax-policy survey experiment and reports the results. Section 7 offers directions for future work and some concluding thoughts.

2 Reference points and optimal taxation

We begin by briefly examining how reference-dependent utility affects the prescriptions of a standard model of optimal taxation, focusing primarily on a single-period static case. We close with a discussion of how dynamics and possible adaptation would change the analysis.

2.1 Redistribution in a static optimal tax model

2.1.1 With standard utility assumptions

Even in the absence of moral hazard the standard optimal tax exercise would be intractable without assumptions on individual utility. As Mankiw *et al.* (2009) point out, despite the general caveat that utility cannot be compared across individuals, the typical optimal tax exercise does exactly that and further assumes that individuals have identical utility functions: a function $f : x \to u$ maps consumption x (which in the static model is identical to income or wealth) into utility u, with $f' > 0$ and $f'' < 0$.

With these assumptions, optimal taxation in the utilitarian framework with no moral hazard is a simple exercise. Assuming a single convex function can capture all agents' utility, then welfare is maximized at the point of total equality.

Figure 1 displays this logic for two individuals p (poor) and r (rich), who ex-ante are endowed with unequal wealth x_p and x_r, $x_p < x_r$. Given r's lower marginal utility of wealth, a social planner can increase total utility by transferring some Δ from r to p. In fact, the

social planner can continue to do so and still increase total welfare up to the point where $f'(x_r) = f'(x_p)$, that is, $x_r = x_p$.

2.1.2 With reference-dependent utility

If utility is reference-dependent, we lose the analytical convenience of both a common utility function across individuals as well as differentiability. Figure 2 shows typical reference-dependent utility functions.[5] Importantly, r and p no longer share a utility function, as each has a point of non-differentiability at their status-quo position. Transferring Δx from r to p no longer guarantees welfare improvements. As shown in the figure, moving to complete equality is welfare-reducing, as is any smaller perturbation that transfers endowments from r to p.

While we have drawn gains as concave and losses as convex in Figure 2, such restrictions are not necessary for the local result. So long as losses loom larger than gains relative to the reference point, then small amounts of redistribution will reduce welfare. It is harder to make specific claims about much larger changes, which will depend on third derivatives. The main point is that reference dependence weakens the claim that, without moral hazard, redistribution is necessarily welfare-enhancing in a utilitarian framework.

2.2 Reference points in a dynamic setting

The analysis in the previous subsection assumes a static setting where an individual forms a single reference point from her ex-ante endowment and then experiences a one-time change in utility based on the ex-post distribution. In reality, the government sets tax and transfer policy continuously. While a decision to substantially redistribute income or wealth in a given year might well have the utility consequences described above in the year that follows, if individuals adapt to their new reference points then the formerly rich will experience only momentary disutility in subsequent years.

It is an open question how quickly people adapt to changes in income and thus how malleable reference points are over time.[6] Strictly speaking, our empirical work will sidestep this question: we merely test whether individuals appear to respect others' reference points when deciding whether to redistribute. Finding, as we do, that they in fact respect reference points implicitly implies that *individuals think* that reference points are not completely malleable, or that they think myopically when making redistributive decisions. Thus, the

[5] We draw the shape of the utility functions in Figure 2 to roughly approximate those calibrated in Abdellaoui *et al.* (2007).

[6] See, e.g., Di Tella *et al.* (2010) and citations therein. They find that individual happiness measures return to baseline roughly four years after an income shock.

actual adaptation of reference points is not relevant to the positive question of whether they affect how individuals choose to redistribute.

However, while not our goal, we imagine that adaptation might be central to drawing normative implications from this positive result. If adaptation is slow and losses loom larger than gains, then voters who want to limit redistribution are plausibly maximizing welfare in the utilitarian model. While past work (Chetty and Szeidl, 2007) has shown that reference points (in their case, built on the micro-foundation of consumption commitments) imply higher levels of social insurance against adverse events (as losses have greater utility cost with reference points than with standard utility), by the same logic reference points might also suggest lower levels of redistribution for the purposes of condensing the current income or wage distribution.

On the other hand, if adaptation is rapid, such a voter might not be maximizing welfare, or at least not in a longer-run, steady-state sense. If voters or policy-makers overestimate the persistence of reference points, then their chosen level of redistribution may be lower than the optimal, welfare-maximizing point.

3 Experimental design

We recruited and compensated our subjects through Amazon's Mechanical Turk (mTurk) marketplace (which we describe in detail in Section 4), but redirect them to surveys that we built with Qualtrics' online survey software, adding functionality with JavaScript as needed.

We collected data from nine distinct sessions. In six of these, respondents proceeded through modules of the survey in the following order: (1) the main redistribution experiment; (2) questions on loss-aversion; (3) background questions on political beliefs and demographics. We describe each below. In three of the later surveys, subjects were presented first with questions on income taxation, which we describe in detail in Section 6. (See Appendix Table 1 for the dates of the surveys and the attributes of each.)[7]

3.1 Main redistribution experiment

The centerpiece of the survey presents respondents (whom we term "redistributors" in the paper, though, as noted earlier, at no time do we use this term in the survey itself) with the opportunity, in most cases hypothetical, to transfer money between two other anonymous

[7]We had technical problems in one session and thus do not include it in the main analysis. A description of the problem as well as results from that session appear on the last page of the Appendix.

mTurk participants. In all cases, *the redistributor received only his show-up fee regardless of his decision, so has no direct self-interested motivation.*

Respondents randomized into the control arm of the survey encountered the following instructions:

> Consider two other participants on mTurk, person A and person B. Based on a coin flip, we have given $5 to person A and $15 to person B.
>
> You can now transfer money between persons A and B. *Persons A and B are not told how much money they were initially given*. If you decide to give Person A $X instead of $5, he or she will simply be told that they have been given $X, and will not know how much they started with. Nor will they know that there is another person (Person B) involved, or that a third party (you) determined the money they received.
>
> Please indicate below what transfer, if any, you would make.

A slider and interactive bar graph (which reflects in real time movements of the slider) appeared directly below these instructions, allowing respondents to easily and transparently transfer money between players. The default position of the slider was on the ex-ante ($5, $15) distribution. Appendix Figure 1 provides a screenshot.

For those randomized into the treatment arm, the second paragraph of the control instructions was modified as follows:

> You can now transfer money between persons A and B. *Persons A and B have already been told how much money we have given them*. If you decide to give Person A $X instead of $5, they will be told that they now have $X instead of $5. They will not know that there is another person (Person A) involved, or that a third party (you) determined the money they received.

Appendix Figure 2 provides a screenshot.

To gather within-subject comparisons, we also performed the "reverse experiment" and so immediately after answering the treatment (control) version of the question, treatment (control) respondents answer the control (treatment) version of the question (with the labels "Persons A and B" replaced with "Persons C and D").

Finally, to ensure the robustness of our findings, we ran two variants on the above experiment. First, for one group of subjects, none of the text was italicized or underlined, and the underlined reminder message (see screenshots in the Appendix) placed next to the slider

was removed. In a second group of subjects, the subjects were informed, prior to seeing the instructions, that there was a 10 percent chance their decision would be implemented for real stakes:

> The next two questions will give you the opportunity to determine the payments to two other mTurk participants.
>
> After you make your decisions, the computer will pick at random whether or not to implement your decision. There is a 10% percent chance that one of your decisions *will actually be implemented*. Because you do not know ahead of time whether your decision will be chosen, *you should make your decisions as if they were for real money*.

Two presentational aspects of the main experiment deserve mention. First, the use of the slider requires a default position, which we set to the status-quo allocation of \$5 and \$15 dollars. As such, we suspect that anchoring bias could lower the amount of redistribution. This bias should affect both the treatment and control groups and thus not bias the treatment effect. Second, to illustrate clearly the treatment scenario, note that we write: "If you decide to give Person A \$X instead of \$5, they will be told that they now have \$X instead of \$5." By using the poorer person as the illustration, if anything we should prime redistributors to think of the pleasant surprise that the person starting with \$5 will experience, again biasing the experiment against finding our hypothesized effect.

3.2 Questions to determine respondents' own sensitivity to reference points

As noted in the introduction, we conjecture that subjects who are themselves more sensitive to reference points will exhibit a larger treatment effect, as they project their own loss-aversion onto recipients' outcomes. Such a test requires an individual-level measure of loss-aversion. However, the vast majority of papers on loss-aversion use between-subject analysis—for example, the classic endowment-effect experiments are demonstrated by one group's willingness to accept being higher than another group's willingness to pay which, by construction, is a between-subject exercise. Similarly, in their work showing that respondents judge the fairness of market transactions based on reference points, Kahneman *et al.* (1986) use comparisons between groups assigned to read different vignettes.

Our approach is largely to take the questions asked in between-subject designs and present both versions to each subject. Our main measure of loss aversion is based on the following

sets of questions that relate directly to the central claim of loss-aversion, that individuals are risk-loving over losses and risk-averse over gains:

> Imagine that you face the following decisions. For each decision, please examine both options and indicate the one that you prefer.
>
> - Decision One:
>   - A sure gain of $5.
>   - 50% chance to gain $8 and 50% chance to gain nothing.
> - Decision Two:
>   - A sure gain of $5.
>   - 50% chance to gain $11 and 50% chance to gain nothing.
> - Decision Three:
>   - A sure gain of $5.
>   - 50% chance to gain $14 and 50% chance to gain nothing.

Subjects then faced the same gambles, but in the loss domain:

- Decision One:
  - A sure loss of $5.
  - 50% chance to lose $8 and 50% chance to lose nothing.
- Decision Two:
  - A sure loss of $5.
  - 50% chance to lose $11 and 50% chance to lose nothing.
- Decision Three:
  - A sure loss of $5.
  - 50% chance to lose $14 and 50% chance to lose nothing.

We supplement this approach with additional questions related to behaviors that prior researchers have suggested relate to loss-aversion, but are less directly related to its fundamental predictions. The first is motivated by the empirical results documented in Genesove and Mayer (2001) showing that sellers of owner-occupied homes appear to eschew offers below the price at which they bought their own house to avoid realizing a loss:

> Suppose you bought a house for $250,000 a few years ago. The housing market in your neighborhood has since declined, and you have seen houses very similar to yours sell for $200,000, though some sell for a bit more and some sell for a bit less. You expect the current housing market conditions in your neighborhood to remain relatively stable. You are planning to relocate in the coming year for a new job. Someone is interested in buying your house. What is the least you would be willing to accept as a sale price?

Subjects were then confronted with a question that was identical, except that the house's purchase price was changed from $250,000 to $300,000.

Second, we take a question from Kahneman *et al.* (1986) that relates to reference-dependent preferences over an employee's wages:

> A small coffee shop has one employee who has worked there for six months and earns $10 per hour. The shop continues to do fairly good business, though unemployment in the area has increased due to a factory closure nearby. As a result, other small restaurants have now hired reliable workers at $8 an hour to perform jobs similar to those done by the coffee shop employee. The owner of the coffee shop reduces the employee's wage to $8. The owner's actions were: ∘ Completely fair ∘ Acceptable ∘ Unfair ∘ Very unfair.

Subjects were then presented with a version of the question where penultimate sentence is changed to: "The current employee leaves, and the owner decides to pay a replacement worker $8 an hour."

Appendix Table 2 shows the correlation among these measures. Intriguingly, the answers to the vignettes appear largely uncorrelated with the classic measure of being risk-loving over gains, suggesting that they are capturing distinct phenomena. In general, we focus on the classic risk-loving-over-losses measure, but show results with all loss-aversion measures.

3.3 Demographic and political opinion questions

The survey ended with standard demographic questions, along with questions on political affiliation and whether "the government should reduce income differences between the rich and poor." These questions allow us to examine whether our treatment effect is larger for certain groups, and also to compare the mTurk sample to more representative populations such as the American Community Survey. Given that we collect these questions *after* the experiment (so as not to prime the results in the experiment), it is possible that some answers may be primed by the experiment itself; comparisons by these covariates should thus be viewed with this potential priming in mind. The final questions of the survey relate to whether respondents felt any part of the survey was confusing or biased and also asked for any other feedback they wished to share.

We now describe the data-collection process on mTurk and the resulting sample of respondents.

4 Data

4.1 Data collection procedures

All of our subjects were recruited through mTurk. mTurk is an online labor market where "requesters" can post *human intelligence tasks* (HITs) to be completed by "workers." As of this writing, mTurk advertises that requesters can "access more than 500,000 workers." The most common posted HIT currently is "extract purchased items from a shopping receipt" and pays 8 cents (the requester would pay in addition roughly ten percent to Amazon).[8]

Over the past few years, social scientists have increasingly used mTurk to perform experiments and collect survey data (see Kuziemko *et al.*, 2013 and papers cited therein for a review). We registered as a "requester" and posted the following HIT: "Short (less than ten minutes) opinion survey on a variety of topics." We tried to use a neutral description that would limit selection bias while also giving workers an honest description of the work. Compensation was set to $1, which approximated minimum wage assuming that subjects took seven or eight minutes to complete it. Actual median completion time was 10.1 minutes, implying an hourly wage of $6.09. Though we cannot find official data on average wages on mTurk, reading through worker forums suggest that we are paying a very generous wage (and indeed when we post a request for 300 survey takers, the full sample is typically gathered within an hour).

Each worker logs in with an mTurk worker ID. Recall that we collect data over nine separate sessions, and thus we drop any worker who has taken a previous survey with the same ID so as to gather a fresh sample each time. Of course, if workers have multiple worker IDs then some individuals may have participated in a previous session. Outside of surveys (which appear to make up a very small share of all HITs), in which case requesters would typically want unique workers, there is little financial incentive for mTurk workers to create multiple mTurk IDs, but we cannot completely eliminate the possibility that some have and thus could pass through our screening process.

Another issue that arises on mTurk is the possibility of 'robots,' algorithms who masquerade as individuals. To address this concern, we begin each survey with a "captcha" (non-standard writing difficult for computers to interpret).[9] To the extent some remain in our sample, they would attenuate any treatment effect.

To limit heterogeneity of the sample, we collect all data on workdays during daylight hours on the East Cost of the United States. Individuals were automatically prompted for a

[8]Based on viewing a list of HITS on 10:56 AM EDT, August 12, 2014.

[9]Examples of "captchas" can be found here: `http://www.fileflash.com/graphics/screens/Captcha_Creator_PHP_Script-69.gif`.

response when they tried to skip questions (to discourage robots or inattentive respondents). Particularly given our focus in some parts of the survey on American tax policy, we limited the survey's availability to those with U.S. billing addresses; we further asked respondents to confirm their residency in the United States. To further ensure the attentiveness of our subjects, we limit respondents to those with positive ratings from at least 90 percent of past "requesters."

The data pass basic reality checks (for example, subjects that identify as Republican tend to be white and male, mirroring patterns observed in polling data). Over three quarters of respondents went on to answer an open-ended "feedback" question, with the vast majority providing positive feedback on the survey.[10]

In Appendix Tables 3 we show how questions on perceived political bias of the survey vary with treatment status. About 87 percent of respondents felt that the survey was unbiased, with about 9 percent finding it had a liberal slant and 3 percent a conservative slant. There is a borderline significant difference in perceived bias among treatment and control subjects (column 3), though the bias is split between those that say the experiment was left- and right-wing biased (columns 1 and 2), and neither directional bias is significant. Similarly, survey fatigue should not affect our estimates of the treatment effect, as the average number of minutes taken to complete the survey is also independent of treatment status.

4.2 Data sample

Table 1 provides detail on the mTurk workers who completed our survey, comparing them to the (weighted) population of adults sampled in the 2010 American Community Survey. Consistent with past work using mTurk, we find that younger, male, and college-educated subjects are over-represented in our sample, while minorities are under-represented.

Table 2 provides a longer list of covariates, while Table 3 examines differences between the control and treatment groups. Overall, there appears to be good experimental balance, with no variable showing a statistically significant difference at the five-percent level. In particular, a variable that would be expected to have an impact on redistributive decisions—an indicator for supporting President Obama in the 2012 election—is essentially identical between the treatment and controls groups (p values of 0.89).

Our loss-aversion measures also appear relatively balanced across survey arms. *R-loving in losses* is coded as 0 if the respondent rejects all three gambles in the loss domain, 1 if she accepts only the $8 gamble, 2 if both the $8 and $11, and 3 if she accepts all three gambles.

[10]The positive feedback likely reflects the tedium of most other mTurk tasks. The vast majority of negative feedback concerned the difficulty of some of the "captchas," suggesting that it would be difficult for robots to pass through this screen.

R-loving in gains is defined in a parallel manner. Our respondents are more risk-loving in losses than gains, but there are no differences appear across survey arms. *L-G risk* subtracts *R-loving in gains* from *R-loving in losses* as an additional loss-aversion measure: in this case, how much *more* risk-loving a respondent is in losses than in gains.

The last two variables refer to our loss-aversion vignettes. *Higher WTA* indicates that the respondent demanded a higher house price with a \$300,000 initial price than the \$250,000 price. 80% of our respondents felt that the initial price was relevant to their decision. Finally, Δ *wage unfairness* takes the 1-4 rating of unfairness that the employer would cut the current employee's wage to \$8, and subtracts the same rating for the new employee. It is thus a measure of how much *more* unfair the cut is to the current employee and, indeed, our respondents view it as substantially more unfair. Thus, for both measures the reference point effect survives a within-subject comparison, despite subjects seeing the scenarios one after the other.

5 Results from the redistribution experiment

5.1 Main results

Table 4 shows, for the full sample, the main between-subject differences in total redistribution for those first assigned to the control versus those first assigned to the treatment. Recall that redistributing \$5 from the "richer" to "poorer" recipient would result in complete redistribution. Column (1) shows the treatment effect controlling only for session fixed effects. Those in the control group achieve nearly complete redistribution, shifting an average of \$4.55 from the richer recipient to the poorer one, or 91% of level of redistribution required for strict equality. Recall that the default position of the slider was the status quo (\$5 and \$15) allocations, suggesting that anchoring may, if anything, bias the control group results against inequality-reducing redistribution. Those assigned to the treatment redistribute on average \$0.83 (or 17 percent) less than those in the control.

Column (2) drops subjects who finished the survey in less time than one could reasonably be capable of completing it.[11] The control group mean increases slightly, consistent with the view that some rapid finishers simply clicked thoughtlessly through the redistributive decisions, leaving the sliders in their default positions. The treatment effect increases when these subjects are dropped.

Column (3) continues to exclude extremely quick finishers, and also drops subjects for

[11]Specifically, less than three minutes for the first session (as it did not have the module on loss-aversion), and six minutes for the other sessions.

whom the \$5/\$15 experiment was not the first item of the survey. This restriction removes subjects that may be contaminated by exposure to our income tax survey experiment (which we discuss in Section 6 below). We take the results in this column as our preferred specification: the control group closes 94% of the gap ($4.692 \div 5$), whereas the treatment group only 76 ($(4.692 - 0.888) \div 5$). Thus, our treatment reduces redistribution by 19 percent.

The rest of the table provides a number of additional robustness checks on the results in column. to additional robustness checks. In column (4) we drop subjects who choose to make inequality-*increasing* reallocations—what we term "odd choices" in the table. We present results with these subjects excluded to ensure that our average treatment effect is not being driven by them (though one could in fact imagine such choices as utility-maximizing under, say, a convex utility function). Column (5) includes a number of demographics control variables; given the balance across the control and treatment arms documented in Table 3, it is not surprising that the treatment effect is unaffected by the inclusion of these controls. Column (6) presents results for the "real stakes" subsample of redistributors who were informed that there was a 10 percent chance that their decision would be implemented, while in Column (7), we show results for the group of subjects where none of the text in the experiment was highlighted or underlined. We find that our basic result continues to hold in each of these subsamples, though the effect is attenuated, particularly in the non-underlined version.

Our main results are consistent across sessions more generally, as shown in Appendix Figure 4. Here, for the Column (3) sample, we plot the coefficients and 95 percent confidence intervals for the treatment effect, disaggregated by session date. The only treatment effect that stands out is the very first survey, which is larger in magnitude than the others. While, as noted, we drop individuals who took previous sessions with the same mTurk ID, the large treatment effect for the first survey potentially suggests that individuals who already took the survey using a different worker ID may have attenuated the measured treatment effect in later sessions. The session-by-session results highlight how the "real stakes" and "unemphasized" versions of the survey have nearly identical treatment effects as the two standard surveys that followed the initial session.

Figure 3 shows histograms (from the Column (3) sample) of the final allocation for the ex-ante "poorer" player, both for the treatment and control groups. For both groups, the distribution is bimodal, with most of the mass at (10, 10) but also a second, shorter peak at (5,15). By contrast, there is almost no mass in between these two points. Thus, most of the treatment effect occurs at the *extensive* margin—the decision to redistribute at all—as opposed to increases in partial redistribution. If losses are "more convex" than gains are "concave," then after no redistribution, complete redistribution is the utility-maximizing

outcome. As such, the lack of partial redistribution is consistent with subjects responding to the convexity of losses.

Table 5 shows these extensive margin results in a regression framework. Approximately four-fifths of control group respondents set final allocations at (10, 10), as compared to only 60 percent of the treatment group, for a treatment effect of 25 percent along the extensive margin. As before, this result is highly robust to the addition of controls, as well as excluding suspiciously short surveys, and "odd" choices that increase inequality.

As noted in Section 3, all subjects are presented with the "reverse" experiment: those first assigned to the treatment also face the control scenario, and vice versa. Our emphasis is on the between-subjects analysis presented above, as respondents are likely to anchor at least partially on their first response, and past work has further shown that the tendency to anchor may be related to loss-aversion.[12] Nonetheless, the within-subject treatment effect is highly significant, and near-identical in magnitude to the between-subject estimates. Table 6 shows an average treatment effect of -0.873 (column 1), with a somewhat smaller effect (column 2) for those who start with the treatment scenario than those who start with the control scenario (column 3). The final column shows that the within-subject result holds also for the real-stakes subsample.

5.2 Magnitude of the reference point effect

To gauge the magnitude of the effect we document above, we compare our treatment effect to the impact of having endowments earned by merit (score on SAT questions) rather than luck (a coin flip). This comparison is motivated by prior work using polling data, which has shown that one of the most important determinants of support for redistribution is whether an individual considers income to be a function of merit versus luck (Alesina and Angeletos, 2005).

To make this comparison, we ran a session where the control arm was kept the same (the \$5 and \$15 endowments were determined by a coin flip, and the strangers will only know their final allocations), while in the treatment arm respondents were told "the initial amounts given to Persons A and B *were based on their performance on SAT questions* [emph. in original]." As with the control version of this experiment, the redistributor is told that Persons A and B would only learn of their own final allocation.

There results are reported in Table 7. Consistent with the large effect of perceived merit on redistribution found in past work, the coefficient on the treatment variable is larger than in our reference-point experiment. Dividing the coefficients in this table by their analogues

[12]See Beggs and Graddy (2009).

in Table 4 suggests that our reference-point effect is between 51 and 55 percent of the luck-versus-merit difference.[13]

Interestingly, the histogram of final outcomes takes a somewhat different shape than for the main experiment (see Appendix Figure 5). There are still mass points at \$5 and \$10, but other intermediate choices are now more popular. Whereas only nine percent of those assigned to the original treatment chose an "in between" final allocation for the reference-point experiment, twenty percent do so in the luck-versus-merit experiment, highlighting that respondents are more sensitive to the assigned reference points in our main loss aversion treatment.

5.3 Are treatment effects larger for the loss-averse?

We hypothesize that individuals make decisions in the original redistribution game by projecting their own loss-aversion onto other subjects. We therefore expect to see a stronger treatment effect among those who are themselves more loss-averse.

In Table 8, we show results from estimating regressions of the form

$$Redistribution_i = \beta_1 Loss\text{-}Averse_i \times Treated_i + \beta_2 Treated_i + \beta_3 Loss\text{-}Averse_i + \gamma X_i + \epsilon_i.$$

We conjecture that $\beta_1 < 0$,; that is, those who are loss-averse will redistribute even less than others when exposed to the treatment condition.

Table 8 shows support for the notion that the treatment is strongest among loss-averse subjects. In column (1) we replicate our preferred specification from the main experiment, for ease of comparison. In column (2) we add the interaction between treatment status and our measure of risk-seeking *over losses* (along with its main effect, though we do not report this coefficient in the interests of space). The coefficient on the interaction term is negative and highly significant and, moreover, reduces the main effect of the treatment essentially to zero.[14] Column (3) shows that risk attitudes *over gains* has no significant effect on the treatment. When we subtract risk attitude in gains from the risk attitude in losses (to compare risk-taking over losses *relative* to risk-taking over gains) we also find a significant

[13]Our luck-versus-merit results are similar to those found in preliminary work by Chevanne et al. In an undated online draft of their paper, they find that third-party dictators redistribute roughly 32 percent less when they are told initial inequality was due to effort. In our experiment, the effect is closer to fifty, depending on the specification.

[14]It is worth noting that this regression presents a demanding test: as Table 2 shows, a large majority of individuals give answers consistent with loss aversion in the surveys. Given that the main treatment effect acts primarily through the extensive margin, it is plausible that individuals with varying degrees of loss aversion over own payoffs will have similar treatment effects, so long as they are above some threshold level of loss-aversion.

interaction effect in the expected direction (col 4). Neither of our alternative loss aversion measures approaches statistical significance (and in fact both point estimates go against the hypothesized direction); though as we noted earlier both 'vignette'-based proxies are weakly correlated with our "classic" loss-aversion measure.

The strong effect of the loss-aversion interaction could merely reflect a differential treatment effect along another margin that correlates with loss-aversion. In Appendix Table 4 we show results for our main covariates interacted with the treatment effect. Of the eleven variables, only two (age and college) are even marginally significant. As such, differential treatment effects appear very small (outside of loss-aversion).

To further assess the robustness of our finding that own loss-aversion interacts significantly with the treatment effect, we examine whether the coefficient on the *Risk-loving in losses*$\times$ *Treatment* effect in column (2) remains large and significant when we simultaneously control for *all* the interactions documented in Table 8 (along with all underlying main effects). Column (7) shows that the interaction of interest is virtually unchanged when these additional controls are included; column (8) shows this to be the case as well when we use our measure of risk-loving over losses *relative* to gains: including all the $Covariate \times Treatment$ interactions does not affect the coefficient on the interaction term of interest.

While the results in Table 8 indicate that our main results are driven by loss-averse subjects, it may be that loss-averse individuals are more sensitive to treatment effects more generally. To test this claim, in Appendix Table 5 we replicate the analysis in Table 8, using data from the SAT/coin-toss version of the experiment. While the sample size is small and our estimates correspondingly noisier, the pattern of coefficients offers no hint that loss aversion affects the merit treatment effect. The coefficient on the treatment's main effect is virtually unchanged when its interaction with the loss-aversion measures are added, in sharp contrast to the patterns observed in Table 8.

5.4 Discussion

We find very robust evidence that in their role as social planner, subjects' decisions are affected by whether recipients are aware of their initial endowments. The high levels of redistribution we observe in the control condition suggest that subjects have little innate respect for property rights as the dominant principle in their redistributive decision-making, as in Nozick (1974). While the fact that the electorate does not vote for complete redistribution in the real world is almost certainly over-determined (and as such the ideas put forth in, for example, Saez and Stantcheva (2013), Weinzierl (2012) surely play an important role) at least part of the explanation—reference points—can be reconciled with classic utilitarianism,

albeit with non-classical utility functions. The results from the control arm of the experiment are thus interesting in their own right.

As noted earlier, Chetty and Szeidl (2007) present a model of consumption commitments that could similarly diminish redistribution by a social planner who takes into account the commitments of relatively well off individuals. As they in fact note, however, in a context such as ours it is implausible that *actual* consumption commitments could drive subjects' decisions—the individuals over whom they were making decisions were given money that, by construction, had not yet been spent.[15]

Our findings to this point indicate that subjects are sensitive to others' reference points in redistributive decisions in laboratory settings over small stakes. In order to relate our findings more directly to policy-relevant questions, we now turn to results from a survey experiment on preferences over income tax rates.

6 Survey results on reference points and preferred top income tax rates

The question of the what constitutes an appropriate income tax rate on high-income households is a much-discussed issue in American politics today. A threshold of $250,000 has become a focal point in this discussion, and surveys often ask about support for higher taxes on households with annual incomes of at least that level.[16] While the majority of respondents in surveys tend to support higher income taxes on this group, the strength of those preferences has been debated. For example, while survey respondents in 2010 exhibited strong support for letting the so-called "Bush tax cuts" (those specified in 2001 and 2003 tax relief acts) expire individuals earning over $250,000, in that year's midterm Congressional elections Republicans won handily despite their position in favor of extending the cuts, despite this being a major issue in many campaigns.[17]

The survey experiment below tests whether respect for reference points might weaken respondents' preference to tax high-income households.

[15]If individuals develop reference-dependent heuristics as a result of commonly observing consumption commitments in their day-to-day lives, it could help to provide an underlying model for reference dependent preferences. Examining this possibility may be an interesting direction to pursue but is outside the scope of our paper.

[16]See, e.g., `http://politicalticker.blogs.cnn.com/2012/12/06/trio-of-polls-support-for-raising-taxes-on-wealthy/`.

[17]Larry Bartels discussed this tension in a 2010 online post: `http://today.yougov.com/news/2010/10/26/taxes-energized-minority/`.

6.1 The survey experiment

Subjects were presented with a vignette describing an individual that had received an unexpected increase in earnings. In most waves, the source of the increase was a corporate takeover of the company where the individual is employed (the "takeover" vignette). Subjects were randomly assigned to either a treatment or control arm, which differed only in the *timing* of when the earnings increase took place.

The "control" arm of the vignette took the following form:[18]

> There has been much talk about whether wealthy families are paying their fair share in taxes.
>
> Consider the following person. He has been working for about five years as a regional sales manager at a medium-sized firm. *This year*, his firm was taken over by a larger corporation. While he will be doing the same job as before, to make his pay compatible with the earnings of employees in his position at the larger firm, his salary is now doubled, to $250,000.
>
> If it were up to you, how much of his salary should he pay in taxes? (As a basis of comparison, the average American pays about 22 percent in taxes on the income they make.)

In the treatment variant, we attempt to make the protagonists' reference income of $250,000 more deeply embedded. Instead of receiving the raise just this year, he received it five years ago. Specifically, the second paragraph in the treatment vignette reads:

> Consider the following person. He started five years ago as a regional sales manager at a medium-sized firm. *Soon after starting*, his firm was taken over by a larger corporation. While he did the same job as before, his salary was doubled to make his pay compatible with the earnings of employees in his position at the larger firm. Since then, his annual salary has been roughly steady and is now $250,000.

After reading either the control of treatment version of the vignette, subjects provided their response using a slider positioned immediately below the vignette, with values in the

[18] In the vignettes we reference a tax rate of 22 percent on the "average American." We base this figure on NBER Taxsim estimates for combined federal and state income tax, and then add the employee side of payroll taxes.

range [0,100] percent and the default set to zero. (See Appendix Figures 3 and 4 for screen shots.)

In later sessions, to assess the generalizability of our findings, we changed the reason for the individual's increase in income. In the control version of this final wave, the second paragraph of the vignette above is replaced with:

> Consider the following person. *This year*, he won the state lottery. As a result, he will receive $250,000 a year for the rest of his life (note that lottery winnings are treated as taxable income).

As before, in the treatment version, to strengthen the reference point, we simply replaced "*This year*" with "*Five years ago*" and changed the verb tense (from "will receive" to "receives") as appropriate.

Our analysis in this section is motivated by models of habit formation whereby individuals acclimate to conditions—financial or otherwise—over time (see, for example, Bowman *et al.* (1999)). Thus, we conjecture that subjects presented with a vignette where the protagonist has been receiving $250,000 for nearly five years will set a lower ideal tax rate than those presented with a protagonist that has received high earnings for only a short time (and hence is not yet accustomed to it).

6.2 Results

We begin by presenting results based on the pooled sample of both takeover and lottery vignettes. In the first column of Table 9, we present the basic "Treatment effect," where *Treated* denotes that a subject was presented with the vignette where the protagonist's earnings (via corporate acquisition or lottery) increased five years ago, using the full sample of mTurk participants (even those that did not see the tax survey experiment first). *Treated* subjects choose a tax rate for the protagonist that is 1.09 percentage points lower than control subjects, significant at the 10 percent level; by comparison, the control group mean is 28.0 percent. As with the $5/$15 experiment, our preferred sample includes only those who saw the tax experiment first (and thus cannot be contaminated with the redistribution experiment). When we focus on the subsample where the tax vignette appeared first (column 2), the treatment effect increases to 1.71 percentage points (with a control group mean of 27.8).

A small fraction of subjects choose extreme values: about one percent of subjects selected a tax rate of zero while a few chose tax rates of 99 and 100 percent. In column (3) we omit the

top 1 percent and bottom 5 percent of observations to limit the influence of these extreme observations.[19] This restriction has only a slight impact on the size of the treatment effect. In Appendix Table 6, we show that the treatment effect is robust to a number of alternative ways of dealing with outliers, including estimates based on median regressions, winsorizing instead of dropping outliers, dropping only zero tax rates, and dropping regressive (<21 percent) tax rates. In all specifications, the treatment effect's magnitudes are comparable to the figures presented in Table 9. In column (4) we include controls, which has very little impact on the treatment effect.

While we tried to hold everything constant in the treatment and control arms except the strength of the reference point, it is possible respondents read other differences into the stories. We suspect that the most likely bias pushes against finding our result—in the five-year scenario, the protagonist would have had a great capacity to save and thus could cover the costs of a greater tax burden. Moreover, we suspect respondents might think it unfair that in the five-year scenario the protagonists enjoys the large raise after barely having worked for the company.

However, a concern that pushes in the opposite direction in the take-over vignette is that individuals confronted with the five-year scenario credit the protagonist with greater merit because he has worked at the larger corporation for longer. While we emphasized that in both cases the individual would receive a raise even though he would be doing the *same job* as before, in the five-year scenario the individual has apparently managed to fit in at the larger corporation, at least to the point that he has kept his (high-paying) job for half a decade. In the one-year scenario, the future performance of the protagonist at the new corporation is left unclear.

For this reason, it is useful to examine the estimates from the take-over and lottery vignettes separately, as only in the former case would this merit argument apply. Column (5) shows estimates from the take-over vignette and Column (6) the lottery vignette. While the treatment effect for the take-over vignette is larger, both are negative and are statistically indistinguishable from one another.

6.3 Discussion

While the survey experiments have documented that the strength of an individual's reference point reduces the tax rates assigned by our subjects, the precise mechanism is unclear. While a literal application of consumption commitments cannot account for the results of the \$5/\$15

[19] Another reason to exclude zero in particular is that it is the default position of the slider and thus many of these individuals may have been simply skipping through the survey.

experiment (as the money had not yet been spent and the stakes were modest), respondents could well be responding to the perceived consumption commitments of the protagonists in the vignettes, in the spirit of Chetty and Szeidl (2007). It is not unreasonable to imagine that the person who became rich five years earlier would since have taken on a hefty mortgage and enrolled her children in private schools. This consumption commitments view presents a possible foundation for the existence of a loss aversion heuristic, whether for oneself or, as is the case in our experiment, on behalf of others. In this paper we aim to document how asymmetric responses to gains and losses affect redistributive preference more generally and upper-income tax policy in particular, rather than attempting to distinguish amongst underlying explanations for this behavior.

It is also interesting to note that, at least along some policy dimensions, there appears to be respect for reference points in the distribution of *transfer* policies as well. For example, a policy that has gained popularity during the Great Recession would require welfare and food stamp recipients to pass drug tests.[20] Interestingly, in a 2011 Rasumussen poll, while 53% of respondents supported mandatory drug test for *new applicants* to welfare, only 29% supported that same requirement for *current recipients*.[21] As such, respondents seemed to view individuals' current benefits as more of an entitlement. The fact that many cuts in benefits are grandfathered in can also be viewed as an implicit respect for beneficiaries' reference points.

7 Conclusion

Past work has established that in many contexts, individuals behave as though losses relative to their reference points loom larger than analogous gains. In this paper, we consider the implications of such reference-dependent utility for standard models of optimal taxation, maintaining the assumption of a utilitarian social welfare function. As losses—even for the wealthy—loom larger than gains, much if not all of the welfare gain from redistribution may be erased. If individuals project their own loss-aversion onto others when forming their redistributive preferences, then loss-aversion might help explain the gap between voters' stated policy preferences and the more egalitarian normative prescriptions of optimal tax models or the positive predictions from political economy models.

We provide robust results that, in a laboratory setting, individuals who are given the opportunity to redistribute between two recipients with unequal endowments are highly sen-

[20]The constitutionality of such a requirement has been challenged in federal court.

[21]See `http://stopthedrugwar.org/chronicle/2011/jul/22/national_poll_finds_support_welf`.

sitive to the recipients' reference points. When the recipients do *not* know their initial endowments, the redistributor erases close to the full ex-ante income gap. However, redistribution is reduced by nearly twenty percent when the recipients *do* know their ex ante endowments. This reference-point effect is large in magnitude, more than one-half of the effect of having endowments determined via performance on an academic test versus a coin flip. Moreover, it is strongest among respondents who themselves appear the most loss-averse.

Our findings suggest that it may be useful to draw on recent findings in behavioral economics more broadly in understanding voters' preferences for redistribution. There is a rich literature—both theoretical and experimental—which aims to describe individuals' distributional preferences, almost always in the form of modified dictator games. It may be possible to link this body of research more directly to public finance questions by examining analogous questions in the type of social planner experiment we use in this paper. Similarly, the various cognitive biases that have been documented by social psychologists may have consequences for attitudes toward taxation. (One obvious example of this is optimism bias, which would lead low-income individuals to vote against progressive redistribution because of incorrect beliefs about their own future wealth.)

The existence of reference-dependent preferences in redistributive decisions may also help to explain some puzzling aspects of tax policy. For example, if wealth is a more salient reference point than income, it could help to explain the lack of broad-based support for wealth taxes (and may be reason for skepticism that recent wealth tax proposals will get much traction). Our findings may suggest that tax increases—whether based on wealth or income—might be better-received if policymakers can commit to them several years in advance of their implementation, thus allowing individuals to adjust their reference wealth or income ahead of the actual change. A fuller analysis of the consequences of reference-dependent utility for taxation—how reference points are set and evolve in response to policy changes or pronouncements; whether there are circumstances that attenuate or intensity the role of reference-dependence in redistributive preferences; and so forth—is a further area for future research.

References

Abdellaoui, M., Bleichrodt, H. and Paraschiv, C. (2007). Loss aversion under prospect theory: A parameter-free measurement. *Management Science*, **53** (10), 1659–1674.

Alesina, A. and Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, pp. 960–980.

— and Passarelli, F. (2014). Loss aversion and politics.

Allcott, H., Mullainathan, S. and Taubinsky, D. (2014). Energy policy with externalities and internalities. *Journal of Public Economics*.

Ariely, D. and Norton, M. I. (2011). Building a better americaone wealth quintile at a time. *Perspectives on Psychological Science*, **6** (1), 9–12.

Battigalli, P., Charness, G. and Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, **93**, 227–232.

— and Dufwenberg, M. (2007). Guilt in games. *The American Economic Review*, pp. 170–176.

Beggs, A. and Graddy, K. (2009). Anchoring effects: Evidence from art auctions. *The American Economic Review*, pp. 1027–1039.

Benabou, R. and Ok, E. (2001). Social mobility and the demand for redistribution: The POUM hypothesis. *Quarterly Journal of Economics*, **116** (2), 447–487.

Bernheim, B. D. and Rangel, A. (2004). Addiction and cue-triggered decision processes. *The American Economic Review*, **94** (5), 1558–1590.

Boskin, M. J. and Sheshinski, E. (1978). Optimal redistributive taxation when individual welfare depends upon relative income. *The Quarterly Journal of Economics*, pp. 589–601.

Bowman, D., Minehart, D. and Rabin, M. (1999). Loss aversion in a consumption–savings model. *Journal of Economic Behavior & Organization*, **38** (2), 155–178.

Chetty, R. and Szeidl, A. (2007). Consumption commitments and risk preferences. *The Quarterly Journal of Economics*, **122** (2), 831–877.

Di Tella, R., Haisken-De New, J. and MacCulloch, R. (2010). Happiness adaptation to income and to status in an individual panel. *Journal of Economic Behavior & Organization*, **76** (3), 834–852.

Diamond, P. and Saez, E. (2011). The case for a progressive tax: from basic research to policy recommendations. *The Journal of Economic Perspectives*, **25** (4), 165–190.

Ellingsen, T., Johannesson, M., Tjøtta, S. and Torsvik, G. (2010). Testing guilt aversion. *Games and Economic Behavior*, **68** (1), 95–107.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *The Quarterly Journal of Economics*, **116** (4), 1233–1260.

Harms, P. and Zink, S. (2003). Limits to redistribution in a democracy: a survey. *European Journal of Political Economy*, **19** (4), 651–668.

Kahneman, D., Knetsch, J. L. and Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, pp. 728–741.

— and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pp. 263–291.

Kuziemko, I., Norton, M. I., Saez, E. and Stantcheva, S. (2013). *How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments*. Working Paper 18865, National Bureau of Economic Research.

Lee, W. and Roemer, J. E. (2006). Racism and redistribution in the United States: A solution to the problem of American exceptionalism. *Journal of Public Economics*, **90** (6), 1027–1052.

Mankiw, N. G., Weinzierl, M. and Yagan, D. (2009). Optimal taxation in theory and practice. *The Journal of Economic Perspectives*, pp. 147–174.

Marshall, J. D., Knetsch, J. L. and Sinden, J. A. (1986). Agents' evaluations and the disparity in measures of economic loss. *Journal of Economic Behavior & Organization*, **7** (2), 115–127.

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

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

Nozick, R. (1974). *Anarchy, state, and utopia*, vol. 5038. Basic books.

O'Donoghue, T. and Rabin, M. (2006). Optimal sin taxes. *Journal of Public Economics*, **90** (10), 1825–1849.

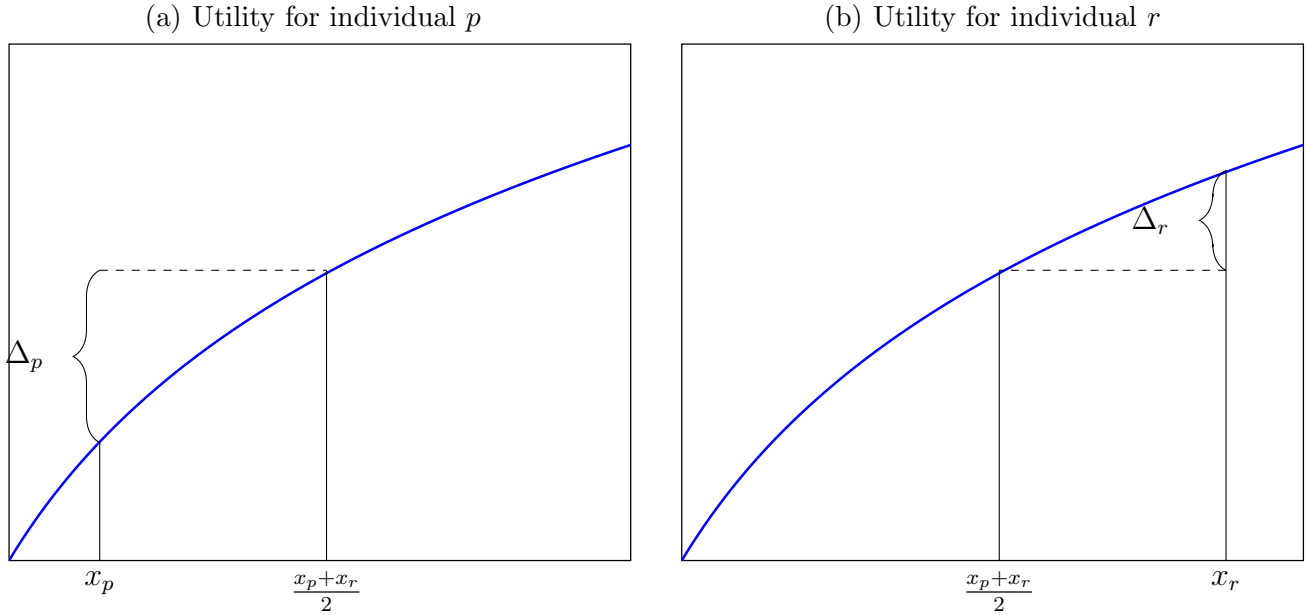
Oswald, A. J. (1983). Altruism, jealousy and the theory of optimal non-linear taxation. *Journal of Public Economics*, **20** (1), 77–87.

Roemer, J. E. (1998). Why the poor do not expropriate the rich: An old argument in new garb. *Journal of Public Economics*, **70** (3), 399–424.

Saez, E. and Stantcheva, S. (2013). *Generalized Social Marginal Welfare Weights for Optimal Tax Theory*. Working Paper 18835, National Bureau of Economic Research.

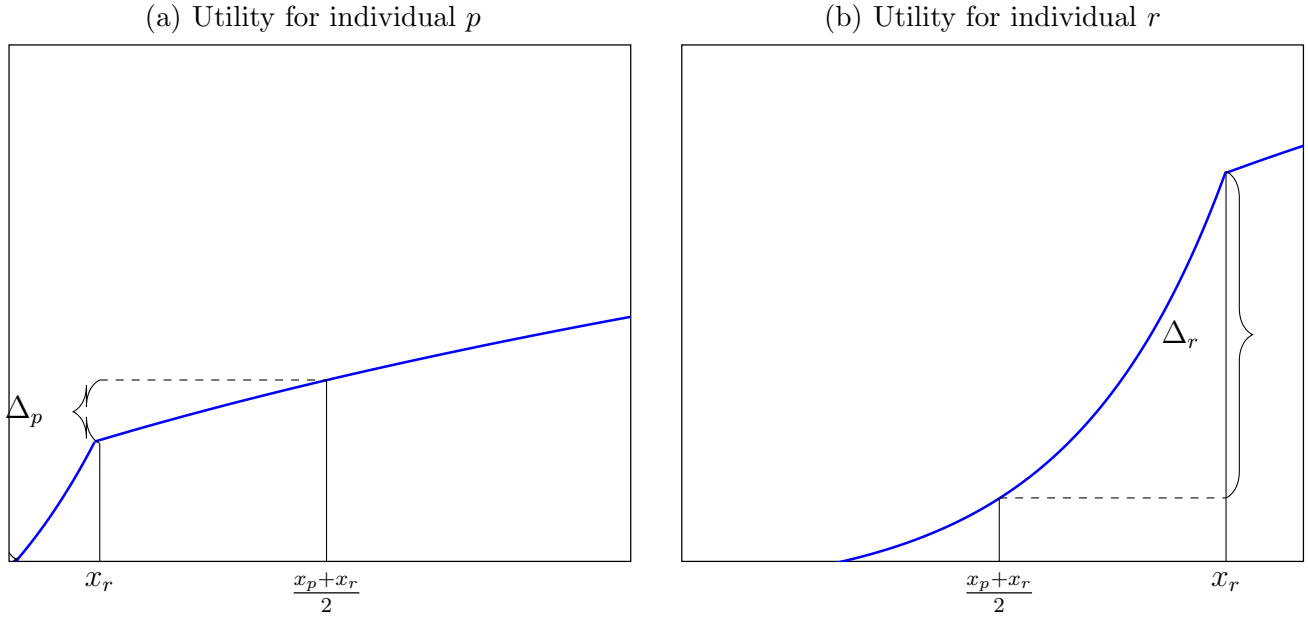
Weinzierl, M. C. (2012). *The Promise of Positive Optimal Taxation: A Generalized Theory Calibrated to Survey Evidence on Normative Preferences Explains Puzzling Features of Policy*. Working Paper 18599, National Bureau of Economic Research.

Figure 1: Redistribution with standard utility functions



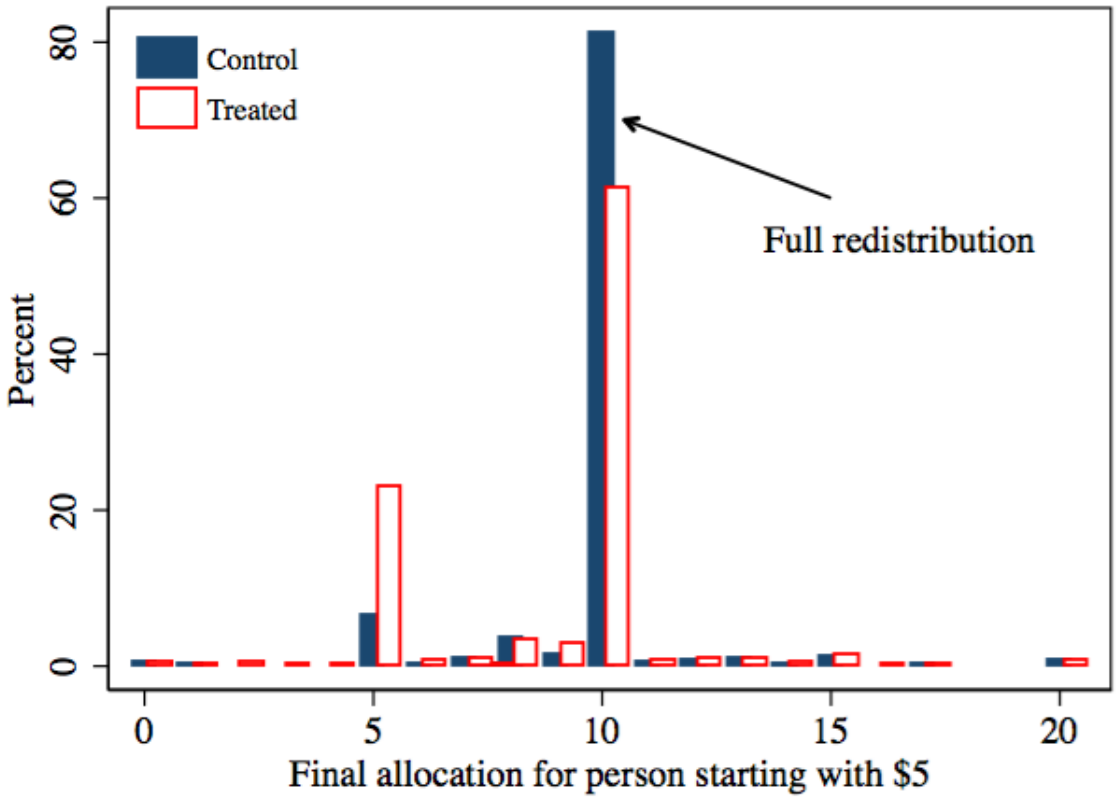
Notes: A depiction of the optimal tax solution under a utilitarian social welfare function when utility (y-axis) is a positive and strictly concave function of consumption (x-axis).

Figure 2: Redistribution with reference-dependent utility



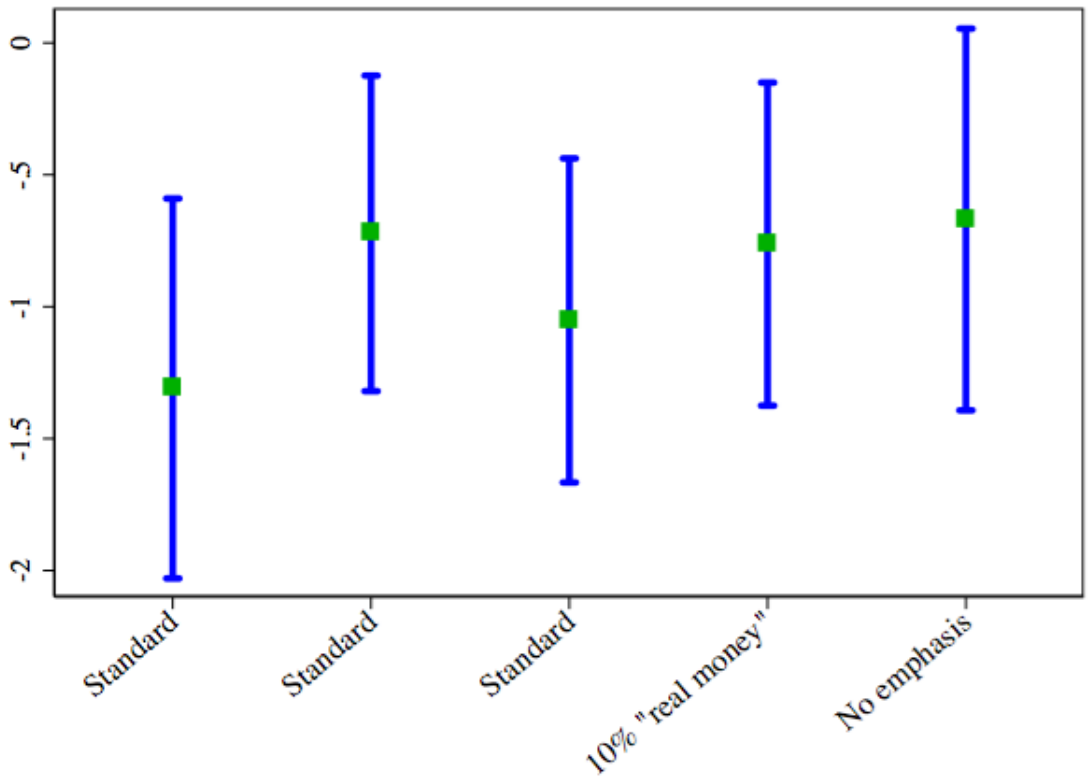
Notes: A depiction of changes in utility after full redistribution when utility functions exhibit loss-aversion. In this example, we have drawn losses from the reference point as convex and gains as concave.

Figure 3: Histogram of ex-post allocations for the ex-ante poorer player



Notes: Sample used in the figure is that in our "preferred" analysis sample (Column 3, Table 4).

Figure 4: Treatment effects and ninety-five-percent confidence intervals for the reference-point experiment, by session



Notes: We show by-session the between-subject treatment effects for all rounds in which the reference-point money-transfer experiment appears firstI (the sample in col. 3 of Table 4.. As noted in the text, there are a total of six sessions where the money-transfer experiment appears first, but one is a version where instead of the reference-point contrast, we contrast the effects of earned (via SAT questions) and unearned (via a coin toss). Those results are reported in Table 7 but are not plotted here.

Table 1: Basic summary statistics in mTurk sample compared to ACS

	(1) mTurk Sample	(2) ACS sample
Age	33.03	46.40
Female	0.437	0.515
White	0.776	0.669
Black	0.0724	0.120
Hispanic	0.0568	0.143
Asian	0.0741	0.0503
College	0.455	0.257
Income	4.969	7.132
Obervations	1,850	2,369,395

Notes: Col. 1 includes all nine sessions of the experiment. Col. 2 includes all adults in the 2010 American Community Survey (weighted with the provided individual-level weights). "Income" refers to household income (in units of $1,000).

Table 2: Full summary statistics in mTurk sample

	Mean	Std. Dev.	N
Age	33.03	11.17	1850
Female	0.44	0.50	1850
White	0.78	0.42	1850
Black	0.07	0.26	1850
Hisp	0.06	0.23	1850
Asian	0.07	0.26	1850
Income	4.97	4.00	1849
Fulltime	0.42	0.49	1850
Partime	0.14	0.34	1850
College	0.46	0.50	1850
Student	0.11	0.32	1850
Supported Obama in 2012	0.65	0.48	1849
R-loving (losses)	1.43	1.06	1348
R-loving (gains)	0.82	0.92	1383
Δ Wage unfairness for current v. new worker (cont)	1.40	1.01	1666
Higher WTA if bought house at $300K (binary)	0.79	0.41	1391

Notes: See Section 3.2 for a detailed description of the loss-aversion variables (the last four variables jn the Table). Briefly, "R-loving (losses)" takes integer values from $[0, 3]$, increasing in how the number of times you choose the lottery option over the risk free option over options involving losses. "R-loving (gains)" is defined analogously, but over gains. "Δ wage unfairness" is increasing in how much more unfair a respondent deems a wage cut to a current versus a new employee. "Higher WTA" refers to measures of anchoring bias to the original sales price of a house.

Table 3: Further summary statistics and experimental balance

	(1) Cont. mean	(2) Tr. mean	(3) Diff.	(4) P-val
Age	33.43	32.77	0.664	0.208
Female	0.435	0.432	0.00271	0.908
White	0.785	0.768	0.0173	0.378
Black	0.0687	0.0785	-0.00976	0.427
Hisp	0.0505	0.0614	-0.0110	0.311
Asian	0.0741	0.0717	0.00244	0.842
Income	4.936	4.932	0.00416	0.982
Fulltime	0.427	0.421	0.00656	0.778
Partime	0.147	0.126	0.0209	0.197
College	0.451	0.456	-0.00507	0.829
Student	0.101	0.126	-0.0253	0.0895
Supported Obama in 2012	0.646	0.642	0.00317	0.888
R-loving (losses)	1.457	1.409	0.0483	0.412
R-loving (gains)	0.805	0.850	-0.0446	0.376
Δ Wage unfairness for current v. new worker (cont)	1.359	1.434	-0.0748	0.136
Higher WTA if bought house at \$300K (binary)	0.787	0.792	-0.00504	0.821
Observations	934	891	1825	1825

Notes: Observation totals are the shared non-missing observations across all variables. Col. (1) displays means for those randomized into the control version of the \$5/\$15 money-transfer experiment (where recipients do not know their original endowment) and col. (2) displays means for the treatment version (where recipients do know their original endowment). Col. (3) subtracts col. (2) from (1) and Col. (4) is the p-value associated with $H_0 : Diff = 0$.

Table 4: Main between-subject results (using only first-stage observations)

	Dept. var: Amount redistributed						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated in first stage	-0.829*** [0.110]	-0.863*** [0.113]	-0.888*** [0.148]	-0.934*** [0.115]	-0.877*** [0.148]	-0.763** [0.312]	-0.669* [0.369]
Cont. gp. mean	4.545	4.581	4.692	4.512	4.694	4.451	4.787
Controls	No	No	No	No	Yes	No	No
Ex. short duration	No	Yes	Yes	Yes	Yes	Yes	Yes
Ex. presented second	No	No	Yes	Yes	Yes	Yes	Yes
Ex. odd choices	No	No	No	Yes	No	No	No
Ex. hypothetical	No	No	No	No	No	Yes	No
Ex. emphasis	No	No	No	No	No	No	Yes
Observations	1850	1712	1036	968	1031	195	183

Notes: All regressions include session fixed effects. Ex. short duration: exclude subjects who finish the survey in a suspiciously short amount of time. Ex. presented second: exclude survey sessions where the main redistribution experiment was not presented first. Ex. odd choices: exclude subjects who choose to make inequality-*increasing* reallocations. Ex. hypothetical: exclude survey sessions where the redistribution experiment was entirely hypothetical. Ex. emphasis: exclude survey sessions where the instructions to the redistribution experiment included underlined and italicized text and a reminder to the right of the person A and B bar chart.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 5: Between-subject results (extensive margin)

	Dep. v.: Complete redistribution			
	(1)	(2)	(3)	(4)
Treated in first stage	-0.199*** [0.0275]	-0.199*** [0.0266]	-0.198*** [0.0266]	-0.211*** [0.0627]
Cont. gp. mean	0.811	0.861	0.862	0.824
Controls	No	No	Yes	No
Ex. odd choices	No	Yes	Yes	No
Ex. hypothetical	No	No	No	Yes
Observations	1036	968	963	195

Notes: All regressions include session fixed effects. Both subjects who finished the survey very quickly and subjects not presented the distribution experiment first were excluded from these regressions. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table 6: Within-subject results

	Dept. var: Amount redistributed				
	(1)	(2)	(3)	(4)	(5)
Treatment stage	-0.853*** [0.0890]	-0.763*** [0.143]	-0.927*** [0.112]	-0.646*** [0.193]	-0.727*** [0.235]
Cont. gp. mean	4.681	4.667	4.692	4.364	4.973
Sample	All	T→C	C→T	All	All
Ex. hypothetical	No	No	No	Yes	No
Ex. emphasis	No	No	No	No	Yes
Observations	1991	941	1050	390	366

Notes: All regressions include respondent fixed effects. Subjects who finished the survey very quickly and were not presented the distribution experiment first were excluded. $C \to T$ denotes the subsample that was first randomized into the control scenario and *then* the treatment scenario. $T \to C$ denotes the subsample that was first randomized into the treatment scenario and *then* the control scenario. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table 7: Luck (control) versus merit (treatment), between-subject results

	Dep. v.: Amount redistributed		Dep. v.: Complete redistribution	
	(1)	(2)	(3)	(4)
Treated in first stage	-1.806*** [0.360]	-1.641*** [0.368]	-0.301*** [0.0664]	-0.257*** [0.0682]
Cont. gp. mean	4.515	4.510	0.699	0.696
Controls	No	Yes	No	Yes
Observations	206	205	206	205

Notes: All regressions include session fixed effects. Subjects who finished the survey very quickly or were not presented the distribution experiment first were excluded from these regressions. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table 8: Interacting loss-aversion measures with treatment status

	Dept. var: Amount redistributed							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated in first stage	-0.888*** [0.148]	-0.122 [0.326]	-0.904*** [0.262]	-0.518** [0.207]	-0.983** [0.406]	-1.063*** [0.282]	0.110 [0.701]	-0.431 [0.658]
Tr. x R-loving (losses)		-0.424** [0.184]					-0.425** [0.186]	
Tr. x R-loving (gains)			0.186 [0.209]					
Tr. x R-loving (L-G)				-0.311** [0.134]				-0.305** [0.135]
Tr. x Higher WTA					0.401 [0.460]			
Tr. x Δ Wage unfairness						0.182 [0.161]		
Cont. gp. mean	4.692	4.656	4.657	4.617	4.686	4.710	4.656	4.617
Incl. Covar x Treat	No	No	No	No	No	No	Yes	Yes
Observations	1036	607	624	599	623	865	605	597

Notes: All regressions include session fixed effects. "Tr. x R-loving (losses)", "Tr. x R-loving (gains)", and "Tr. x R-loving (L-G)" refer to the interaction of the the risk-loving over losses, risk-loving over gains, and risk-loving over losses *relative* to that over gains variables with $Treat$. "Tr. x Higher WTA" refers to the interaction of the variable that indicates that the respondent demanded a higher house price with a \$300,000 initial price than the \$250,000 price and $Treat$. "Tr. Δ Wage unfairness" refers to the interaction of the difference between the fairness ratings of the cut to the wages of the current and new coffee shop employees and $Treat$. "Inc. Covar x Treat" means that interactions with $Treat$ and the following list of variables are all included simultaneously: age, female, white income, student status, full-time status, Obama support, college degree, and support for government redistribution. The main effects of these interactions are controlled for in all regressions. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Table 9: Preferred tax for person who became rich five versus one year ago

	Dept. var: Chosen tax rate					
	(1)	(2)	(3)	(4)	(5)	(6)
Treated (rich for five yrs.)	-0.0117* [0.00644]	-0.0171** [0.00751]	-0.0168*** [0.00630]	-0.0165** [0.00640]	-0.0185*** [0.00709]	-0.0120 [0.0134]
Cont. gp. mean	0.280	0.278	0.283	0.283	0.286	0.274
Ex. if presented second	No	Yes	Yes	Yes	Yes	Yes
Adj. for outliers?	No	No	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	No	No
Vignette	Both	Both	Both	Both	Takeover	Lottery
Observations	1097	721	694	682	513	181

Notes: All regressions include session fixed effects. Subjects who finished the survey very quickly are excluded from the regression. "Adj. for outliers" indicates that the lowest five percent and the highest one percent of chosen tax rates are dropped (the asymmetry is due to a small mass of zeros, the default position of the slider). "Vignette" refers to the brief description of the event that led to the sudden increase in earnings. $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Appendix Table 1: Survey Session Details

Session	Date	Obs	First Experiment	Hypothetical	Emphasis	$250,000 Tax Q
One	Feb 13, 2014	187	Redistribution	Yes	Yes	No
Two	Feb 27, 2014	295	Redistribution	Yes	Yes	No
Three	Mar 10, 2014	250	Redistribution	Yes	Yes	No
Four	Mar 21, 2014	282	Tax	Yes	Yes	Yes
Five	Mar 24, 2014	303	Tax	Yes	Yes	Yes
Six	Apr 25, 2014	228	Redistribution	Yes	Yes	No
Seven	May 28, 2014	207	Redistribution	No	Yes	Yes
Eight	May 30, 2014	216	Redistribution	Yes	No	Yes
Nine	Jun 19, 2014	200	Tax	Yes	Yes	Yes

Appendix Table 2: Correlation across loss-aversion measures

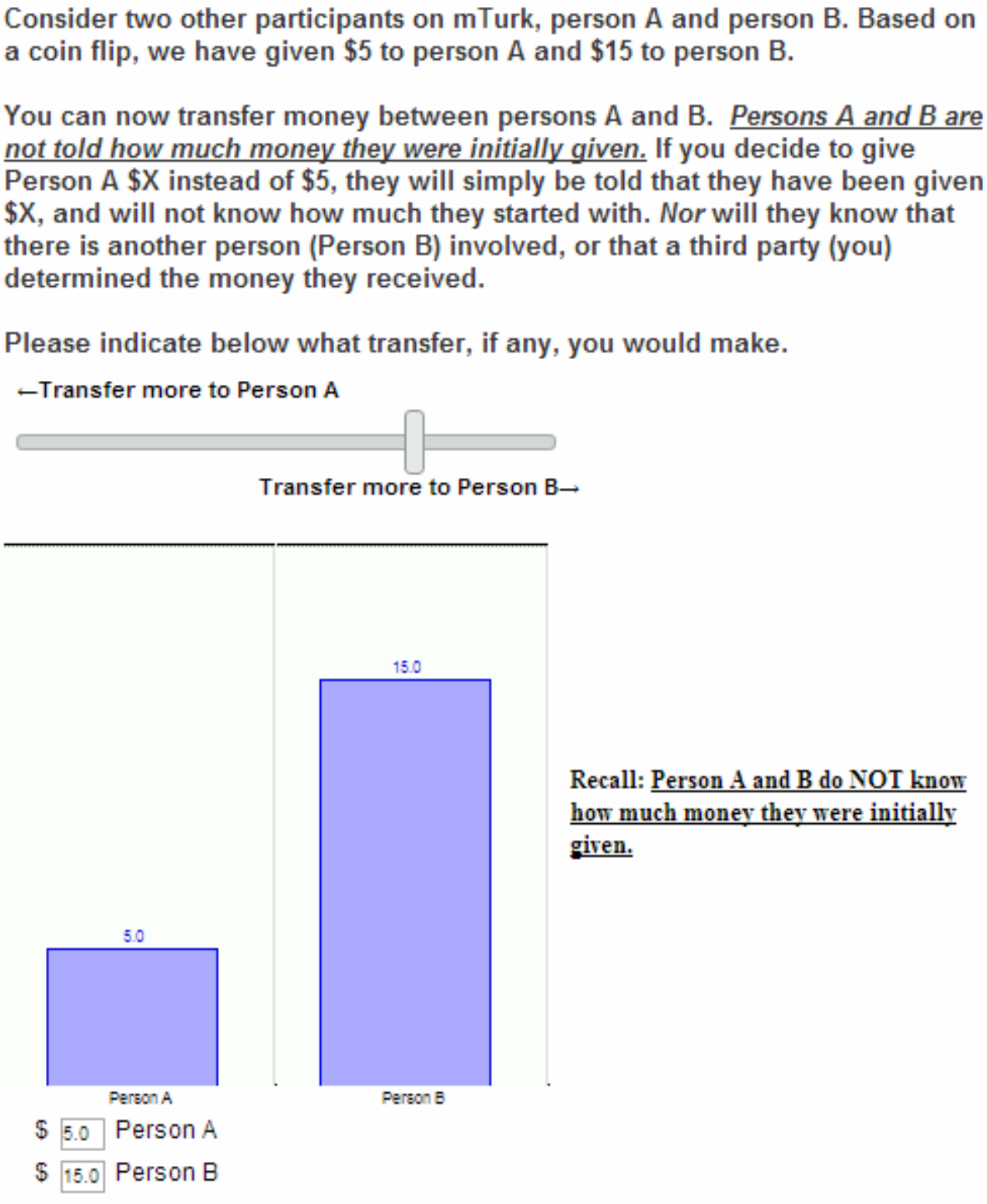
	R-loving (losses)	R-loving (gains)	R-loving (L-G)	Higher WTA	Δ Wage unfairness
R-loving (losses)	1				
R-loving (gains)	-0.0253	1			
R-loving (L-G)	0.762***	-0.667***	1		
Higher WTA	0.0953***	0.00342	0.0683**	1	
Δ Wage unfairness	-0.00875	0.0308	-0.0330	0.0842***	1
Observations	1679				

Appendix Table 3: Assessing perceptions of bias and survey fatigue, by redistribution experiment survey arm

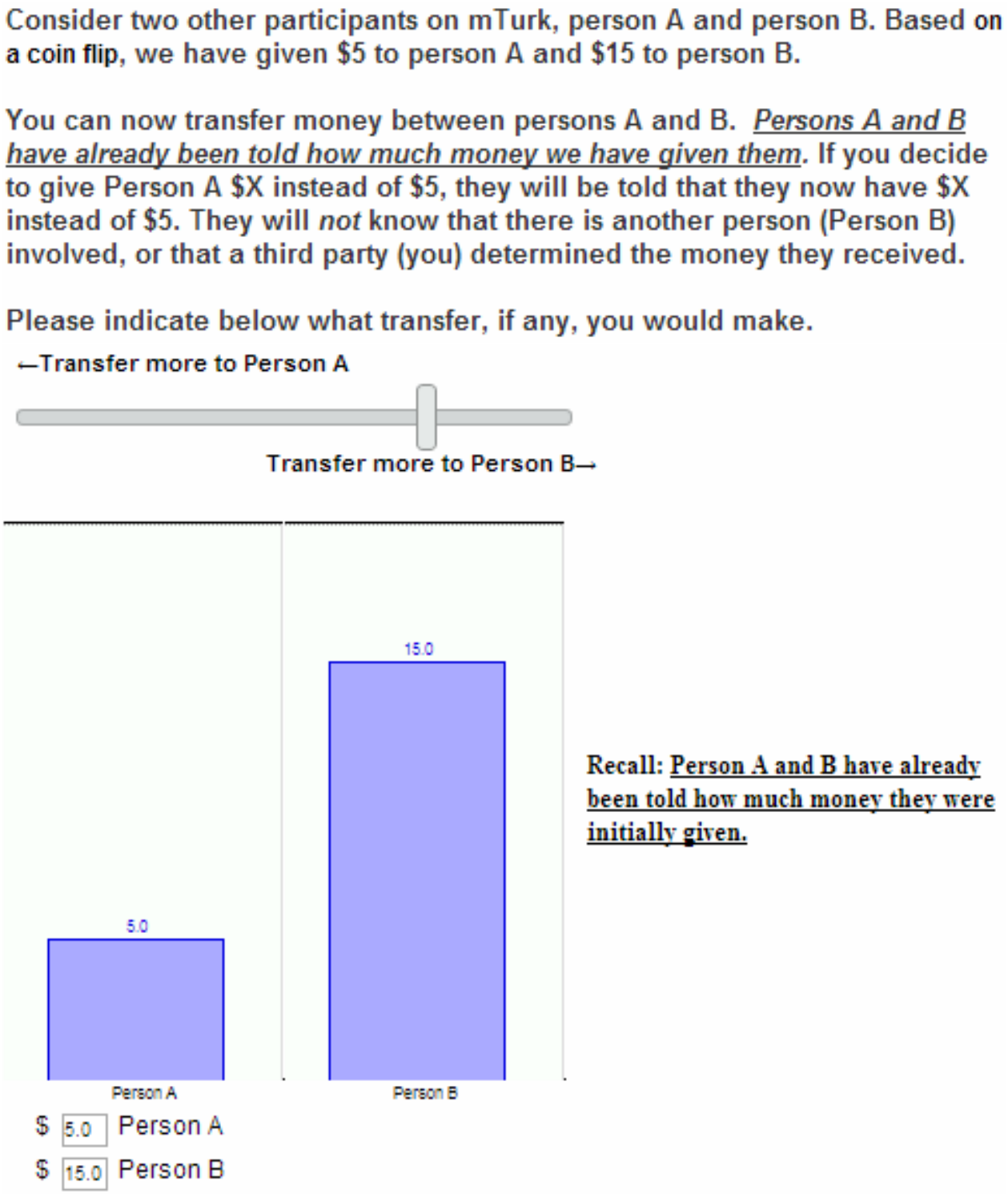
	(1) LW bias	(2) RW bias	(3) No bias	(4) Minutes
Treated in first stage	0.0253 [0.0177]	0.0102 [0.0113]	-0.0392* [0.0210]	0.604 [0.433]
Cont. gp. mean	0.0763	0.0286	0.889	12.21
Observations	1031	1031	1031	866

Notes: Subjects who finished the survey very quickly and were not presented the distribution experiment first were excluded from these regressions.

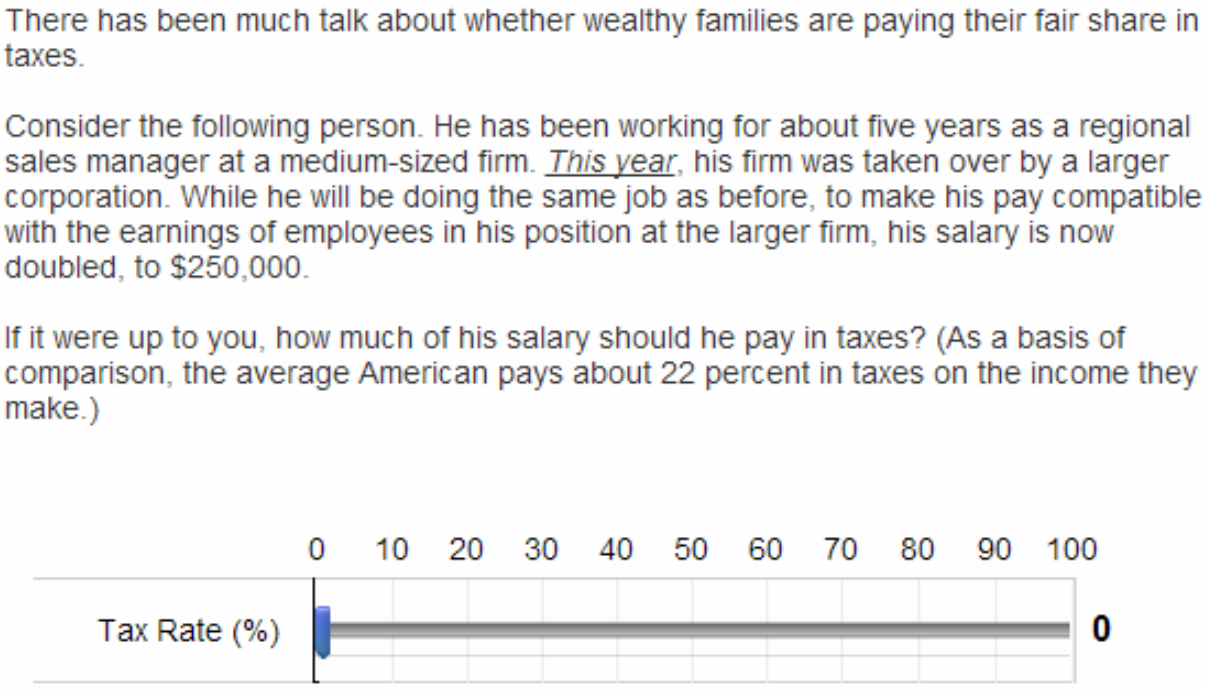
Appendix Figure 1: Main redistribution experiment (control arm)



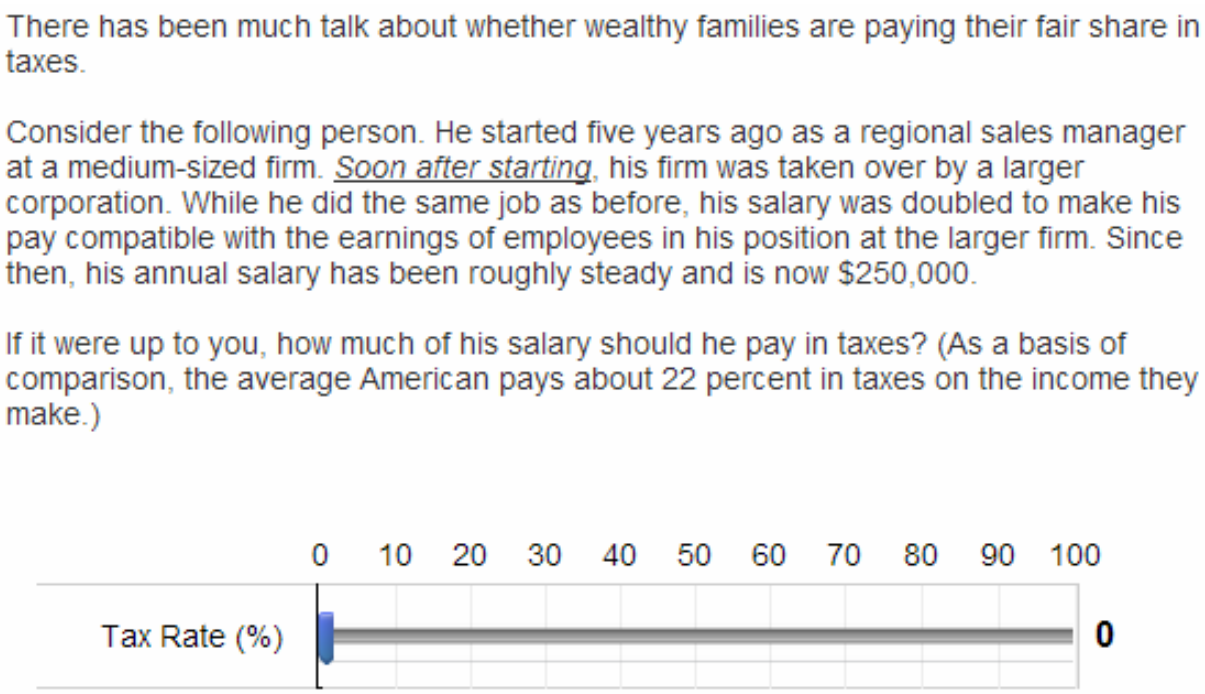
Appendix Figure 2: Main redistribution experiment (treatment arm)



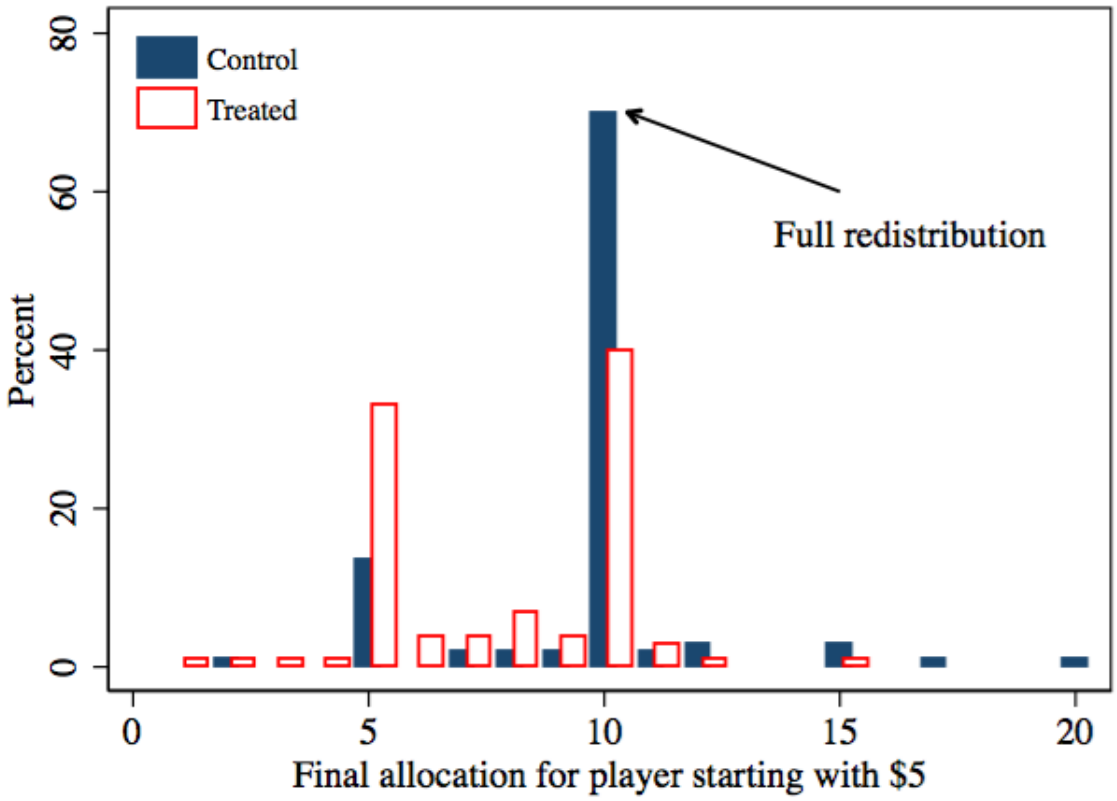
Appendix Figure 3: Income tax experiment (control arm)



Appendix Figure 4: Income tax experiment (treatment arm)



Appendix Figure 5: Histogram of ex-post allocations for the ex-ante poorer player, luck (control) versus merit (treatment) session



.

Appendix Table 4: Differential treatment effects from between-subject results

	Dept. var: Amount redistributed										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Treated in first stage	-1.490*** [0.470]	-0.900*** [0.197]	-1.003*** [0.327]	-0.885*** [0.153]	-0.900*** [0.154]	-0.848*** [0.153]	-0.739*** [0.243]	-0.926*** [0.156]	-0.906*** [0.196]	-0.845*** [0.161]	-0.700*** [0.204]
Tr. x Age	0.0179* [0.0132]										
Tr. x Female		0.0373 [0.300]									
Tr. x White			0.150 [0.367]								
Tr. x Black				0.0251 [0.613]							
Tr. x Hisp					0.243 [0.601]						
Tr. x Asian						-0.582 [0.618]					
Tr. x Income							-0.00000302 [0.00000401]				
Tr. x Student								0.464 [0.512]			
Tr. x Fulltime									0.0526 [0.302]		
Tr. x Partime										-0.261 [0.423]	
Tr. x College											-0.393* [0.300]
Cont. gp. mean	4.694	4.694	4.694	4.694	4.694	4.694	4.694	4.694	4.694	4.694	4.694
Observations	1032	1032	1032	1032	1032	1032	1032	1032	1032	1032	1032

Notes: All regressions include session fixed effects and the controls listed in Table 4, excluding Obama support and support for government redistribution. Subjects who finished the survey very quickly and were not presented the distribution experiment first were excluded from these regressions. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$

Appendix Table 5: Interacting loss-aversion measures with treatment status (SAT v. coin toss experiment)

	Dept. var: Amount redistributed							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated in first stage	-1.806*** [0.360]	-1.963*** [0.612]	-2.135*** [0.479]	-1.695*** [0.400]	-1.230 [0.819]	-2.155*** [0.673]	-1.848 [1.252]	-1.760 [1.223]
Tr. x R-loving (losses)		0.0914 [0.361]					-0.146 [0.372]	
Tr. x R-loving (gains)			0.438 [0.381]					
Tr. x R-loving (L-G)				-0.200 [0.264]				-0.296 [0.259]
Tr. x Higher WTA					-0.758 [0.913]			
Tr. x Δ Wage unfairness						0.239 [0.381]		
Cont. gp. mean	4.515	4.545	4.510	4.541	4.510	4.515	4.541	4.515
Incl. Covar x Treat	No	No	No	No	No	No	Yes	Yes
Observations	206	195	203	192	205	206	194	195

Notes: All regressions include session fixed effects and the controls listed in Table 4. "Tr. x R-loving (losses)", "Tr. x R-loving (gains)", and "Tr. x R-loving (L-G)" refer to the interaction of the the risk-loving over losses, risk-loving over gains, and risk-loving over losses *relative* to that over gains variables with $Treat$. "Tr. x Higher WTA" refers to the interaction of the variable that indicates that the respondent demanded a higher house price with a \$300,000 initial price than the \$250,000 price and $Treat$. "Tr. Δ Wage unfairness" refers to the interaction of the difference between the fairness ratings of the cut to the wages of the current and new coffee shop employees and $Treat$. "Inc. Covar x Treat" means that interactions with $Treat$ and the following list of variables are all included simultaneously: age, female, white income, student status, full-time status, Obama support, college degree, and support for government redistribution. The main effects of these interactions are controlled for in all regressions. $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$

Appendix Table 6: Replicating the main tax result with various outlier adjustments

	Dept. var: Chosen tax rate				
	(1)	(2)	(3)	(4)	(5)
Treated (rich for five yrs.)	-0.0168*** [0.00630]	-0.0300*** [0.00883]	-0.0169** [0.00671]	-0.0158** [0.00739]	-0.0127** [0.00562]
Cont. gp. mean	0.283	0.278	0.279	0.280	0.304
Outlier adjustment	Drop (orig. spec.)	Qreg	Winsorize	Drop 0s	Drop regr.
Observations	694	721	721	717	602

Notes: The first column replicates the preferred specification from Table 9 (col. 3), where outliers below the bottom fifth percent and above the higher one percentile are dropped. "Qreg" refers to median regression on the entire sample. "Winsoring" winsorizes the outliers in Col. (1) instead of dropping them. Col. (4) merely drops those who choose a zero tax rate while col. (5) drops anyone who chooses a regressive tax rate (i.e., a rate less than the average rate of 22 percent). $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$

Details from the excluded experimental session

In one session (April 17, 2014), instead of having individuals transfer money between the strangers using the slider, we tried to use a drop-down menu (which listed all twenty possible money transfers). In all cases, "no transfer" was listed first, which likely caused strong anchoring effects. More seriously, we only realized ex post that the drop-down menu covered up the graphic of the two strangers' endowments (which in the slider version respondents could see change in real time as they moved from one allocation to the other).

We report results from this session below but do not include them in our main results.

	Dept. var: Amount redistributed		
	(1)	(2)	(3)
Treated in first stage	-0.151 [0.336]	-0.467 [0.318]	-0.521 [0.338]
Cont. gp. mean	3.747	3.713	3.699
Controls	No	No	Yes
Ex. short duration	Yes	Yes	Yes
Ex. odd choices	No	Yes	Yes
Observations	209	200	197

Notes: Subjects who finished the survey very quickly were excluded.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$