Willingness to Pay and Willingness to Accept are Probably Less Correlated Than You Think*

Jonathan Chapman
NYUAD
jchapman@nyu.edu
www.jnchapman.com

Mark Dean
Columbia
mark.dean@columbia.edu
www.columbia.edu/~md3405/

Pietro Ortoleva
Princeton
pietro.ortoleva@princeton.edu
ortoleva.mycpanel.princeton.edu/

Erik Snowberg
Caltech, UBC, CESifo, NBER
snowberg@caltech.edu
hss.caltech.edu/~snowberg/

Colin Camerer
Caltech
camerer@hss.caltech.edu
hss.caltech.edu/~camerer/

January 31, 2019

Abstract

An enormous literature documents that willingness to pay (WTP) is less than willingness to accept (WTA) a monetary amount for an object, a phenomenon called the *endowment effect*. Using data from an incentivized survey of a representative sample of 3,000 U.S. adults, we add one (probably) surprising additional finding: WTA and WTP for a lottery are, at best, slightly correlated. Across all participants, the correlation is slightly negative. We also collect data from published, incentivized studies, all run on university students, to analyze the correlation between WTA and WTP, which those studies did not examine. We document a correlation of 0.15–0.2, which is consistent with the correlation for high-IQ participants in our own data. While poorly related to each other, WTA and WTP are closely related to different measures of risk aversion, and relatively stable across time. Models of reference dependence can explain these correlations, but are inconsistent with other aspects of our data, suggesting the need for more theories and empirical studies of the processes of buying and selling.

JEL Classifications: C90, D81, D91
Keywords: Willingness To Pay, Willingness To Accept, Endowment Effect, Loss Aversion

*We thank Douglas Bernheim, Benedetto De Martino, Stefano Della Vigna, Xavier Gabaix, Daniel Gottlieb, Eric Johnson, David Laibson, Graham Loomes, Matthew Rabin, Jan Rivkin, Peter Wakker, Michael Woodford, and the participants of seminars and conferences for their useful comments and suggestions. Evan Friedman and Khanh Ngoc Han Huynh provided research assistance. Camerer, Ortoleva, and Snowberg gratefully acknowledge the financial support of NSF Grant SMA-1329195.

1 Introduction

In standard economic theory, willingness to pay (WTP) and willingness to accept (WTA) a monetary amount for an object are the same.[1] An enormous experimental and empirical literature finds this is not the case: WTA is generally higher than WTP, a phenomenon referred to as the *endowment effect*.[2]

We add one, probably surprising, finding to this enormous literature: in our data WTA and WTP are not only different, but are, at best, slightly correlated. Our results come from three large, incentivized, and representative surveys of U.S. adults totaling 3,000 people. Overall, the correlation between WTA and WTP is slightly negative. For subgroups of the population, the correlation is almost always small, typically statistically insignificant, and often negative. This is despite the fact that there is an endowment effect (WTA > WTP) for a majority of participants, and that there is substantial heterogeneity in both measures. Each incentivized survey elicits WTA and WTP for two different lottery tickets, which allows us to account for measurement error (using both averaging and ORIV; Gillen et al. Forthcoming). This does not change the results. Nor does excluding participants that give extreme answers. Nor do a number of other refinements. As we have repeated measures six months later from half our participants, we can show that our measures are relatively stable across time. This suggests that the low correlation between WTA and WTP is not due to excessive measurement error. Moreover, as described below, WTA and WTP are each strongly related to *different* aspects of risk aversion.

Our results suggest that the use of a buying or selling frame is of first-order importance in determining preferences. All of our measures show substantial heterogeneity across participants. Yet, individual attitudes towards risk are relatively stable within a frame: valuations of different lotteries have correlation coefficients of 0.67–0.71 for WTA and 0.74–0.79 for WTP. In comparison, individual attitudes are almost completely unrelated across

[1]This assumes that the value of the object is small enough that wealth effects are irrelevant.
[2]See Camerer (1995, p. 665), Dhami (2016, p. 217), and Marzilli Ericson and Fuster (2014) for reviews.

frames: the correlations between WTA and WTP for the same lottery are between -0.09 and -0.01. Indeed, even implicitly framing a question as buying or selling a lottery—using the approach of Sprenger (2015)—is enough to produce a very small correlation between different measures of risk aversion, as described below. Together with evidence from psychology and neuroscience, reviewed in our final section, this suggests that it may be useful to consider valuations in buying and selling as coming from distinct processes, rather than from a shading up or down of some core value given by the standard model.

The data in this paper come from two studies, one of which contained two waves with the same participants. The first study consisted of 2,000 participants. This was designed to explore the relationships between a wide variety of behavioral and political measures over time (Camerer et al., 2019a). We re-contacted the participants six months later. Of the original 2,000 participants, 1,465 participated in this second wave. The surprising relationship between WTA and WTP, and a few other factors, encouraged us to run a second study of 1,000 independent participants (Camerer et al., 2019b). Our large sample size allows us to document relevant facts precisely, even for small sub-groups. As we have repeated measures from half our participants, we can show that our measures are relatively stable across time, again reducing concerns about measurement error.

To understand the plausibility of our findings in light of the existing literature, we examine the correlation of WTA and WTP in published studies containing within-person, incentivized measures of the endowment effect for lotteries. Although several studies satisfy these criteria, none examine or report on the relationship between WTA and WTP. Only three (total $N = 466$) have available data. These studies differ from ours in both elicitation methodology and participant pools. Re-examining their data, we find the correlation between WTA and WTP differs depending on the study, but across all three the average is around 0.2—still small, but higher than in our data. This is, however, consistent with the correlation in our data for those most similar to lab participants—those with high IQs and post-secondary education—suggesting that the difference between our study and prior work is due to differences in participant populations, rather than methodology. Moreover, it seems that correlations between WTA and WTP are quite meager, even in the lab with university students.

Having established the robustness of our primary finding in several ways, we turn to exploring potential explanations, and their implications. As WTA and WTP are measured for lottery tickets, we begin by examining the relationship between them and other measures of risk aversion. Our second study contained six such measures. Analyzing their correlational structure, we find two *clusters*: groups of variables that tend to be strongly related to each other, but poorly related to variables in other groups. In particular, these clusters are largely determined by whether a question asks for a certainty equivalent or a lottery equivalent.[3] Interestingly, our measure of WTA is highly related to measures of risk aversion in the certainty-equivalent cluster, while WTP is related to probability-equivalent measures. Overall, this suggests that WTA and WTP, although unrelated to each other, are instead strongly related to other (different) measures. This, once again, suggests that WTA and WTP are measuring something, just not the same thing.

The specific clusters can be interpreted using the framework of Sprenger (2015), in which question structure serves as an implicit frame and determines reference points. In particular, Sprenger shows that certainty and probability equivalents yield different levels of risk aversion. He explains this finding by appealing to the different reference points these frames induce. This interpretation helps to make sense of the clusters described in the preceding paragraph. This approach also provides an alternative way of measuring the correlation between WTA and WTP using a different hypothesis of reference-point formation that does not rely on explicit buying and selling frames. The fact that we find a zero or negative correlation between certainty equivalents and lottery equivalents supports our main finding. Our results also provide additional support for Sprenger's approach to reference-point formation by testing if the endowment effect obtained from questions with explicit framing (WTA/P) is related to the one obtained from implicit framing (certainty vs. probability equivalents). We

[3] See Camerer et al. (2019b) for a careful analysis. Imperfect correlations between different measures of risk preferences are unsurprising based on the current literature, see Section 6.

find strong support for Sprenger's approach: a correlation of 0.54 between these quantities.

In our final analytical section, we compare our data to existing theories of reference dependence. In the standard model, WTA and WTP are perfectly correlated, and both determined by the curvature of the utility function. In the standard versions of Cumulative Prospect Theory (CPT) and the Kőszegi Rabin (KR) model that allow for an endowment effect, the curvature of the utility function affects both WTP and WTA, while loss aversion usually affects only one of the two. If curvature and loss aversion are independently distributed, then variation in utility curvature leads to a strong positive correlation between WTA and WTP, which is incompatible with our findings.

Another possible theoretical explanation—one that also appears to be incompatible with our data—is that there is no variance in utility curvature. This is the case, for example, in the popular version of reference-dependent models that assumes linear utility over small-stakes gambles, and thus ascribes all risk averse behavior towards such gambles to loss aversion. This has often been suggested in the literature to address the Rabin (2000) paradox (for example Kőszegi and Rabin, 2007). However, if loss aversion affects only one of WTA and WTP, the other should exhibit variation only due to noise. But this is incompatible with the fact that we find large heterogeneity for both measures, and that, as we have seen, neither is primarily noise (as they are both stable over time and related to other measures). Alternatively, one could model people as encoding the amount spent in WTP as a loss in CPT (in addition to encoding the object sold as a loss in WTA). This implies loss aversion increases WTA and decreases WTP; however, this would lead to a strong negative correlation, which is also inconsistent with our data.

A final theoretical possibility is to allow for both variation in utility curvature—which would lead to positive correlation—and for loss aversion to enter both WTA and WTP. In this case, precisely calibrated variation in loss aversion could offset variation in utility curvature, leading to a zero correlation. This last approach is part of a more general point: taking full advantage of all degrees of freedom in these models can induce a wide range of correlations. In particular, allowing for arbitrary joint distributions of loss aversion and the curvature of utility can yield almost any correlation between WTA and WTP—from large and negative to large and positive—including zero.

Although both CPT and KR models are flexible enough to be compatible with almost any correlation, the richness of our data allows further tests of these models. Our data also contains the common finding that risk aversion changes depending on whether lotteries contain gains, losses, or both—the phenomenon that motivated loss aversion in prospect theory. To capture this finding together with the endowment effect, hybrid versions of the CPT or KR models are needed. In particular, the reference point needs to be affected both by the relationship of payments to a fixed reference point of zero, and by the framing of the question. Crucially, these models make one additional prediction: that the endowment effect is correlated with the increase in risk aversion for lotteries with both gains and losses, since both phenomena are driven by loss aversion. We test this predicted relationship in two ways, and find no evidence for it.

Models of reference dependence can rationalize all our results—including the last—by positing that the reaction to losses in lotteries and the endowment effect are driven by different types of loss aversion. In that case, our results have several implications for reference dependent theories. First, in line with Sprenger (2015), they tell us something about how reference points work: they can be set either by the structure of the MPL or by external framing. Second, hybrid theories are needed to allow for the multiple reference points that can exist in a single choice. Third, our results put specific constraints on the distribution of parameters in the population. Fourth, and finally, they rule out some popular special cases of these theories: as noted above, our results rule out models that attribute all small-stakes risk aversion to loss aversion.

However, models of reference dependence have long tied together the reaction to losses in lotteries and the endowment effect (Kahneman et al., 1990). Thus, our final result suggests the usefulness of considering other psychological theories of the endowment effect, summarized in our last section. In particular, these theories focus on the distinct *processes*—cognitive, attentional, emotional, and social—underlying the activities of buying and selling, rather than simply modeling them as gains and losses. This focus is consistent with our data, and with data from psychology and neuroscience that is not addressed by economic theories of the endowment effect. While this literature is still developing, it suggests, together with our findings, the need for more research into the central economic activities of buying and selling.

2 Design and Data

Our data come from three incentivized, representative surveys of U.S. adults conducted online by YouGov. YouGov maintains a panel of participants. It continually recruits new people, especially from hard to reach and low socio-economic-status groups. To generate a representative sample, it randomly draws people from various Census Bureau products, and matches them on observables to members of their panel. Differential response rates lead to the over-representation of certain populations. Thus, YouGov provides sample weights to recover estimates that would be obtained from a fully representative sample.[4] We use these weights throughout the paper, including when assessing the percentage of participants with a certain response or trait. Unweighted results are in Appendix A.

All three surveys were incentivized. That is, participants were paid based on outcomes associated with their choices. These choices identified a large number of attitudes, including

[4]According to Pew Research, YouGov's sampling and weighting procedure yields better representative samples than traditional probability sampling methods with non-uniform response rates, including Pew's own probability sample (Pew Research Center, 2016, YouGov is Sample I). As economists rarely run their own surveys in representative populations, it is worth explaining how the survey research literature uses the term "representative." With few exceptions—censuses, and samples in rural areas of developing countries based on a census—representative samples are representative on observables, not on unobservables. While random samples have the potential to be representative on both observables and non-observables, low response rates render these samples less representative on both observables and expressed preferences, as the Pew study documents. Commonly-used representative surveys in economics, such as the Current Population Survey, use weighting to account for non-response. The CPS also uses imputation to adjust for item non-response, which is not present in our survey (see www.census.gov/programs-surveys/cps/technical-documentation/methodology/imputation-of-unreported-data-items.html).

WTA and WTP (see Camerer et al., 2019a,b, for more details on these other questions). Two choices were selected for payment after the participant completed the entire survey.[5] We elicited WTA and WTP for lottery tickets. Following the literature, we used framing to endow participants with those tickets. This has been shown to reliably produce an endowment effect (see, for example, Isoni et al., 2011).

All outcomes were expressed in points, an internal YouGov currency. Points can be converted to U.S. dollars, or prizes, using the approximate rate of $0.001 per point. The average payment to participants was around $9 (9,000 points). The survey took participants between 45 and 60 minutes. This compensation level is quite high for an internet survey, and represents a rate of pay approximately three times the average for YouGov surveys.

The ordering of many of the questions, including WTA and WTP, was randomized.[6] The first survey, which we call Study 1, Wave 1 (W1) contained 2,000 participants and was conducted between March 27 and April 3, 2015. A second wave (Study 1, Wave 2 or W2), recontacted the same population and received 1,465 responses between September 21 and November 23, 2015. The attrition rate of $\sim 25\%$ is lower than most online surveys. This is due, in part, to YouGov's panel management, and in part to the large incentives we offered. The third survey (Study 2) used an independent sample of 1,000 participants, and was run between March 30 and April 14, 2016. Combining the first wave of the first study with the second study gives us a total population of 3,000 participants, with two observations for approximately half of them.

2.1 Measuring WTA and WTP

[5] This is incentive compatible under Expected Utility, but not necessarily under more general risk preferences, where no such mechanism may exist (Karni and Safra, 1987). A growing literature suggests this theoretical concern may not be empirically important (Beattie and Loomes, 1997; Cubitt et al., 1998; Hey and Lee, 2005; Kurata et al., 2009), but there are some exceptions (Freeman et al., 2015).

[6] The average WTA, WTP, and endowment effect are the same regardless of whether WTA was randomly selected to appear before WTP, or WTP before WTA. As shown in Table 4, the correlation between WTA and WTP is also independent of question ordering.

Each study contained incentivized measures of WTA and WTP for two lottery tickets. This allows a *within*-subject design. This is necessary for measuring the correlation between WTA and WTP, but unusual in the literature (see Section 5). Using two lottery tickets allows us to correct for measurement error, as discussed in Section 2.2.[7]

Each elicitation is performed using a multiple price list (MPL; Holt and Laury, 2002).[8] An MPL consists of a table with two columns of outcomes. In each row, the participant is asked to make a choice between the outcomes in the columns. One column contains the same outcome in all rows, while outcomes in the other column vary. In the latter column, the outcome becomes more attractive as one moves down the table. Participants who understand the question should choose the former option for early rows, and at some point switch to choosing the latter (varying) option. In all rows below that point, the participant should also choose the latter option.[9] To increase participants' understanding, the first and the last row of the MPL always involved a dominated option—for example, 2,000 points or a lottery that pays 2,000 or 8,000 with equal probability—with the undominated option pre-selected.[10]

[7]Approximately half of the participants received the WTA questions first, and the other half the WTP questions first. In Study 1, Wave 1, these questions appeared in the second and seventh block of questions (out of 11). The same positions were used in Wave 2, but the randomization over the ordering of WTA and WTP was conducted independently of the first wave. In Study 2, these questions appeared in the third and seventh block (out of 12). A qualitative question about an unrelated topic was asked between the two elicitations of WTA, and also between the two elicitations of WTP.

[8]The MPL is generally considered easier to explain to participants than, for example, incentivized pricing tasks (see, for example, Cason and Plott, 2014).

[9]Participants who do not understand the question may switch back and forth between columns multiple times. However, in our implementation, participants were not allowed to proceed if there were multiple switches in their choices. They were also not allowed to proceed unless there was a choice in every row. This approach is based on the observation that choices with multiple switches are often thrown out by researchers. This design decision limits our ability to study other types of behaviors that may be indicated by multiple switches, but those are not our focus. In practice, very few instances of multiple switches were observed. This is due to extensive training before the participant answered their first MPL.

[10]We use the average value of the varying option over the two rows where the participant's choices switched columns. Using the minimum or maximum value does not alter results.

The elicitation of WTA had the following form:[11]

> For this question, you are given a lottery ticket that has a 50% chance of paying you 10,000 points, and a 50% chance of paying you 0 points. You have two options for this lottery ticket:
>
> - Keep it, or
> - Sell it for a certain amount of points (for example, 2,000 points).

Participants were then asked "For each row in the table below, which option would you prefer?" and were presented with an MPL with the option "Keep it" and the option "Sell it for x points," where x changes with the row.

There were two lottery tickets with different payoffs. Table 1 describes the lotteries used in each study. Both waves of Study 1 used the same lotteries. In Study 2, neither lottery had the possibility of a 0 payoff. This feature was included as a robustness check. Note that all lotteries are mean-preserving spreads of each other.

For WTP, the same lottery tickets were used, but participants were instead told:

> For this question, you have been given 10,000 points. You will be offered the opportunity to exchange some of these points for a lottery ticket. This lottery ticket has a 50% chance of paying you 10,000 points, and a 50% chance of paying 0 points.
>
> For example, if you choose to pay 1,000 points for a lottery ticket, and this question is chosen for payment, you will:
>
> - Pay 1,000 points for the lottery ticket;
> - Keep 9,000 points for yourself; and
> - Earn whatever proceeds you get from the lottery ticket (if any).

Participants were shown an MPL with the options "Keep 10,000 points" and "Buy the lottery ticket for $(10,000 - x)$ points and keep the remaining x points", with x varying by row.

In order to identify participants that chose dominated options without carefully reading the question, an additional design change was made to Study 2. In particular, we included

[11] This is taken from our study design documents. These, and screenshots of the study, may be found at hss.caltech.edu/~snowberg/wep.html. Screenshots of the WTA and WTP questions may be found in Appendix D.

Table 1: Lotteries Used

	Lottery 1	Lottery 2
Study 1, Wave 1 & 2	$0.5 * 0 \oplus 0.5 * 10{,}000$	$0.5 * 2{,}000 \oplus 0.5 * 8{,}000$
Study 2	$0.5 * 1{,}000 \oplus 0.5 * 9{,}000$	$0.5 * 2{,}000 \oplus 0.5 * 8{,}000$

two dominated options at the top and bottom of each MPL. One was pre-selected, consistent with Study 1. The second could be chosen or avoided by participants.[12] If a participant chose a dominated option in Study 2, he or she was alerted to this on the next screen. He or she was given the option to go back and change their choice or to continue. For WTA and WTP, 66% of those alerted to a dominated choice did not change it. We present some results below which exclude those who chose a dominated option.

In order to elicit the endowment effect we explicitly use the language of buying and selling in the WTA and WTP questions, as in Isoni et al. (2011). That paper finds a robust endowment effect for lotteries that survives various checks designed to reduce participant misperceptions. Another line of research, including Sprenger (2015), uses implicit framing in questions to manipulate reference points and elicit the endowment effect. As noted in the Introduction, in Section 6 we use Sprenger's approach and data from other risk elicitations in our study to show that our results are robust to this design choice.

We do not explicitly control for the expectations of participants (see Marzilli Ericson and Fuster, 2011; Heffetz and List, 2014, for examples of explicitly controlling for expectations). However, we see substantial differences between WTA and WTP, so our manipulation is clearly changing something important about the choice environment. Viewed through the lens of reference dependent theories, this implies we are changing the reference point. In particular, the reference point is either determined entirely by the endowment (as supported by

[12] In particular, in Study 1, the top and bottom row contained the maximum and minimum payment of the lottery. The lottery was pre-selected when the alternative was the minimum of the lottery. When the alternative was the maximum of the lottery, this was pre-selected instead of the lottery. However, in Study 2, the pre-selected options were in rows with sure payments greater than the maximum payment, and less than the minimum payment, of the lottery.

evidence in Heffetz and List, 2014), or the endowment changes expectations, which changes the reference point (as proposed by Sprenger, 2015). In Section 7 we discuss the implications of different assumptions about reference point formation for our data, and how these interact with various models of reference-dependent behavior.

2.2 Measurement Error

A common concern with statistically insignificant findings is that they are caused by attenuation bias due to measurement error. Eliciting WTA and WTP for two lotteries for each participant allows us to substantially reduce this concern. We approach this in two ways.

First, we average together the two measures of WTA and WTP in each study. This reduces, but does not eliminate, attenuation due to measurement error. The estimate of the correlation is both biased and inconsistent.

Second, we can use the ORIV technique (Gillen et al., Forthcoming). This adapts an errors-in-variables instrumental variables (IV) approach to our data, and produces consistent estimates of correlations. The main difficulty with a standard IV approach is that each quantity measured is equally valid as a left- or right-hand-side variable. Moreover, for a given right-hand-side variable, either observation could be the instrument or the instrumented variable. In essence, ORIV stacks all four possible IV regressions in order to maximize the information in the estimate, and then applies adjustments to get from a regression coefficient to a correlation and to ensure that standard errors are calculated efficiently.

Formally, with two measures of WTA (WTA^*) measured with error ($\text{WTA}^a = \text{WTA}^* + \eta^a$ and $\text{WTA}^b = \text{WTA}^* + \eta^b$ with $\mathbb{E}[\eta^a] = \mathbb{E}\left[\eta^b\right] = \mathbb{E}\left[\eta^a \eta^b\right] = 0,$), and two of WTP (WTP^*) measured with error ($\text{WTP}^a = \text{WTP}^* + \nu^a$ and $\text{WTP}^b = \text{WTP}^* + \nu^b$ with $\mathbb{E}[\nu^a] = \mathbb{E}\left[\nu^b\right] =$

$\mathbb{E}\left[\nu^a \nu^b\right] = 0$), ORIV stacks the values, and estimates:

$$\begin{pmatrix} \text{WTA}^a \\ \text{WTA}^a \\ \text{WTA}^b \\ \text{WTA}^b \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} + \beta \begin{pmatrix} \text{WTP}^a \\ \text{WTP}^b \\ \text{WTP}^a \\ \text{WTP}^b \end{pmatrix} + \varepsilon$$

$$\text{with instruments} \quad W = \begin{pmatrix} \text{WTP}^b & 0_N & 0_N & 0_N \\ 0_N & \text{WTP}^a & 0_N & 0_N \\ 0_N & 0_N & \text{WTP}^b & 0_N \\ 0_N & 0_N & 0_N & \text{WTP}^a \end{pmatrix}.$$

This yields regression estimate β^*. The regression coefficient is then transformed into a correlation using

$$\text{Corr}[\text{WTA}^*, \text{WTP}^*] = \hat{\beta}^* \sqrt{\frac{\widehat{\text{Cov}}[\text{WTP}^a, \text{WTP}^b]}{\widehat{\text{Cov}}[\text{WTA}^a, \text{WTA}^b]}}.$$

Gillen et al. (Forthcoming) shows that this produces consistent estimates of $\text{Corr}[\text{WTA}^*, \text{WTP}^*]$.

We use a similar stacking of alternative measures in Section 7 when regressing WTA, WTP, and the endowment effect on loss aversion, controlling for risk aversion. As each of the three measures we use as controls has multiple elicitations, we create a structure that allows each measure to act as the instrument and the instrumented value, maximizing the efficiency of the estimates.

3 Results

We begin by showing the limited correlation between WTA and WTP in our data, in Table 2. Two patterns are worth noting. First, WTA and WTP are not correlated. If there is a relationship, it is negative, although small. That is, observing a high willingness to pay for a lottery ticket conveys very little information about willingness to accept. Second, the

Table 2: Correlations

	Correlation between WTA and WTP				Correlation within Type	
	Lottery 1	Lottery 2	Averages	ORIV	WTA	WTP
Study 1, Wave 1	-0.06^{*} (.037)	-0.06^{*} (.037)	-0.08^{**} (.037)	-0.09^{**} (.044)	0.71^{***} (.023)	0.74^{***} (.029)
Study 1, Wave 2	-0.01 (.050)	-0.02 (.049)	-0.02 (.053)	-0.02 (.064)	0.67^{***} (.033)	0.79^{***} (.022)
Study 2	-0.09^{*} (.050)	-0.06 (.055)	-0.09 (.057)	-0.11 (.068)	0.70^{***} (.034)	0.75^{***} (.044)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

two measures of WTA are highly correlated with each other. This is also true of the two measures of WTP. This suggests that these questions are measuring something, just not the same thing. Both of these patterns exist in all three of our surveys.

The middle two columns of Table 2 take steps to reduce concerns about the role of measurement error in our results. First, we standardize the two lottery measures of WTA, then average them, and do the same for WTP.[13] The correlations between the two resultant measures (in the Averages column) grow slightly more negative, consistent with a reduction in the attenuation bias.

Next, we implement the ORIV procedure described in Section 2.2. ORIV produces consistent estimates of correlations, reducing attenuation even further (to zero, in large samples). Once again, the negative correlations grow in magnitude. However, the statistical significance decreases, as ORIV, like all IV techniques, also increases standard errors.

[13]Standardizing these measures implies CRRA utility. This is a reasonable assumption given the small differences in stakes between lotteries.

Table 3: The endowment effect exists for a majority of participants in our data.

	Lottery 1			Lottery 2		
	WTA<WTP	WTA=WTP	WTA>WTP	WTA<WTP	WTA=WTP	WTA>WTP
Study 1, Wave 1	32%	12%	57%	25%	15%	61%
Study 1, Wave 2	30%	13%	57%	24%	16%	61%
Study 2	28%	12%	59%	24%	14%	62%

3.1 The Endowment Effect

Our data exhibit an endowment effect (WTA > WTP), consistent with existing research. Across the studies, the average WTA was 90% of the expected value of the lottery, with a median of 88%, and standard deviation of 37%. The mean WTP is statistically significantly lower at 66%, with a median of 63%, and standard deviation of 33%. The WTP also has a higher proportion of participants reporting extremely low values—38% of participants reported a WTP of less than half of the expected value, compared to 13% for WTA. The prevalence of such low values for WTP may concern some readers. Thus, in Section 3.3 and Section 4, we examine various ways to remove outliers.

As shown in Table 3, the distribution of the endowment effect is consistent across surveys. Approximately 60% of participants exhibit an endowment effect, and 10% exhibit no endowment effect. The final 30% of participants exhibit a negative endowment effect. This can be reconciled with prospect theory if participants are loss loving—which is not uncommon in individual-level studies (see, for example, Sokol-Hessner et al., 2009).[14] These figures are in line with the few previous within-subject studies, see Section 5.

The existence of a negative endowment effect is not likely to be due solely to measurement error or random utility. The correlation between the two different measures of the endowment effect (corresponding to each of the two lottery tickets) is 0.75. This correlation is relatively similar for those with positive and negative endowment effects on a single lottery. To put this another way, of those with a negative endowment effect on a single lottery, there is a 65–75% chance they have a negative endowment effect on the other lottery. The range depends on the study examined, and how those choosing dominated options are treated. By comparison, those with a positive endowment effect for one lottery ticket had a 70–80% chance of having a positive endowment effect for the other. Thus, there seems to be only a bit more noise among participants with a negative endowment effect.[15]

[14] Other studies have found consistent negative endowment effects when physical proximity of a good is counterintuitive: that is, buyers have objects in front of them and sellers do not (Knetsch and Wong, 2009).

[15] Note that the MPL elicitation method does not allow us to distinguish between those with small positive,

3.2 Stability Across Time

A final unique feature of our data is that one of our studies (Study 1) was administered twice to the same participants, six months apart. This allows us to examine the stability over time of WTA and WTP, as well as any of our other measures. As noted in Section 2, only 1,465 of the original participants were successfully resurveyed. Using sampling weights allows for representative statistics. However, we cannot rule out the possibility that unobserved factors correlated with WTA and WTP are driving attrition. Thus, it is important to compare the stability of WTA and WTP to the stability of other attitudes.

WTA for the two lotteries are correlated 0.25 and 0.22 across waves. For WTP, the correlations are 0.29 and 0.26.[16] Using ORIV produces correlation estimates of 0.33 for WTA, and 0.34 for WTP. All of these correlations are statistically significantly different from zero, but are not significantly different from each other.

The stability of WTA and WTP is comparable to the stability of other measures of risk preferences. Study 1 contained three other measures of risk aversion. Two were elicited twice, and have ORIV correlations of 0.37 and 0.33. A third was measured only once per wave, and has a correlation of 0.33 across time. This is similar to the stability of measures of risk preferences over approximately the same amount of time in Caltech undergrads (Gillen et al., Forthcoming).[17] Study 2 also contains two measures of ambiguity aversion (ORIV correlation 0.21 within-participant, across time), time preferences (ORIV correlation 0.33), and amount given in the dictator game (ORIV correlation 0.50).

Three other patterns deserve mention. First, the correlations between WTA in one wave and WTP in the other are nonexistent. This is, perhaps, unsurprising given the lack of correlation within a single study, and the relative stability of WTA and WTP across time.

small negative, or zero endowment effects. The results in Table 3 assume a valuation that is the midpoint of the certainty equivalents on either side of the MPL switch. Alternative codings produce somewhat different results. The minimum percentage of participants with negative endowment effects is approximately 10 percentage points below the numbers shown in Table 3; that is, 15–22%, depending on the lottery and study.

[16] Recall from Table 1 that both waves employed the same lotteries. Unweighted correlations are 0.24 and 0.23 for WTA, and 0.27 and 0.24 for WTP.

[17] Gillen et al. (Forthcoming) did not elicit WTA and WTP.

Second, the endowment effect has a (within-participant) correlation of 0.30 across time. Third, we can examine the correlation for WTA and WTP on the subset of people who gave the same valuations in both Wave 1 and Wave 2. For the first lottery, this restricts the sample to 101 participants.[18] The correlation between WTA and WTP for these participants is 0.10 ($p = 0.30$) for the first lottery, and 0.01 ($p = 0.90$) and 0.02 ($p = 0.87$) for the second lottery in the first and second waves, respectively. For the second lottery, the sample is 122 participants. The correlation between WTA and WTP for these participants is -0.02 ($p = 0.84$) in the second lottery, and 0.04 ($p = 0.70$) and 0.01 ($p = 0.92$) in the first lottery, in the first and second waves, respectively.

Thus, WTA, WTP, and the endowment effect are as stable as other measures. Once again, this suggests that the WTA and WTP questions are measuring something, just not the same thing.

3.3 Summary of Basic Findings

These findings can be summarized, and extended, in a graph. Figure 1 shows average WTA and WTP for all participants in both studies.[19] We also display a non-parametric fit of the data with 95% confidence intervals. Finally, the 45-degree line separates those exhibiting positive endowment effects from those exhibiting negative ones.

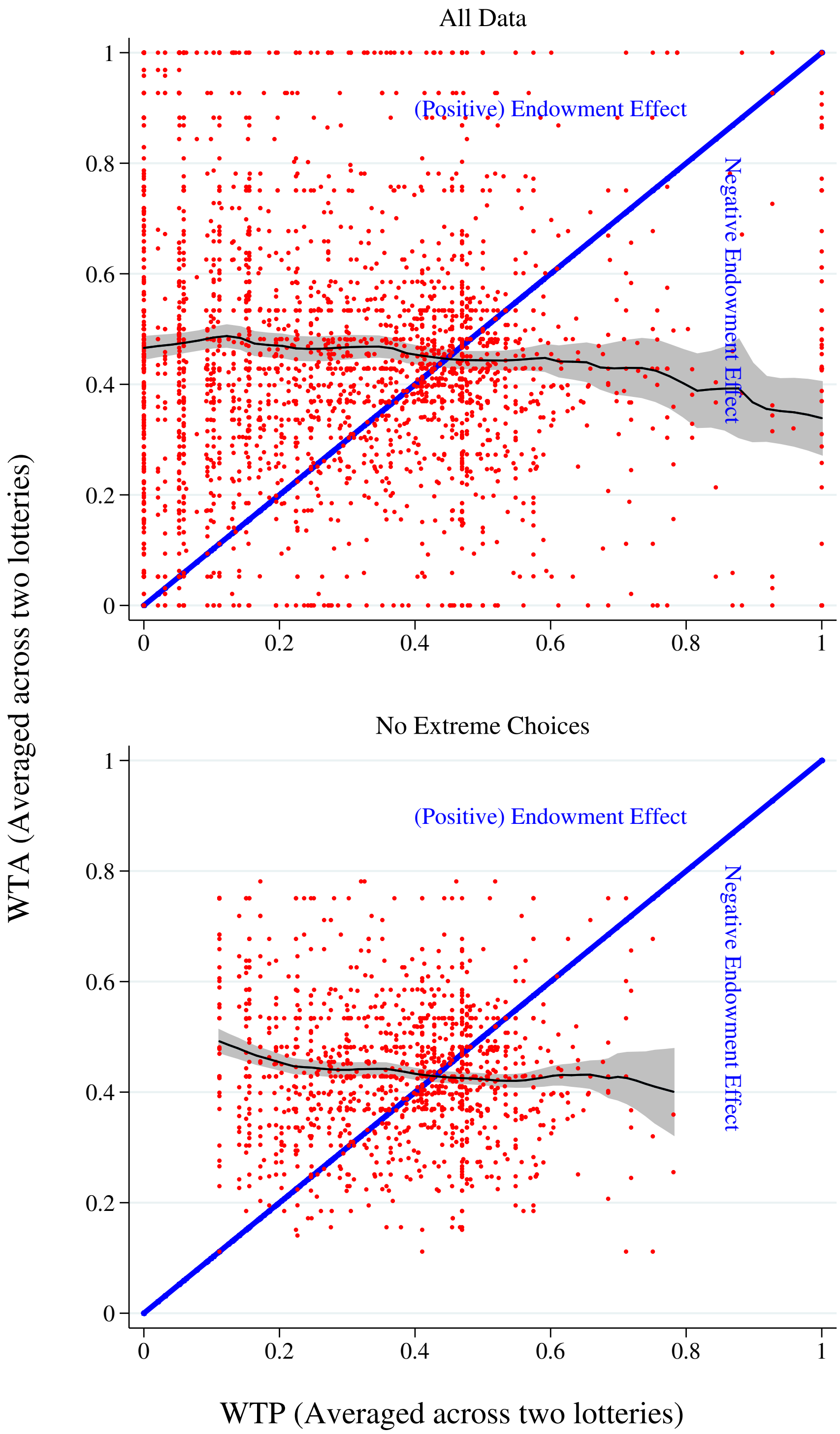
The previously discussed patterns are immediately apparent. WTA and WTP are poorly correlated. A majority of participants exhibit a positive endowment effect. New patterns emerge from the figure as well. First, to the extent that there is any correlation between WTA and WTP, it is due to participants with above risk-neutral values expressed for WTP. As this describes only ∼25% of participants, this negative relationship is weak, as indicated by the expanding standard errors. Second, there is wide variation in both WTA and WTP.

This wide variation includes participants that give extremely high or low values for WTA

[18] Almost no one had the same valuations for WTA and WTP for both lotteries.

[19] This excludes Wave 2 of Study 1, as these are repeated observations of participants from Wave 1. All results are shown separately by survey in Appendix B. A histogram of the endowment effect underlying these figures is shown in Appendix Figure E.1

Figure 1: WTA and WTP are uncorrelated, even after removing dominated choices.



and/or WTP. Thus, in the second panel of Figure 1, we remove all participants that reported either a dominated option, or the next lowest or highest value for any of the four measures of WTA or WTP. The figure displays the same patterns.

The next section considers a number of other subgroups to assess the consistency of our findings across different populations.

4 Subgroups

One may believe that a correlation between WTA and WTP would be restored if we focus on specific subgroups: for example, participants with higher IQ or post-secondary education, or participants who took their time on the survey, indicating that they were paying attention. We examine the correlation between WTA and WTP for all these subgroups, and some others, in Table 4. Correlations are examined by lottery, for the average of both lotteries, and using ORIV. To maximize statistical power, we combine Study 1, Wave 1 and Study 2. This gives us a total of 3,000 independent observations.[20]

Most of the subgroups in the table need no explanation. However, a few do. "Not too fast" removes the 10% of participants who completed the WTA and WTP questions in the least amount of time. Similar to our approach in Figure 1, we show three additional ways of removing "extreme" choices in the table: one removes dominated options, one removes participants whose switching point was just before the first selectable item or just after the last; the third removes participants if their switching point was in the top three or bottom three rows.

In all subgroups, correlations between WTA and WTP are small in magnitude and/or negative. There is a consistent (small) positive correlation, in the range of 0.1–0.2, for those with high CRT or IQ scores. In both studies, we administered the Cognitive Reflection Test of Frederick (2005). Around 50% of participants got one or more of the three questions

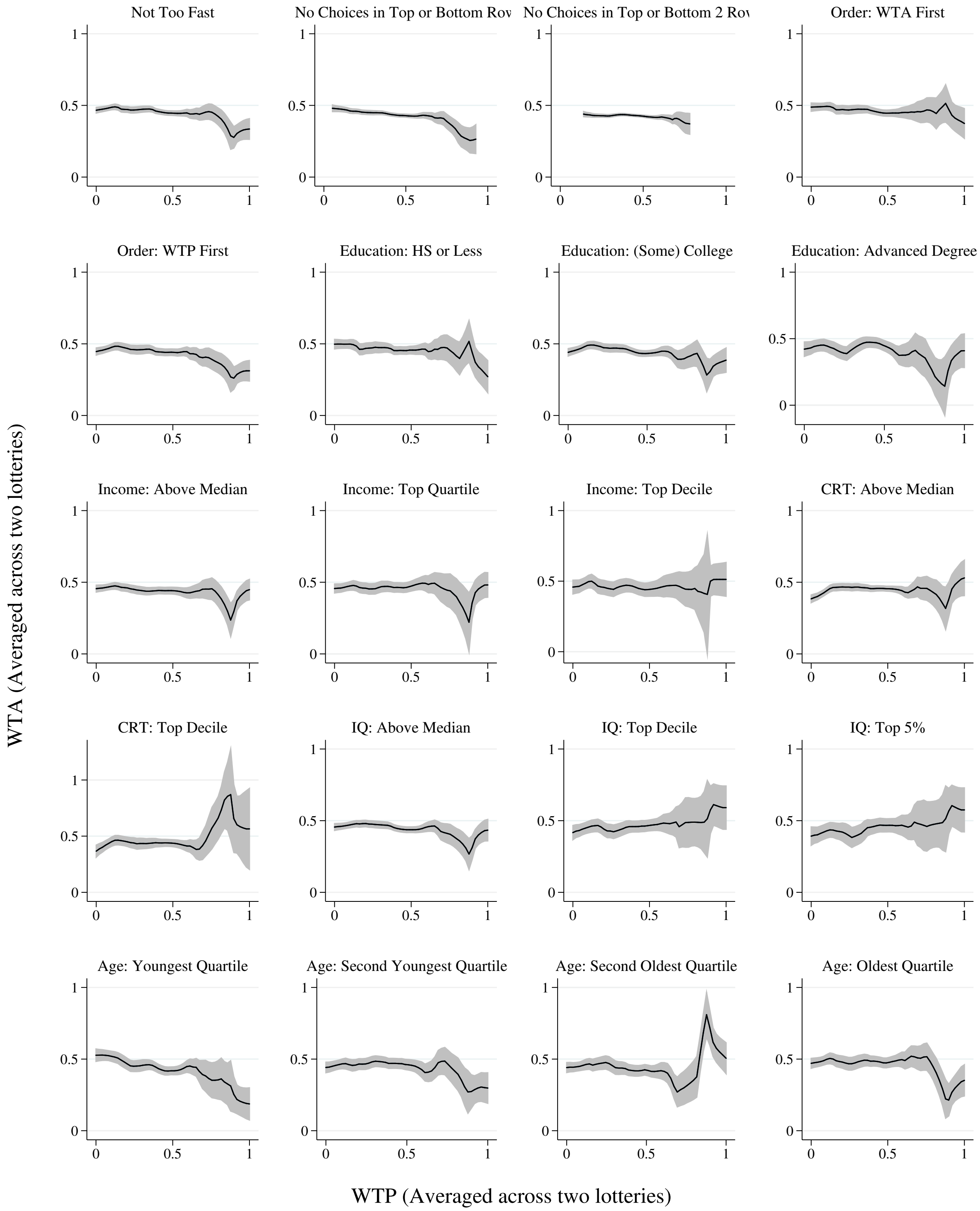
[20] Wave 2 consists of the same participants as Wave 1, hence the observations are not independent. For results by survey, see Appendix B.

Table 4: Correlations for Subgroups. Data from Study 1, Wave 1 and Study 2

	N	Lottery 1	Lottery 2	Average	ORIV
All	3,000	−0.07**	−0.06**	−0.08**	−0.10***
		(.030)	(.031)	(.031)	(.038)
Not Too Fast	2,700	−0.06*	−0.06**	−0.07**	−0.09**
		(.031)	(.031)	(.031)	(.037)
No Dominated Choices	2,672	−0.07**	−0.08***	−0.09***	−0.10***
		(.032)	(.031)	(.032)	(.037)
No Switches in Top or Bottom Two Rows	1,595	−0.11***	−0.14***	−0.14***	−0.19***
		(.043)	(.043)	(.042)	(.051)
No Switches in Top or Bottom Three Rows	1,028	−0.05	−0.05	−0.05	−0.10
		(.058)	(.050)	(.056)	(.079)
Question Order: WTA First	1,501	−0.06	−0.06	−0.07	−0.08
		(.042)	(.046)	(.043)	(.052)
Question Order: WTP First	1,499	−0.08*	−0.07*	−0.09**	−0.12**
		(.043)	(.043)	(.044)	(.052)
Education: HS or Less	1,128	−0.11**	−0.05	−0.11**	−0.13**
		(.049)	(.055)	(.054)	(.066)
Education: Some College	1,538	−0.03	−0.06*	−0.05	−0.07
		(.039)	(.037)	(.039)	(.046)
Education: Advanced Degree	334	−0.01	−0.09	−0.04	−0.07
		(.074)	(.063)	(.071)	(.086)
Income: Above Median	1,520	−0.03	−0.06	−0.04	−0.06
		(.037)	(.041)	(.038)	(.046)
Income: Top Quartile	813	−0.01	−0.01	0.01	0.00
		(.054)	(.058)	(.053)	(.064)
Income: Top Decile	399	0.06	−0.04	0.01	0.01
		(.071)	(.067)	(.073)	(.088)
CRT: Above Median	1,371	0.11***	0.05	0.09**	0.10**
		(.042)	(.041)	(.043)	(.052)
CRT: Top Decile	288	0.10	0.11	0.10	0.11
		(.084)	(.084)	(.085)	(.10)
IQ: Above Median	1,694	−0.03	−0.06*	−0.05	−0.07
		(.038)	(.035)	(.037)	(.043)
IQ: Top Decile	337	0.10	0.07	0.09	0.10
		(.089)	(.075)	(.086)	(.10)
IQ: Top 5%	146	0.28***	0.13	0.20**	0.25*
		(.094)	(.10)	(.098)	(.13)
Age: Youngest Quartile	611	−0.20***	−0.23***	−0.26***	−0.32***
		(.069)	(.062)	(.070)	(.083)
Age: Second Youngest Quartile	798	−0.03	−0.04	−0.03	−0.05
		(.056)	(.056)	(.053)	(.063)
Age: Second Oldest Quartile	783	−0.02	0.04	0.00	0.00
		(.057)	(.067)	(.062)	(.074)
Age: Oldest Quartile	808	−0.04	−0.05	−0.04	−0.07
		(.048)	(.053)	(.049)	(.060)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Figure 2: In most subgroups, WTP provides no information about WTA.



correct, representing all those above the median. The top decile answered all three questions correctly. We also administered an in-study IQ test. This test took a fixed set of six questions from the International Cognitive Ability Resource (ICAR; Condon and Revelle 2014). These six questions were chosen from the matrix and three-dimensional rotation modules such that they were progressively more difficult. Only 5% of participants correctly answered all six questions.[21] The top decile represents five or more correct answers, and above the median is two or more correct answers.

The only robust positive correlation is for those who answered all six of the IQ questions correctly. Here, the correlation goes as high as 0.28, although including the next 5% of participants in terms of IQ reduces the correlation to around 0.1. In Section 5, we will compare our results to those obtained in the lab—that is, we will examine studies conducted on undergraduate students of elite universities. For that comparison, we isolate those that are college educated (2+ years), younger than 45, and have an IQ plus CRT in the top 20% ($N = 163$). In this subgroup, the correlation between WTA and WTP is 0.16.[22]

These subgroup results can be illustrated graphically as well. Figure 2 contains 20 panels in the style of Figure 1, showing a non-parametric fit of WTA to WTP. Once again, to the extent that there is any correlation, positive or negative, between WTA and WTP it is for those whose WTP values indicate risk love. As with Figure 1, this region of the figures contains very few participants, so any apparent pattern is likely spurious.

5 Prior Studies

The prior sections lead to a natural question: Why has the limited correlation between WTA and WTP not been documented before? This is likely due to the fact that almost all the studies in the enormous literature on the endowment effect use between-subject designs. In order to examine the correlation between WTA and WTP, one needs a within-subject

[21] Consistent with this, none of the authors of this study correctly answered all six questions.

[22] For those in the top 50% of IQ + CRT with a college education, the correlation is 0.12 ($N = 730$); for those in the top 20% of IQ + CRT with a college education, the correlation is 0.17 ($N = 399$), and so on.

design. A few studies have collected such data, however we are aware of only one study that examines this correlation: Brown et al. (2017) elicits valuations for two hypothetical annuities, and finds negative correlations between WTA and WTP of −0.11 and −0.15.

We perform a meta-analysis of the three studies that use within-person incentivized measures of WTA and WTP and have data available ($N = 466$).[23] We emphasize that none of the studies we found—whether we were able to obtain their data or not—studied, or even reported, the correlation between WTA and WTP (with the exception of the one un-incentivized study mentioned above). Of these three studies from which we have data, two used the BDM method (Becker et al., 1964) to elicit WTA and WTP for several lotteries. A third used a median-price auction, repeated six times for each of two lotteries, to elicit WTA and WTP. The average correlation between WTA and WTP across all three studies is small, but larger than in our study. This difference is probably driven by different samples—representative versus those that come to the lab—rather than differing methodologies.

Table 5 re-analyzes the data from the two studies which use the BDM method. The data from Isoni et al. (2011) shows little or no correlation between WTA and WTP. On the other hand, the data from Fehr et al. (2015) suggests a correlation around 0.2. It is unclear what is driving these differences. Taking an average—weighted by number of participants—of correlations across these studies indicates a correlation of 0.15.[24]

While this average correlation is somewhat larger than we observe in our data, it is probably smaller than one would expect. It is also in-line with what we observe, in Section

[23] We searched all papers published in top economics journals. We also consulted the comprehensive annotated bibliography by Peter Wakker (http://people.few.eur.nl/wakker/refs/webrfrncs.docx). This yielded nine studies. Tunçel and Hammitt (2014) conducts a similar search and finds five studies with within-subject designs—all of which were also found by our search. Three no longer had data available (Harless, 1989; Eisenberger and Weber, 1995; Borges and Knetsch, 1998). Three others we excluded from the meta-analysis: Schmidt and Traub (2009) and Schmidt and Trautmann (2014) use the same data, which contains 23 participants making choices over 50 lotteries. The range of correlations of WTA and WTP in those lotteries is from −0.67 to 0.86, with an average of 0.19. Most of these correlations are statistically insignificant due to the very small sample size. Plott and Zeiler (2005) measures WTA and WTP for lotteries in training rounds, but does not report this data due to concerns about reliability. Although we have similar concerns we include an analysis of this data in Appendix C for completeness.

[24] Dropping dominated choices, or replacing them with undominated options, results in similar overall patterns, although the value of particular correlations changes, sometimes substantially.

Table 5: The correlation between WTA and WTP for lotteries over gains is limited in prior studies.

Study	Group (N)	Lottery	Correlation	WTA < WTP
Isoni et al. (2011)	1 (100)	$0.3 * 1 \oplus 0.7 * 4$	0.01 (.10)	16%
		$0.5 * 1.5 \oplus 0.5 * 3.5$	0.03 (.10)	37%
		$0.6 * 1 \oplus 0.4 * 3$	0.20** (.10)	21%
		$0.7 * 0.1 \oplus 0.3 * 0.8$	0.03 (.10)	26%
		$0.7 * 1 \oplus 0.3 * 5$	0.10 (.10)	31%
Fehr et al. (2015)	1 (95)	$0.3 * 1 \oplus 0.7 * 8$	0.15 (.10)	25%
		$0.5 * 1 \oplus 0.5 * 1.5$	0.26** (.10)	35%
		$0.5 * -3 \oplus 0.5 * 9$	0.35*** (.10)	19%
		$0.6 * 1 \oplus 0.4 * 6$	0.20* (.10)	24%
		$0.7 * -0.1 \oplus 0.3 * 0.8$	0.22** (.10)	33%
		$0.7 * 1 \oplus 0.3 * 11$	0.11 (.10)	32%
	2 (96)	$0.5 * 1 \oplus 0.5 * 1.5$	0.15 (.10)	28%

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Correlations with standard errors in parentheses.

4, from participants who are most like those in lab studies. Moreover, the percentage of participants expressing a negative endowment effect—around 25%—is quite similar to what we observe in our data (see Table 3). Together with the fact that these two studies use a different elicitation method than we do, this suggests the difference in results is due to our use of a representative participant pool.

Table 6: Correlations from repeated measurements of WTA and WTP in (Loomes et al., 2003, $N = 175$).

	Lottery 1 $0.2 * 0 \oplus 0.8 * 12$		Lottery 2 $0.8 * 0 \oplus 0.2 * 12$	
	Correlation	WTA < WTP	Correlation	WTA < WTP
Round 1	0.18^{**} (.075)	38%	0.20^{***} (.075)	32%
Round 2	0.22^{***} (.074)	34%	0.14^{*} (.075)	30%
Round 3	0.28^{***} (.073)	32%	0.28^{***} (.073)	33%
Round 4	0.27^{***} (.073)	36%	0.26^{***} (.073)	38%
Round 5	0.36^{***} (.071)	33%	0.15^{**} (.075)	39%
Round 6	0.23^{***} (.074)	37%	0.13^{*} (.075)	38%

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Correlations with standard errors in parentheses.

Next, we examine the single study that repeatedly elicited WTA and WTP (Loomes et al., 2003). In this study, WTA and WTP for two lotteries were elicited from $N = 175$ participants using a median-price auction. This was done six times, and after each round the price of each lottery was posted. This repetition was used to study whether the endowment effect, and other biases, were reduced by repeated play and learning. The correlations for these two lotteries and negative endowment effect, round by round, are shown in Table 6. Although there is some variation in the estimated correlations, no two correlations for the same lottery are statistically different from each other, or from the average correlation across both lotteries and all rounds of 0.23. This suggests, again, that our results are not due to the elicitation method we use. Moreover, the fact that results are fairly stable from round to round suggests that these findings are unlikely to be due to "mistakes."

A final (unpublished) study worth noting elicited WTA and WTP from a non-representative sample of the British population ($N = 1{,}990$; Carvalho et al., 2017). This study consisted of a subset of the authors here, and used the same setup and lotteries as Study 2. The correlation for both lotteries was 0.10 (s.e. of .027 for both lotteries). This reinforces the contention that the particular (small) correlation observed in different studies is due to the participants, rather than elicitation method. Moreover, across all these different participant pools, the correlation between WTA and WTP is much, much less than 1—and probably much less than one would expect.

6 Relation with Measures of Risk Aversion

WTA and WTP for lottery tickets are relatively stable and largely independent in all the data we have considered. We now investigate their relationships with other measures of risk attitudes from Study 2.[25] Each of the risk aversion measures is highly correlated with WTA or WTP, but rarely both. Moreover, the measures correlated with WTA also tend to be highly correlated with each other, but not with those correlated with WTP, and vice-versa. This reinforces the view that WTA and WTP are capturing something, but not the same thing. Following Hershey and Schoemaker (1985) and Sprenger (2015), a plausible reading of these results is that how questions are asked may implicitly determine a reference point.

The risk attitude measures in Study 2 fall into two broad categories: one where the MPL features a fixed lottery, and one where it features a variable lottery. The former category contains four measures that vary in terms in the domain—gains, losses, or both—over which risk is measured, or in the way that randomization in the lottery is conducted—with an urn or with a lottery ticket:[26]

Urn: This measures the certainty equivalent for a draw from an urn with an equal number of two colors of balls, one representing a large payoff, and the other a zero payoff.

[25] As we are examining WTA and WTP for lotteries, these might also be thought of as measures of risk attitudes.

[26] Screenshots and details of the study design can be found at hss.caltech.edu/∼snowberg/wep.html. All questions were elicited twice with different lotteries in order to control for measurement error.

Gain: A lottery where one payoff was a small gain (or zero), and the other a large gain.

Mixed: A lottery where one payoff was a moderate gain, and one was a moderate loss. Participants could choose between these lotteries and sure losses/sure gains.

Loss: A lottery where one payoff was a small loss (or zero), and the other was a large loss. Participants could choose between these lotteries and sure losses.

The second category of measures includes those with a variable lottery. They are:

FM: The monetary amount is fixed (FM = fixed money), and the participants choose which lotteries (with fixed probabilities but a variable prize) they prefer over this amount.

2L: There is a fixed lottery, and the participants choose which of the variable lotteries (with fixed probabilities) they prefer over this fixed lottery (2L = two lotteries).

We also administered a qualitative self-assessment of risk attitudes, from Falk et al. (2013). The question reads: "How do you see yourself: are you a person who is generally willing to take risks or do you try to avoid taking risks?", followed by clickable horizontal boxes with the numbers 0 through 10 in them. To the left of 0, the text reads, "Completely unwilling to take risks," and to the right of 10 the text reads,"Very willing to take risks." All measures of risk aversion are (re-)coded so that higher values correspond to more risk aversion. Thus, the expected (and usually observed) sign of the correlation between WTA or WTP and these measures is negative.

An obvious pattern jumps out from the ORIV correlations between these measures, displayed in Table 7: there are two clusters of risk attitudes. These clusters feature large within-cluster correlations, and smaller correlations with measures in the other cluster.[27]

[27] The fact that theoretically equivalent measures of risk attitudes produce poorly correlated responses is consistent with a large literature. For recent results from the lab or field, see Berg et al. (2005); Bruner (2009); Crosetto and Filippin (2016); Deck et al. (2010, 2013); Harbaugh et al. (2010); He et al. (2017); Isaac and James (2000); Loomes and Pogrebna (2014); Lönnqvist et al. (2015); Nielsen et al. (2013). See Weber and Johnson (2008) for a summary of the psychology literature on this topic. Gillen et al. (Forthcoming) suggests that findings of low correlations between measures of risk attitudes may be due to measurement error—we adopted its techniques to rule this out.

Table 7: ORIV Correlations between WTA/P and Risk Measures, Study 2

				Fixed Lottery				Variable Lottery	
		WTA	Urn	Gain	Mixed	Loss	WTP	FM	2L
Fixed Lottery	Urn	-0.66^{***}					0.07		
		(.042)					(.067)		
	Gain	-0.66^{***}	0.65^{***}				0.04		
		(.051)	(.051)				(.070)		
	Mixed	-0.58^{***}	0.51^{***}	0.60^{***}			0.19^{***}		
		(.054)	(.053)	(.049)			(.069)		
	Loss	-0.27^{***}	0.26^{***}	0.39^{***}	0.65^{***}		0.30^{***}		
		(.076)	(.067)	(.070)	(.056)		(.077)		
Variable Lottery	FM	-0.03	0.05	0.09	-0.14^{*}	-0.19^{***}	-0.45^{***}		
		(.070)	(.066)	(.069)	(.069)	(.075)	(.041)		
	2L	0.12^{*}	-0.17^{***}	-0.13^{*}	-0.21^{***}	-0.15^{*}	-0.28^{***}	0.41^{***}	
		(.071)	(.063)	(.071)	(.073)	(.078)	(.061)	(.061)	
Qualitative		-0.24^{***}	0.18^{***}	0.18^{***}	0.17^{***}	-0.05	-0.15^{***}	0.15^{***}	0.13^{*}
		(.062)	(.058)	(.077)	(.070)	(.089)	(.064)	(.062)	(.065)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses. All measures except WTA and WTP are (re-)coded so that higher values correspond to more risk aversion.

Particularly notable is that WTA and WTP are part of different clusters. Loosely speaking, measures in the fixed lottery cluster are all fairly highly correlated with each other, as are the two variable lottery measures. The fixed lottery measures are generally highly correlated with WTA, but not WTP.[28] The variable lottery measures show the opposite pattern. The only measure that appears to be (somewhat) correlated across the board is the qualitative measure. Clearly, WTA and WTP are measuring something, just not the same thing.

It is well known that subtle aspects of how questions are structured, or the the set of choices, can create implicit frames that change choices. The fixed element of the MPL may implicitly become the reference point, an interpretation used in the influential works of Hershey and Schoemaker (1985) and Sprenger (2015). Implicit framing can explain the patterns above: WTA's cluster contains other measures where the lottery option is fixed (on the left-hand side of the MPL) and participants are asked for their certainty equivalent—implicitly, to price the lottery. In the FM question, the fixed option is an amount of money, and this measure is, coherently, highly correlated with WTP.[29] Thus, our findings replicate the differences in risk aversion measures of Sprenger (2015)—although he examined differences, not correlations, between probability and certainty equivalents. Moreover, they provide further support for Sprenger's interpretation, as we show a strong correlation between measures with an assumed implicit frame and corresponding measures with an explicit framing (WTA and WTP).

Using implicit framing provides an alternative way to examine the correlation between WTA and WTP. This relies on a different hypothesis about reference point formation that does not involve an explicit mention of buying or selling. If we assume, in line with Sprenger (2015), that participants consider themselves endowed with a monetary amount in the case of FM, and with a lottery in the case of Urn, Gain, Mixed, and Loss, we again find evidence

[28]A principal components analysis confirms these clusters, and suggests relationships with a broad range of other preference measures, see Camerer et al. (2019b).

[29]However, this reading is not perfect: the 2L measure contains both a fixed and a variable lottery option, and resides in the WTP cluster. Moreover, although risk aversion over losses is firmly in WTA's cluster, it is equally correlated with WTP, but with the opposite sign.

for a zero or negative correlation between WTP and WTA: FM is not significantly correlated with the first two measures, and is significantly negatively correlated with the latter two. Moreover, the endowment effect(s) as measured with implicit and explicit framing are correlated 0.54.

These results suggest a possible explanation for the small correlation between WTA and WTP: buying and selling are different processes. Before we return to this idea in Section 8, we next turn to discuss the implications of our findings for theories of reference dependence.

7 Relation to Theories of Reference Dependence

This section relates our data to models of reference-dependent preferences, using the taxonomy of Sprenger (2015). We first discuss theories that are consistent with at least one of the broad features of reference dependence—namely, the endowment effect and the change in risk preferences between lotteries featuring only gains, only losses, and both gains and losses. These patterns are found in our data, and many, many other studies.[30] Applied mechanically, theories of reference dependence have difficulty capturing both behaviors. To explain the former, a context-dependent reference point—which changes with the framing, as in WTA and WTP—is necessary, while, to explain the latter, a context-independent reference point—which determines differential responses to gains and losses within a question, but does not change across questions—is necessary. We therefore also consider hybrid variants of these theories which can capture both phenomena by allowing flexibility in the reference point both within and across decision problems.

Although hybrid theories are compatible with our main finding, they do not pass more stringent tests utilizing data from choices over the other lotteries. In particular, hybrids predict correlations between loss aversion—measured by an increase in risk aversion for lotteries involving gains and losses (relative to those involving only gains and only losses)—

[30] In our studies, participants are, on average, more risk loving over losses than gains. They show similar average levels of risk aversion in the mixed gain/loss domain as in the gain domain. Camerer et al. (2019b) provides details.

and the endowment effect. Using two different ways of measuring loss aversion, we find no support for this or other, subtler, predictions, and some evidence against them.

Throughout the theoretical development that follows, we focus our attention on lotteries like those used in our WTA and WTP questions: they have a high h and low l payoff, each equally likely.

7.1 Context-Independent Reference Points

Theories of reference dependence in which the reference point does not respond to framing are unable to explain the endowment effect. Thus, a zero correlation between WTA and WTP under such a theory can only come about due if all the variation in these choices is noise, which is clearly not the case given evidence in prior sections. This category of models includes the "classic" version of Cumulative Prospect Theory (CPT—see, for example, Tversky and Kahneman, 1992; Wu and Gonzalez, 1996; Prelec, 1998; Abdellaoui, 2000), in which probabilities are weighted according to a function π and utilities are calculated relative to a fixed reference value r (usually 0). The value of a lottery is then

$$V_{CPT}(l,h) = \pi(0.5)u_{CPT}(h|r) + (1-\pi(0.5))\,u_{CPT}(l|r)$$

A standard formulation, from Kahneman and Tversky (1979), for u_{CPT} is $v(x-r)$ for $(x-r)>0$ and $-\lambda v(r-x)$ for $(x-r)<0$, although others have been proposed. $\lambda>1$ captures loss aversion, and v is a concave function, which captures diminishing sensitivity to both gains and losses.

As the value of r is fixed, this model cannot explain different behavior in the WTP and WTA task. This is related to the reason why the standard model cannot accommodate an endowment effect: utility is relative to current wealth, and, in CPT, current wealth is effectively replaced by the reference point r.[31] This model, however, is designed to explain

[31] Note that there are two protocols that could be used to apply this model to the WTP question, but both lead to the same conclusion. Assuming that v is linear (a standard simplification), and $r=0$, the WTA for

differences in risk attitudes when faced with lotteries containing only gains, only losses, or both gains and losses: diminishing sensitivity leads to risk aversion for gains and risk love for losses, while loss aversion increases risk aversion for lotteries over gains and losses.

Models in which the reference point is based solely on expectations are also unable to explain the endowment effect in our data. This includes a well-known strand of work that extends CPT to allow the reference point to be equal to the expected value of the lottery (Bell, 1985; Loomes and Sugden, 1986): as the latter is the same in both WTA and WTP, the reference point is once again the same. This is also true of more general models in which the reference point is based on expectations which are derived from the history of lotteries experienced by a participant: due to the randomized question ordering, the history of participants' experience is the same for both types of questions, in expectation.[32]

As is known, both the unacclimating personal equilibrium and the personal preferred equilibrium versions of the model of Kőszegi and Rabin (KR; 2006) are unable to explain the WTA–WTP gap.[33] The general KR model assumes that a lottery with distribution F over a (finite) prize space X will yield a value

$$\begin{aligned} V_{KR}(F|G) &= \sum_{x\in X}\sum_{r\in X} u(x|r)F(x)G(r) \\ u(x|r) &= u(x) + \mu(u(x) - u(r)) \end{aligned}$$

the lottery is given by

$$\text{WTA}_{CPT} = V_{CPT}(l,h) = \pi(0.5)h.$$

The first way of determining WTP is as the point at which the value of the lottery plus remaining payment is equal to keeping h (recall that subjects in such questions were endowed with h points). In this case,

$$h = (h - \text{WTP}_{CPT}) + V_{CPT}(l,h) \quad\Rightarrow\quad \text{WTP}_{CPT} = V_{CPT}(l,h).$$

Alternatively, one may calculate WTP by imagining that the DM calculates the value of the lottery plus the additional side payment holistically, in which case

$$h = \pi(0.5)(2h - \text{WTP}_{CPT}) + (1 - \pi(0.5))\,(h - \text{WTP}_{CPT}) \quad\Rightarrow\quad \text{WTP}_{CPT} = \pi(0.5)h = V_{CPT}(l,h).$$

Note that all three formulations are the same.

[32]Furthermore, as shown in Table 4, there is no difference in behavior between participants who saw the WTP question first and those who saw the WTA question first. Moreover, as noted in Footnote 6, the average WTA, WTP, and endowment effect are the same regardless of whether WTA was randomly selected to appear before WTP, or WTP before WTAThis suggests that the intervening questions do not have much impact.

[33]As discussed below, other variants of the KR model are compatible with the endowment effect.

when compared to a (potentially stochastic) reference G. Here, $u(x)$ is the consumption utility and μ is the gain/loss utility. We follow others, including examples in KR, in setting the form $\mu(z) = z$ for $z > 0$ and $\mu(z) = \lambda z$ for < 0 and $\lambda > 1$.

The predictions of the KR model depend on assumptions about the reference point. In the unacclimating personal equilibrium version of this model, rational expectations are assumed, so an alternative can potentially be chosen if it maximizes utility when it is itself the reference point. The personal preferred equilibrium refinement assumes that the unacclimating personal equilibrium with the highest utility will be chosen. Since, in these models, the WTA/WTP framing does nothing to alter the set of admissible reference points, there will be no difference between WTA and WTP.

7.2 Context-Dependent Reference Points

The endowment effect is usually captured through models in which the reference point is affected by framing. In the general setup of the KR model, this is captured by the "first focus" idea discussed in Kőszegi and Rabin (2006). Formally, denote by $EU := (u(l) + u(h))/2$— the expected utility of a lottery without reference dependence. Note that, in the WTA frame, keeping a lottery has value $EU + (0/2 + (u(h) - u(l))/2) + (\lambda(u(l) - u(h))/2 + 0/2)$, whereas selling it for $x \in (l, h)$ yields $u(x) + \frac{1}{2}(u(x) - u(l)) + \frac{1}{2}(\lambda(u(x) - u(h)))$. Setting these terms equal, we obtain:

$$\text{WTA}_{KR} = u^{-1}(EU)$$

which is decreasing in the curvature of u.

WTP_{KR} depends on the how much money the participants starts with m. The utility of not buying the ticket is $u(m)$. The utility for buying the lottery for $y \in (l, h)$ is $(u(m - y + h) + u(m - y + l))/2 + (u(m - y + h) - u(m))/2 + \lambda(u(m - y + l) - u(m))/2$. Assuming

$m = h$, as in our experiment, and setting these expressions equal implies:

$$\begin{aligned} u(h)(3+\lambda) &= 2u(2h - \text{WTP}_{KR}) + u(h + l - \text{WTP}_{KR})(1+\lambda) \\ \Rightarrow \frac{d\text{WTP}_{KR}}{d\lambda} &= \frac{u(h + l - \text{WTP}_{KR}) - u(h)}{2u'(2h - \text{WTP}_{KR}) + u'(h + l - \text{WTP}_{KR})}. \end{aligned}$$

As $h > \text{WTP}_{KR} > l$, WTP_{KR} is decreasing in loss aversion λ. Standard intuitions about risk aversion give that WTP_{KR} is decreasing in the curvature of u. Importantly, this version of the KR model can accommodate the endowment effect.

A common assumption is linear utility $u(x) = x$, with risk aversion arising from the loss aversion parameter (see, for example, Kőszegi and Rabin, 2007). With these assumptions, variance in the loss aversion parameter would affect WTP, but WTA would be just the expected value of the lottery. Thus, any variation in WTA would be due to a random component of choice. This is consistent with a zero correlation between WTA and WTP, but it is inconsistent with the fact that WTA is correlated with other measures of risk aversion.

Adding variation in risk aversion drives a positive correlation between WTA and WTP. This can be attenuated if the variation in loss aversion is much larger, and independent of variation in risk aversion. This would imply that most of the variation in WTA (which includes loss aversion) would be random with respect to risk aversion or WTP (which does not contain loss aversion), leading to a very small correlation. Allowing correlation between risk and loss aversion can lead to a zero correlation, whatever the variance in the two measures, provided that the correlation between measures is precisely calibrated. Taking advantage of this degree of freedom allows the KR model to fit many of the moments in our data quite well. In particular, adding the assumption that the fixed element of the MPL acts as a reference point (as in Sprenger, 2015, and reviewed in the prior section), this model can also explain the patterns in the prior section. As Sprenger's approach allows any theory compatible with an endowment effect for risky prospects to also be compatible with the findings in the prior section, it is a useful addition to any model.

The general KR model, however, cannot explain a change in risk attitudes between lotteries with only gains, only losses, or both gains and losses. The key reason is that under this model there is no first order impact of moving from the gain to the loss domain (relative to the fixed reference point of 0), as opposed to CPT. In KR, one can consider these lotteries to be neutrally framed (so there is no reference point), or framed similarly to the WTA question. In either case, the certainty equivalent is just the inverse of the utility of the expected value. Changes in risk aversion between gains and losses would therefore have to come from changes in the curvature of the utility function. It seems unlikely for these changes to be large across the relatively small changes in wealth represented by the different types of lotteries.[34]

There exist versions of prospect theory that suffer from essentially the same problem as KR, such as that described in Schmidt et al. (2008). Either the reference point is assumed to be fixed at zero, in which case the endowment effect cannot be captured, or it is set according to the buying or selling task, in which case it is hard to capture changes in risk aversion when considering gains or losses.

7.3 Hybrids

To capture both the endowment effect and different risk attitudes towards lotteries with gains, losses, or gains and losses, a hybrid approach is needed. Here, we explore what seem to be the two most obvious ones. The first hybrid approach assumes a KR-type model in which the reference point is either set by the context (in the WTA and WTP questions) or to zero if no explicit reference is set (as in questions, including ours, that are typically used to measure differences in risk attitudes between gains and losses). The reference point of zero yields behavior very similar to CPT, with comparative statics on the endowment effect given

[34]Moreover, if changes did occur, it would be surprising if they did so in such a way as to make people *more* risk loving at lower utility levels—for example, constant relative risk aversion implies absolute risk aversion that is higher at lower wealth levels.

in the previous subsection. This matches all moments of our data that we have discussed up to this point, but also results in testable hypotheses about the relationship between the endowment effect and loss aversion.

A second hybrid approach is a model in the style of CPT, but in which a two-stage procedure is used: first, a participant uses CPT to determine the value of the lottery, then acts as if they are endowed with that lottery (in the same way they would if they were endowed with a mug) in the selling task. This model has two variants which depend on whether or not the monetary outlay to buy the lottery ticket is encoded as a loss. Following the standard treatment, in which the outlay is not treated as a loss (Tversky and Kahneman, 1991, p. 1055), yields:

$$\mathrm{WTA}_{HCPT} = u^{-1}(\lambda V_{CPT}(l,h))\,, \quad \text{and} \quad \mathrm{WTP}_{HCPT} = u^{-1}(V_{CPT}(l,h))$$

(HCPT stands for Hybrid CPT). Alternatively (A-HCPT), it could be encoded as a loss (Kahneman et al., 1990; Tversky and Kahneman, 1991). In this case, $\mathrm{WTA}_{A-HCPT} = \mathrm{WTA}_{HCPT}$, but willingness to pay is now given by

$$\mathrm{WTP}_{A-HCPT} = u^{-1}(EU/\lambda)\,.$$

Both variants allow for an endowment effect. The analysis of the correlation between WTA and WTP under HCPT is similar to the KR model.

A-HCPT is different in that it can lead to a zero correlation between WTA and WTP without correlation between risk and loss aversion. Risk aversion induces a positive correlation, whereas loss aversion induces a negative correlation. As such, the variation of loss aversion can be precisely calibrated to offset the correlation induced by the variation in risk aversion, yielding a zero correlation.

7.4 Tests of Hybrid Theories

Hybrid models are compatible with the moments of our data we have explored up until this point, and also make additional testable predictions about the relationship between loss aversion—as measured by the increase in risk aversion for mixed-risk lotteries—WTA, WTP, and the endowment effect. These predictions are given by the equations above, and summarized in the first panel of Table 8.[35] Importantly, all three models predict a positive correlation between the endowment effect and loss aversion (after controlling for risk aversion).

Our data provide two ways to test these predictions, as each study used a different method to elicit loss aversion in risky choice. In Study 1, we used Dynamically Optimized Sequential Experimentation (DOSE, Chapman et al., 2019) to elicit risk- and loss-aversion parameters. This method optimally selects a series of questions for each participant in order to maximize the information about a CRRA-risk parameter and a loss-aversion parameter. Each of these questions is a binary choice between a lottery and a sure amount, and is selected based on past answers, in order to optimally identify the parameters of this model with only 20 questions. The estimated parameters from this method are more stable across time (between Wave 1 and Wave 2) than WTA, WTP, risk aversion, ambiguity aversion, and so on.[36]

DOSE-estimated parameters are used to test theoretical predictions about WTA, WTP, and the endowment effect in Panel B of Table 8. WTA and WTP are significantly negatively related to DOSE-estimated risk aversion, as predicted by all three theories. However, counter

[35] We are aware of two unpublished studies that attempt to study the relationship between the endowment effect and loss aversion. Gächter et al. (2007) relates the endowment effect to a measure similar to our "Mixed" measure. Its results are similar to the unconditional correlation between that measure and the endowment effect in our data, and we find that the endowment effect is unrelated to loss aversion. Dean and Ortoleva (2016) finds a correlation between loss aversion and the endowment effect in lab data from an elite university. However, its approach differs from ours in a number of ways: they measure loss aversion parametrically, without accounting for measurement error, and their loss aversion questions are framed in terms of WTA, which may have important effects.

[36] The cross-time correlation in the risk-aversion parameter elicited via DOSE is 0.45, and the cross-time correlation in loss aversion is 0.40. This compares favorably with the cross-time correlations for the other risk preference measures mentioned in Section 3.2.

Table 8: Relationships between WTA, WTP, and the Endowment Effect, and loss and risk aversion.

Dependent Variable:	WTA	WTP	Endowment Effect
Panel A: Theoretical Relationships with Loss Aversion (Controlling for negative relationship with Risk Aversion)			
KR	0	−	+
HCPT	+	0	+
A-HCPT	+	−	+
Panel B: Study 1, Wave 1			
Loss Aversion	−0.04	−0.03	0.00
	(.033)	(.030)	(.030)
Risk Aversion	−0.73***	−0.38***	−0.28***
	(.093)	(.10)	(.11)
Panel C: Study 2 (IV Specification)			
Loss Aversion (Mixed)	−0.41***	0.05	−0.33***
	(.098)	(.088)	(.091)
Risk Aversion (Gains)	−0.51***	−0.12*	−0.30***
	(.078)	(.062)	(.067)
Risk Aversion (Losses)	0.20***	0.33***	−0.07
	(.073)	(.10)	(.083)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors in parentheses. Panel C displays bootstrapped standard errors from 10,000 simulations. All theories predict a negative relationship between risk aversion and WTA, WTP, and the endowment effect, so predictions in Panel A only apply to loss aversion. These predictions should be compared to the coefficients on loss aversion coefficients in the regressions. These specifications control for risk aversion, as the comparative statics in Panel A are partial correlations.

to all three theories, all three quantities are unrelated to DOSE-estimated loss aversion.[37]

[37] The DOSE technique uses simple binary choices, and a parametric model, to recover parameter estimates. We use the DOSE authors' preferred model, but others are possible: for example one that models the choices in a KR framework. The authors of the DOSE paper have attempted this, and found that the versions of the KR model they tried are badly misspecified: the parameter estimates fit far fewer choices than the model used here. Indeed, the KR models tend to fit around 50% of choices, close to the score one would obtain by guessing randomly.

Study 2 contained additional measures of risk preferences, as noted in Section 6. Three of these—Gain, Loss, and Mixed—can be used to identify loss aversion in risky choice. In Panel C of Table 8, we regress WTA, WTP, and the endowment effect on the certainty equivalents for all three measures. The certainty equivalent for a lottery over gains identifies risk aversion over gains, and the certainty equivalent for a lottery over losses identifies risk aversion over losses. Thus, by controlling for both of these quantities, the residual variation in the certainty equivalent for a lottery over gains and losses encodes loss aversion.[38] Here, again, our results are not compatible with the predictions of any of the three hybrid models, as we obtain a correlation of the *opposite* sign.[39]

These results are contrary to the models discussed above only if the same loss-aversion parameter governs both differences between risk aversion over gains and losses and the endowment effect. While this equivalence is a crucial aspect of current models of the endowment effect, there are other perspectives one could take. For example, one could take the position that, in the KR model, none of the first focus, unacclimating personal equilibrium, personal preferred equilibrium, or other concepts imply that moving from gains to losses will change the reference point, so neither of the measures in Table 8 are picking up the same loss aversion responsible for the endowment effect. This view makes our data consistent with the KR model, but also leaves changing risk attitudes between gain, loss and mixed lotteries—the original evidence for loss aversion—as an unexplained phenomenon. Or, one may posit, under CPT, that the loss-aversion parameters in the two contexts are different and uncorrelated—effectively rejecting the core idea of using the standard notion of loss aversion to explain the endowment effect.

[38] These MPLs are constructed such that both the lotteries, and certainty equivalents, are linear shifts of each other. Thus, the only thing different between the three MPLs is the relationship of the lotteries and certainty equivalents to zero.

[39] As each of these quantities is measured with error, we instrument all of them, as recommended by Gillen et al. (Forthcoming), through a "stacked" regression, as described in Section 2.2. One could also fit a parametric model using these three certainty equivalents. However, this would exclude participants whose choices over these six questions (three quantities, each asked twice) appeared to violate first-order stochastic dominance. As this is expected, and does happen with some frequency due to measurement error, the specification in Panel C allows us to retain all data. Note that all three of the variables in Panel C have been re-coded, as described in Section 6, so that higher values imply more risk and loss aversion.

In sum: there are versions of both CPT and KR that allow these theories to be consistent with the moments of the data presented throughout the previous section. However, these models also predict a positive relationship between the endowment effect and loss aversion, as measured by the change in risk aversion between the gain and loss domain. There is little support for this prediction in our data. Thus, to the extent that one accepts our empirical tests, one must accept that loss aversion is not a unified explanation for these phenomena, and/or search for a model that does not predict a correlation. The next section discusses additional evidence that might aid in this search.

8 Discussion

Using large, representative, incentivized surveys, we document a number of facts about WTA, WTP, and the endowment effect. First, WTA and WTP appear to be, at best, slightly correlated. This is true across a large number of demographic subgroups—except those with the highest IQ and education, where the correlation remains small. Second, the majority of people exhibit the endowment effect, although a substantial minority exhibit the opposite. Third, WTA, WTP, and the endowment effect are as stable across time as other measures of risk and time preferences. Fourth, WTA and WTP are correlated with other measures of risk preferences in a sensible way, depending on whether the fixed option in an MPL is a monetary amount or a lottery, suggesting that whether a question is framed as "buying" or "selling" is important. Fifth, our results are similar to existing studies that contain within-person measures of WTA and WTP. Sixth, these patterns are, in principle, compatible with modeling the endowment effect using hybrid models of reference-dependent preferences, although they require specific assumptions on joint distributions of parameters. Seventh, and finally, these hybrid models make additional predictions about the relationships between WTA, WTP, the endowment effect, and risk and loss aversion that are *not* supported in our data.

Before discussing the implications of our findings, we note that our data only elicited WTA and WTP for lotteries. Whether similar patterns hold for other types of tradable goods is an open question. Lotteries are, however, a natural object of study in this context. First, there is no doubt that there is an endowment effect for lotteries. Second, it is possible to endow participants with lotteries remotely, allowing for studies of the endowment effect on representative populations. Third, lotteries are intimately connected with risk and loss aversion, two standard parts of reference-dependent theories. Finally, and most importantly, lotteries are a key object of interest due to the centrality, in economics, of choices involving risk.

The most popular approach to modeling the disparity between WTA and WTP (in economics) is to link the endowment effect in general, and the patterns of WTA and WTP in particular, to loss aversion.[40] The patterns we document are consistent with some of these theories—for specific parameters—with the exception of the last set of our results. This last set directly tests the positive relationship between loss aversion and the endowment effect predicted by these theories, and finds no evidence of such a correlation.

Disregarding this last result, our data have important implications for reference-dependent theories. First, they tell us something about how reference points work: they can be set either by the structure of the MPL or by external framing. Second, hybrid theories are needed to allow for the multiple reference points that can exist in a single choice. Third, they put specific constraints on the distribution of parameters in the population. Fourth and finally, they rule out some popular special cases of these theories: for example, our results rule out models that attribute all small-stakes risk aversion to loss aversion.

However, our last set of results indicates that the best explanation for the endowment effect may not be a theory based on reference dependence and loss aversion. Instead, it shifts attention to theories in which a zero-correlation between WTA and WTP emerges because

[40] There are exceptions. Among them, a strand of the literature has linked the endowment effect to incomplete preferences (Bewley, 1986; Masatlioglu and Ok, 2005; Ortoleva, 2010; Masatlioglu and Ok, 2014). These models are trivially compatible with any correlation between WTA and WTP, but are incompatible with the presence of a negative endowment effect.

different processes—cognitive, attentional, emotional, and social—are connected to buying and selling. Such processes are the subject of a large literature on the endowment effect in psychology and neuroscience (see Morewedge and Giblin, 2015, for a review). A well-known theory in this literature is based on reference prices: it suggests that people have a both a subjective value and a reference price, and don't want to sell for less than the reference price or buy for more than the subjective value, which is typically more than the reference price. This theory has experimental support (Weaver and Frederick, 2012). It is also generates the endowment effect, and is compatible with a zero-correlation between WTA and WTP.

Other theories focus on how buying and selling frames evoke different cognitive processes. For example, several studies suggest that people access different information—from memory or the environment—depending on the task. In particular, the act of selling increases the availability of information pointing towards keeping the good, while the opposite happens during the activity of buying (Carmon and Ariely, 2000; Nayakankuppam and Mishra, 2005; Johnson et al., 2007; Ashby et al., 2012; Pachur and Scheibehenne, 2012).[41] This is further supported by how different types of information affects WTA and WTP differentially: inducing buyers to consider positive features of the goods increases WTP, while the same information does not affect WTA of sellers; the reverse happens with information related to negative features of the goods (Carmon and Ariely, 2000; Nayakankuppam and Mishra, 2005; Johnson et al., 2007). These theories have empirical support, and can match patterns that are not addressed by, or are incompatible with, loss aversion-based theories.

Recent work in neuroscience offers suggestive evidence supporting the presence of two different mechanistic processes (in the brain) governing buying and selling. An early fMRI study found distinct activity in the medial prefrontal cortex (mPFC) when making a buying decision at a low price and (more weakly) a selling decision at a high price (Knutson et al., 2008). This region encodes abstract integrated types of value, which is consistent with a

[41] For example, buyers of pens and lotteries recall fewer positive and more negative attributes than sellers (Nayakankuppam and Mishra, 2005; Saqib et al., 2010). Similarly, buyers of basketball tickets tend to consider costs of attending the game, while sellers consider the benefits of attending (Carmon and Ariely, 2000).

value for being relatively confident in getting a good deal. Further, the size of the endowment effect was correlated with the amount of differential activity in the insula during selling versus buying. This indicates that selling and buying decisions are processed differently in the brain, and the difference is likely associated with more uncertainty or emotional discomfort during selling (as those are common functions of the insula). Another study found increased activity in the amygdala and caudate nucleus during selling; the former is consistent with enhanced emotional salience during selling (Weber et al., 2007). Finally, there is evidence that WTP is processed by the the medial orbitofrontal cortex (mOFC), while WTA is processed by a more lateral portion of the OFC (the lOFC, see De Martino et al., 2009). These studies do not form a simple, integrated neuro-psychological picture, but all show differences in neural processing during buying versus selling.

Additional evidence consistent with different processes can be found in response times to our own questions. The WTA module had a much shorter response time than the WTP module. The median response time for WTA questions was 88 seconds, while it was 122 seconds for WTP.[42] This difference is large: of the eight risk aversion modules in Section 6, WTP has the longest median response time, while the Risk Aversion Gain module had the shortest (74 seconds). Moreover, although participants take longer on earlier modules, even when the WTP module is randomly selected to be later in the survey, it still takes longer than WTA when WTA is randomly selected to be earlier in the survey. This difference is compatible with the idea that different processes are involved in buying and selling.

In keeping with the goal of this paper—to document facts about WTA, WTP, and the endowment effect, and to explore how leading models in economics account for them—this discussion is far from suggesting a particular formal theory. We leave the development of novel models for future research. However, we believe our results suggest the usefulness of

[42]This difference was true question-by-question as well. Moreover, this difference is unlikely to be driven by the slightly longer instructions of the WTP questions. The instruction screen for the WTA and WTP module are approximately the same length as the difference between the questions, and participants spent a median of 5 seconds on each. These instruction screens are not included in the overall module response times above.

existing results and theories from psychology and neuroscience in deepening our knowledge of the processes involved in buying and selling. These actions are at the center of economic activity. Surprisingly, they are not yet fully understood by economists.

References

Abdellaoui, Mohammed, "Parameter-Free Elicitation of Utility and Probability Weighting Functions," *Management Science*, 2000, *46* (11), 1497–1512.

Ashby, Nathaniel J.S. Stephan Dickert, and Andreas Glöckner, "Focusing on What You Own: Biased Information Uptake due to Ownership," *Judgment and Decision Making*, 2012, *7* (3), 254.

Beattie, Jane and Graham Loomes, "The Impact of Incentives upon Risky Choice Experiments," *Journal of Risk and Uncertainty*, March 1997, *14* (2), 155–68.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak, "Measuring Utility by a Single-Response Sequential Method," *Behavioral Science*, 1964, *9* (3), 226–232.

Bell, David E., "Disappointment in Decision Making under Uncertainty," *Operations Research*, 1985, *33* (1), 1–27.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Risk Preference Instability Across Institutions: A Dilemma," *Proceedings of the National Academy of Sciences of the United States of America*, 2005, *102* (11), 4209–4214.

Bewley, Truman F, "Knightian Decision Theory: Part I," *Cowles Foundation discussion paper*, 1986, *807.*

Borges, Bernhard F.J. and Jack L. Knetsch, "Tests of Market Outcomes with Asymmetric Valuations of Gains and Losses: Smaller Gains, Fewer Trades, and Less Value," *Journal of Economic Behavior & Organization*, 1998, *33* (2), 185–193.

Brown, Jeffrey R. Arie Kapteyn, Erzo F.P. Luttmer, and Olivia S. Mitchell, "Cognitive Constraints on Valuing Annuities," *Journal of the European Economic Association*, 2017, *15* (2), 429–462.

Bruner, David M., "Changing the Probability versus Changing the Reward," *Experimental Economics*, 2009, *12* (4), 367–385.

Camerer, Colin F., "Individual Decision Making," in John H. Kagel and Alvin E. Roth, eds., *Handbook of Experimental Economics*, Vol. 2, Princeton, NJ: Princeton University Press, 1995, pp. 587–703.

_ **, Jonathan Chapman, Pietro Ortoleva, and Erik Snowberg**, "Time Stability of Econographics," 2019. California Insitute of Technology, *mimeo.*

_ **, Mark Dean, Jonathan Chapman, Pietro Ortoleva, and Erik Snowberg**, "Econographics," 2019. California Insitute of Technology, *mimeo.*

Carmon, Ziv and Dan Ariely, "Focusing on the Forgone: How Value can Appear so Different to Buyers and Sellers," *Journal of Consumer Research*, 2000, *27* (3), 360–370.

Carvalho, Leandro, Pietro Ortoleva, Francisco Perez-Arce, and Erik Snowberg, "Education and Preferences," 2017. University of Southern California, *mimeo.*

Cason, Timothy N. and Charles R. Plott, "Misconceptions and Game Form Recognition: Challenges to Theories of Revealed Preference and Framing," *Journal of Political Economy*, 2014, *122* (6), 1235–1270.

Chapman, Jonathan, Erik Snowberg, Stephanie Wang, and Colin F. Camerer, "Dynamically Optimized Sequential Experimentation (DOSE) for for Estimating Economic Preference Parameters," 2019. California Insitute of Technology, *mimeo.*

Condon, David M. and William Revelle, "The International Cognitive Ability Resource: Development and Initial Validation of a Public-domain Measure," *Intelligence*, 2014, *43* (2), 52–64.

Crosetto, Paolo and Antonio Filippin, "A Theoretical and Experimental Appraisal of Four Risk Elicitation Methods," *Experimental Economics*, 2016, *19* (3), 613–641.

Cubitt, Robin, Chris Starmer, and Robert Sugden, "On the Validity of the Random Lottery Incentive System," *Experimental Economics*, 1998, *1* (2), 115–131.

De Martino, Benedetto, Dharshan Kumaran, Beatrice Holt, and Raymond J Dolan, "The Neurobiology of Reference-Dependent Value Computation," *The Journal of Neuroscience*, 2009, *29* (12), 3833–3842.

Dean, Mark and Pietro Ortoleva, "Is it All Connected? A Testing Ground for Unified Theories of Behavioral Economics Phenomena," 2016. Columbia University, *mimeo.*

Deck, Cary, Jungmin Lee, and Javier A. Reyes, "Personality and the Consistency of Risk Taking Behavior: Experimental Evidence," 2010. University of Arkansas, *mimeo.*

_ , _ , Javier Reyes, and Christopher Rosen, "A Failed Attempt to Explain within subject Variation in Risk Taking Behavior using domain specific risk attitudes," *Journal of Economic Behavior & Organization*, 2013, *87*, 1–24.

Dhami, Sanjit, *The Foundations of Behavioral Economic Analysis*, Oxford University Press, 2016.

Eisenberger, Roselies and Martin Weber, "Willingness-to-pay and Willingness-to-accept for Risky and Ambiguous Lotteries," *Journal of Risk and Uncertainty*, 1995, *10* (3), 223–233.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde, "An Experimentally Validated Preference Survey Module," 2013. University of Bonn, *mimeo.*

Fehr, Dietmar, Rustamdjan Hakimov, and Dorothea Kübler, "The Willingness to Pay-Willingness to Accept Gap: A Failed Replication of Plott and Zeiler," *European Economic Review*, 2015, *78*, 120–128.

Frederick, Shane, "Cognitive Reflection and Decision Making," *The Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

Freeman, David, Yoram Halevy, and Terri Kneeland, "Eliciting Risk Preferences using Choice Lists," 2015. University of British Columbia, *mimeo.*

Gächter, Simon, Eric J. Johnson, and Andreas Herrmann, "Individual-level Loss Aversion in Riskless and Risky Choices," *IZA Discussion Paper # 2961*, 2007.

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques and Applications from the Caltech Cohort Study," *Journal of Political Economy*, Forthcoming.

Harbaugh, William T., Kate Krause, and Lise Vesterlund, "The Fourfold Pattern of Risk Attitudes in Choice and Pricing Tasks," *Economic Journal*, 2010, *120* (6), 595–611.

Harless, David W., "More Laboratory Evidence on the Disparity between Willingness to Pay and Compensation Demanded," *Journal of Economic Behavior & Organization*, 1989, *11* (3), 359–379.

He, Pan, Marcella Veronesi, and Stefanie Engel, "Consistency of Risk Preference Measures: An Artefactual Field Experiment from Rural China," *The Journal of Development Studies*, 2017, *Forthcoming.*

Heffetz, Ori and John A List, "Is the endowment effect an expectations effect?," *Journal of the European Economic Association*, 2014, *12* (5), 1396–1422.

Hershey, John C. and Paul J.H. Schoemaker, "Probability Versus Certainty Equivalence Methods in Utility Measurement: Are they Equivalent?," *Management Science*, 1985, *31* (10), 1213–1231.

Hey, John and Jinkwon Lee, "Do Subjects Separate (or Are They Sophisticated)?," *Experimental Economics*, September 2005, *8* (3), 233–265.

Holt, Charles A. and Susan K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Isaac, R Mark and Duncan James, "Just who are you Calling Risk Averse?," *Journal of Risk and Uncertainty*, 2000, *20* (2), 177–187.

Isoni, Andrea, Graham Loomes, and Robert Sugden, "The Willingness to Pay-Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations: Comment," *American Economic Review*, 2011, *101* (2), 991–1011.

Johnson, Eric J., Gerald Häubl, and Anat Keinan, "Aspects of Endowment: A Query Theory of Value Construction.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2007, *33* (3), 461.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, March 1979, *47* (2), 263–292.

_ **, Jack L. Knetsch, and Richard H. Thaler**, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, 1990, *98* (6), 1325–1348.

Karni, Edi and Zvi Safra, "'Preference Reversal' and the Observability of Preferences by Experimental Methods," *Econometrica*, 1987, *55* (3), 675–685.

Knetsch, Jack L. and Wei-Kang Wong, "The Endowment Effect and the Reference State: Evidence and Manipulations," *Journal of Economic Behavior & Organization*, 2009, *71* (2), 407–413.

Knutson, Brian, G.Ẽlliott Wimmer, Scott Rick, Nick G Hollon, Drazen Prelec, and George Loewenstein, "Neural Antecedents of the Endowment Effect," *Neuron*, 2008, *58* (5), 814–822.

Kőszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

_ **and** _ , "Reference-Dependent Risk Attitudes," *The American Economic Review*, 2007, *97* (4), 1047–1073.

Kurata, Hiroshi, Hiroshi Izawa, and Makoto Okamura, "Non-Expected Utility Maximizers Behave as if Expected Utility Maximizers: An Experimental Test," *Journal of Socio-Economics*, 2009, *38* (4), 622–629.

Lönnqvist, Jan-Erik, Markku Verkasalo, Gari Walkowitz, and Philipp C Wichardt, "Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison," *Journal of Economic Behavior & Organization*, 2015, *119*, 254–266.

Loomes, Graham and Ganna Pogrebna, "Measuring Individual Risk Attitudes when Preferences are Imprecise," *The Economic Journal*, 2014, *124* (576), 569–593.

_ **and Robert Sugden**, "Disappointment and Dynamic Consistency in Choice under Uncertainty," *The Review of Economic Studies*, 1986, *53* (2), 271–282.

_ **, Chris Starmer, and Robert Sugden**, "Do Anomalies Disappear in Repeated Markets?," *The Economic Journal*, 2003, *113* (486), C153–C166.

Marzilli Ericson, Keith M. and Andreas Fuster, "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments," *The Quarterly Journal of Economics*, 2011, *126* (4), 1879–1907.

Marzilli Ericson, Keith M. and Andreas Fuster, "The Endowment Effect," *Annual Review of Economics*, 2014, *6* (1), 555–579.

Masatlioglu, Yusufcan and Efe A. Ok, "Rational Choice with Status-Quo Bias," *Journal of Economic Theory*, 2005, *121*, 1–29.

_ and _ , “A Canonical Choice Model with Initial Endowment,” *Review of Economic Studies*, 2014, *81* (2), 851–883.

Morewedge, Carey K. and Colleen E. Giblin, “Explanations of the Endowment Effect: an Integrative Review,” *Trends in Cognitive Sciences*, 2015, *19* (6), 339–348.

Nayakankuppam, Dhananjay and Himanshu Mishra, “The Endowment Effect: Rose-tinted and Dark-tinted Glasses,” *Journal of Consumer Research*, 2005, *32* (3), 390–395.

Nielsen, Thea, Alwin Keil, and Manfred Zeller, “Assessing Farmers’ Risk Preferences and their Determinants in a Marginal Upland Area of Vietnam: A Comparison of Multiple Elicitation Techniques,” *Agricultural Economics*, 2013, *44* (3), 255–273.

Ortoleva, Pietro, “Status Quo Bias, Multiple Priors and Uncertainty Aversion,” *Games and Economic Behavior*, 2010, *69*, 411–424.

Pachur, Thorsten and Benjamin Scheibehenne, “Constructing Preference from Experience: The Endowment Effect Reflected in External Information Search,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2012, *38* (4), 1108–1116.

Pew Research Center, *Evaluating Online Nonprobability Surveys*, www.pewresearch.org, May 2016.

Plott, Charles R. and K. Zeiler, “The Willingness to Pay-Willingness to Accept Gap, The ‘Endowment Effect,’ Subject Misconceptions, and Experimental Procedures for Eliciting Valuations,” *American Economic Review*, 2005, *95* (3), 530–545.

Prelec, Drazen, “The Probability Weighting Function,” *Econometrica*, May 1998, *66* (3), 497–527.

Rabin, Matthew, “Risk Aversion and Expected-Utility Theory: A Calibration Theorem,” *Econometrica*, 2000, *68*, 1281–1292.

Saqib, Najam U., Norman Frohlich, and Edward Bruning, “The Influence of Involvement on the Endowment Effect: The Moveable Value Function,” *Journal of Consumer Psychology*, 2010, *20* (3), 355–368.

Schmidt, Ulrich and Stefan T. Trautmann, “Common Consequence effects in Pricing and Choice,” *Theory and Decision*, 2014, *76* (1), 1–7.

_ and Stefan Traub, “An Experimental Investigation of the Disparity between WTA and WTP for Lotteries,” *Theory and Decision*, 2009, *66* (3), 229–262.

_ , Chris Starmer, and Robert Sugden, “Third-Generation Prospect Theory,” *Journal of Risk and Uncertainty*, 2008, *36* (3), 203–223.

Sokol-Hessner, Peter, Ming Hsu, Nina G. Curley, Mauricio R. Delgado, Colin F. Camerer, and Elizabeth A. Phelps, “Thinking like a Trader Selectively Reduces Individuals’ Loss Aversion,” *Proceedings of the National Academy of Sciences*, 2009, *106* (13), 5035–5040.

Sprenger, Charles, "An Endowment Effect for Risk: Experimental Tests of Stochastic Reference Points," *Journal of Political Economy*, 2015, *123* (6), 1456–1499.

Tunçel, Tuba and James K. Hammitt, "A New Meta-Analysis on the WTP/WTA Disparity," *Journal of Environmental Economics and Management*, 2014, *68* (1), 175–187.

Tversky, Amos and Daniel Kahneman, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 1991, *106* (4), 1039–1061.

_ **and** _ , "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

Weaver, Ray and Shane Frederick, "A reference price theory of the endowment effect," *Journal of Marketing Research*, 2012, *49* (5), 696–707.

Weber, Bernd, Andreas Aholt, Carolin Neuhaus, Peter Trautner, Christian E. Elger, and Thorsten Teichert, "Neural Evidence for Reference-dependence in Real-market-transactions," *Neuroimage*, 2007, *35* (1), 441–447.

Weber, Elke U. and Eric J. Johnson, "Decisions under Uncertainty: Psychological, Economic, and Neuroeconomic Explanations of Risk Preference," in Paul W. Glimcher, Ernst Fehr, Colin Camerer, and Russell Alan Poldrack, eds., *Neuroeconomics: Decision Making and the Brain*, New York, NY: Elsevier, 2008, pp. 127–144.

Wu, George and Richard Gonzalez, "Curvature of the Probability Weighting Function," *Management Science*, 1996, *42* (12), 1676–1690.

Online Appendix—Not Intended for Publication

A Unweighted Specifications

Table A.1: Table 2, unweighted.

	Correlation between WTA and WTP				Correlation within Type	
	Lottery 1	Lottery 2	Averages	ORIV	WTA	WTP
Study 1, Wave 1	-0.05^{**} (.027)	-0.05^{*} (.025)	-0.07^{**} (.027)	-0.08^{**} (.032)	0.69^{***} (.019)	0.74^{***} (.018)
Study 1, Wave 2	-0.02 (.032)	0.02 (.031)	0.01 (.033)	0.01 (.040)	0.66^{***} (.024)	0.75^{***} (.020)
Study 2	-0.08^{**} (.035)	-0.04 (.036)	-0.07^{**} (.036)	-0.09^{**} (.044)	0.70^{***} (.030)	0.78^{***} (.025)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Table A.2: Table 3, unweighted

	Lottery 1			Lottery 2		
	WTA<WTP	WTA=WTP	WTA>WTP	WTA<WTP	WTA=WTP	WTA>WTP
Study 1, Wave 1	32%	12%	57%	24%	15%	61%
Study 1, Wave 2	29%	12%	59%	22%	15%	62%
Study 2	28%	12%	60%	25%	12%	62%

Figure A.1: Figure 1, unweighted.

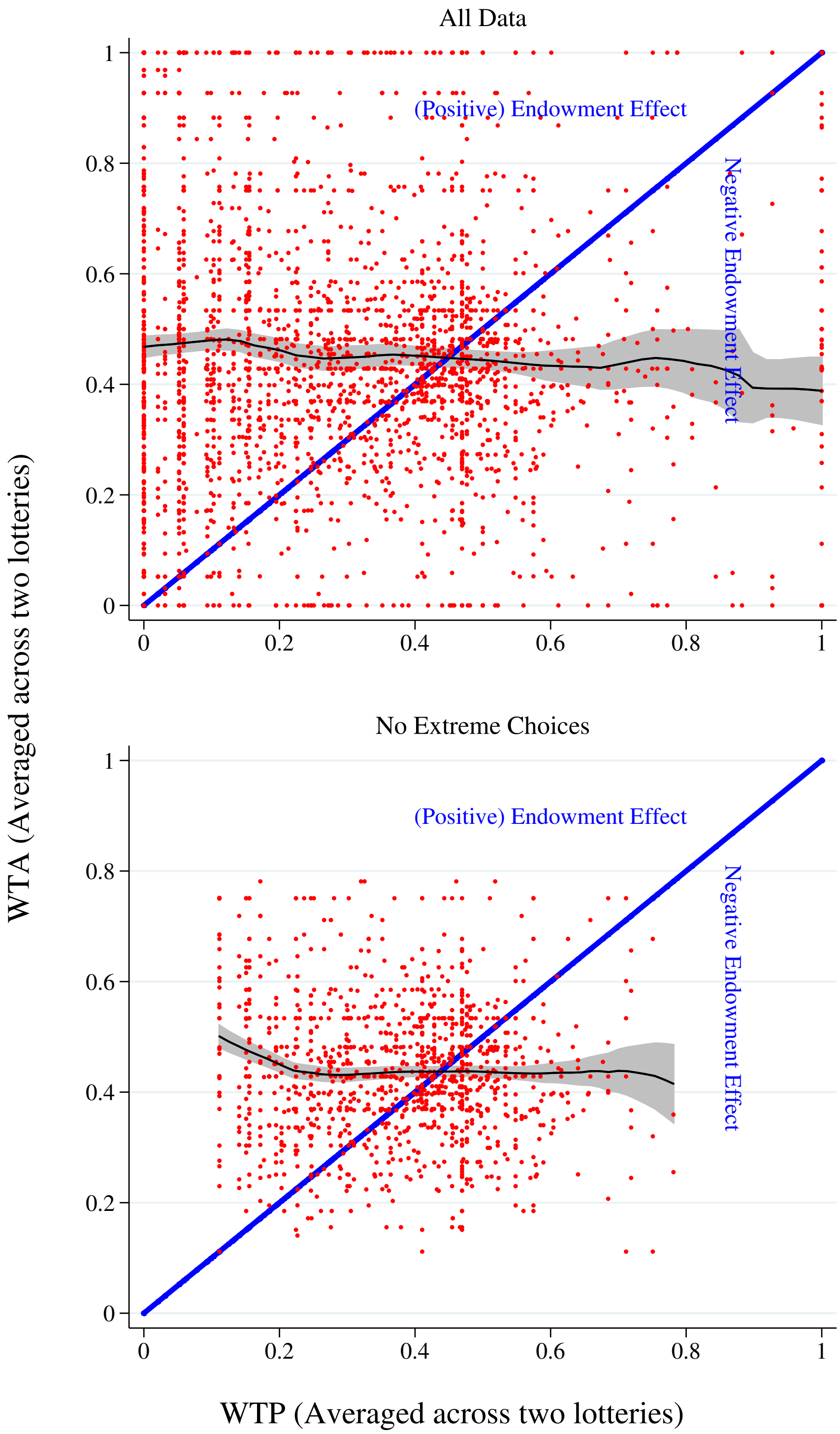


Table A.3: Table 4, unweighted.

	N	Lottery 1	Lottery 2	Average	ORIV
All	3,000	−0.06***	−0.05**	−0.06***	−0.08***
		(.021)	(.021)	(.021)	(.026)
Not Too Fast	2,700	−0.03	−0.03	−0.04*	−0.05*
		(.022)	(.022)	(.022)	(.027)
No Dominated Choices	2,672	−0.05**	−0.05**	−0.07***	−0.08***
		(.024)	(.022)	(.023)	(.027)
No Switches in Top or Bottom Two Rows	1,595	−0.08***	−0.08***	−0.08***	−0.12***
		(.031)	(.033)	(.032)	(.040)
No Switches in Top or Bottom Three Rows	1,028	−0.01	0.02	0.01	−0.01
		(.036)	(.036)	(.037)	(.054)
Question Order: WTA First	1,501	−0.07**	−0.04	−0.06**	−0.08**
		(.030)	(.031)	(.030)	(.037)
Question Order: WTP First	1,499	−0.06*	−0.06*	−0.07**	−0.09**
		(.030)	(.029)	(.031)	(.037)
Education: HS or Less	1,128	−0.11***	−0.06*	−0.11***	−0.13***
		(.033)	(.034)	(.034)	(.041)
Education: Some College	1,538	−0.03	−0.03	−0.03	−0.04
		(.031)	(.030)	(.031)	(.037)
Education: Advanced Degree	334	0.00	−0.04	−0.02	−0.05
		(.068)	(.060)	(.066)	(.080)
Income: Above Median	1,520	0.00	−0.03	−0.01	−0.02
		(.031)	(.030)	(.031)	(.037)
Income: Top Quartile	813	0.04	0.02	0.05	0.05
		(.042)	(.041)	(.042)	(.053)
Income: Top Decile	399	0.09	0.04	0.07	0.08
		(.062)	(.059)	(.063)	(.078)
CRT: Above Median	1,371	0.08**	0.06*	0.07**	0.09**
		(.033)	(.033)	(.034)	(.041)
CRT: Top Decile	288	0.10	0.11	0.10	0.11
		(.078)	(.076)	(.079)	(.093)
IQ: Above Median	1,694	−0.01	−0.02	−0.01	−0.02
		(.029)	(.027)	(.029)	(.034)
IQ: Top Decile	337	0.12*	0.10*	0.12**	0.14*
		(.062)	(.060)	(.061)	(.075)
IQ: Top 5%	146	0.25***	0.15*	0.21**	0.26**
		(.088)	(.088)	(.085)	(.11)
Age: Youngest Quartile	784	−0.06	−0.08**	−0.09**	−0.11**
		(.043)	(.041)	(.043)	(.052)
Age: Second Youngest Quartile	730	−0.08*	−0.01	−0.06	−0.07
		(.043)	(.046)	(.044)	(.052)
Age: Second Oldest Quartile	786	−0.06	0.00	−0.04	−0.05
		(.043)	(.040)	(.041)	(.050)
Age: Oldest Quartile	700	−0.05	−0.10**	−0.07*	−0.10*
		(.042)	(.042)	(.043)	(.052)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Table A.4: Table 7, unweighted.

				Fixed Lottery				Variable Lottery	
		WTA	Urn	Gain	Mixed	Loss	WTP	FM	2L
	Urn	−0.61***					0.05		
		(.038)					(.043)		
	Gain	−0.62***	0.68***				0.05		
Fixed		(.040)	(.032)				(.046)		
Lottery	Mixed	−0.51***	0.47***	0.55***			0.23***		
		(.034)	(.035)	(.033)			(.042)		
	Loss	−0.30***	0.26***	0.36***	0.67***		0.34***		
		(.042)	(.041)	(.043)	(.029)		(.048)		
	FM	0.02	0.04	0.05	−0.16***	−0.18***	−0.38***		
Variable		(.047)	(.045)	(.048)	(.042)	(.043)	(.034)		
Lottery	2L	0.12***	−0.17***	−0.16***	−0.27***	−0.23***	−0.27***	0.42***	
		(.047)	(.045)	(.047)	(.042)	(.043)	(.040)	(.039)	
Qualitative		−0.16***	0.17***	0.17***	0.16***	0.01	−0.11***	0.08*	0.04
		(.044)	(.039)	(.046)	(.041)	(.050)	(.041)	(.042)	(.045)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with standard errors in parentheses.

Figure A.2: Figure 2, unweighted.

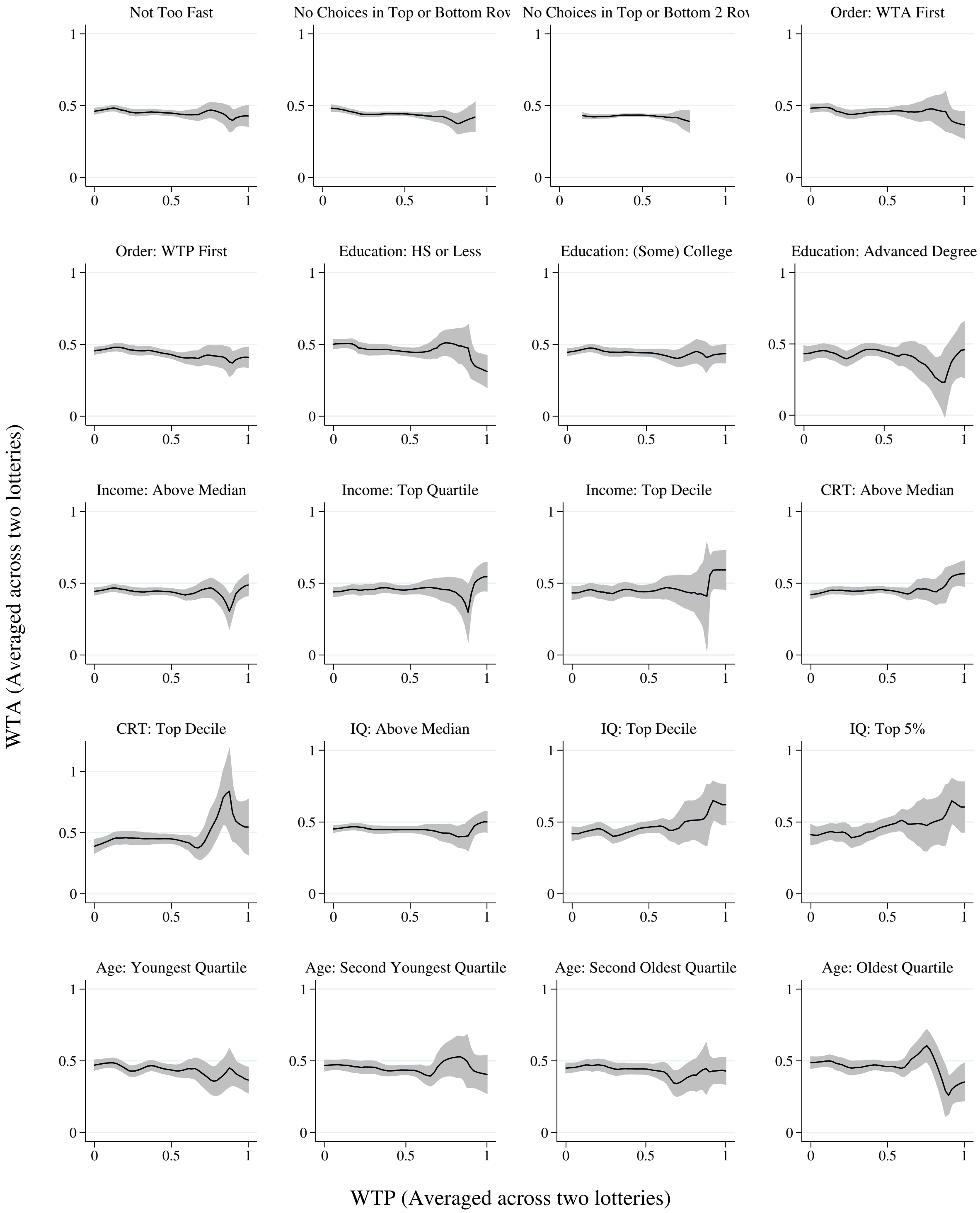


Table A.5: Table 8, unweighted.

Dependent Variable:	WTA	WTP	Endowment Effect
Panel A: Theoretical Relationships with Loss Aversion (Controlling for negative relationship with Risk Aversion)			
Prospect Theory (Standard)	+	0	+
Prospect Theory (Alternative)	+	−	+
Stochastic Reference Dependence	0	−	+
Panel B: Study 1, Wave 1			
Loss Aversion	−0.04*	−0.07***	0.02
	(.021)	(.021)	(.021)
Risk Aversion	−0.66***	−0.52***	−0.13**
	(.059)	(.061)	(.062)
Panel C: Study 2 (IV Specification)			
Loss Aversion (Mixed)	−0.28***	0.06	−0.25***
	(.051)	(.050)	(.049)
Risk Aversion (Gains)	−0.48***	−0.12***	−0.27***
	(.042)	(.039)	(.039)
Risk Aversion (Losses)	0.07	0.38***	−0.19***
	(.051)	(.062)	(.053)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Standard errors in parentheses. Panel C displays bootstrapped standard errors from 10,000 simulations.

B Results for each Study/Wave

Figure B.1: Figure 1 for Study 1, Wave 1 only.

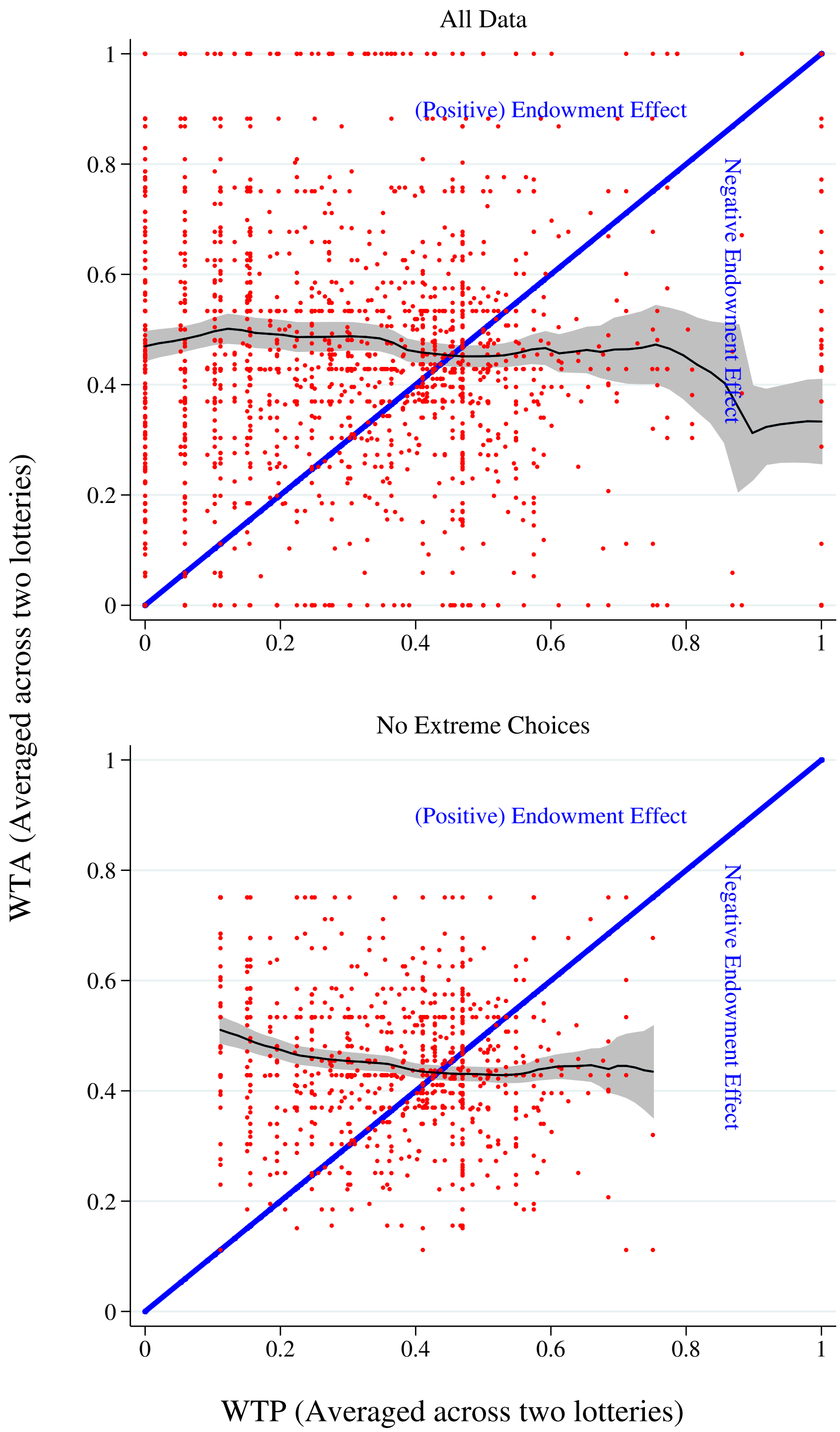


Table B.1: Table 4 for Study 1, Wave 1 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	2,000	−0.06*	−0.06*	−0.08**	−0.09**
		(.037)	(.037)	(.037)	(.044)
Not Too Fast	1,800	−0.03	−0.04	−0.05	−0.06
		(.040)	(.039)	(.039)	(.048)
No Dominated Choices	2,000	−0.06*	−0.06*	−0.08**	−0.09**
		(.037)	(.037)	(.037)	(.043)
No Switches in Top or Bottom Two Rows	923	−0.11*	−0.14***	−0.17***	−0.23***
		(.061)	(.051)	(.052)	(.071)
No Switches in Top or Bottom Three Rows	629	−0.06	0.00	−0.07	−0.12
		(.063)	(.069)	(.077)	(.12)
Question Order: WTA First	1,011	−0.06	−0.07	−0.08*	−0.09*
		(.046)	(.043)	(.044)	(.053)
Question Order: WTP First	989	−0.06	−0.06	−0.08	−0.09
		(.057)	(.058)	(.058)	(.068)
Education: HS or Less	765	−0.13**	−0.09	−0.13**	−0.16**
		(.059)	(.064)	(.061)	(.071)
Education: Some College	1,018	−0.01	−0.04	−0.03	−0.04
		(.050)	(.044)	(.050)	(.057)
Education: Advanced Degree	217	0.05	−0.02	0.00	0.01
		(.093)	(.079)	(.089)	(.11)
Income: Above Median	1,004	−0.04	−0.03	−0.04	−0.05
		(.046)	(.053)	(.046)	(.056)
Income: Top Quartile	540	−0.03	0.04	0.01	0.01
		(.069)	(.077)	(.069)	(.084)
Income: Top Decile	275	0.12	0.00	0.06	0.08
		(.085)	(.080)	(.088)	(.11)
CRT: Above Median	913	0.12**	0.04	0.08	0.10
		(.053)	(.049)	(.052)	(.065)
CRT: Top Decile	189	0.16	0.21**	0.16	0.19
		(.100)	(.097)	(.10)	(.13)
IQ: Above Median	1,117	−0.01	−0.05	−0.04	−0.05
		(.049)	(.043)	(.047)	(.057)
IQ: Top Decile	232	0.10	0.09	0.09	0.11
		(.112)	(.095)	(.11)	(.14)
IQ: Top 5%	97	0.30***	0.16	0.20*	0.29
		(.113)	(.127)	(.12)	(.22)
Age: Youngest Quartile	374	−0.17*	−0.19**	−0.23**	−0.27**
		(.093)	(.087)	(.094)	(.11)
Age: Second Youngest Quartile	563	0.01	0.00	0.00	0.00
		(.070)	(.069)	(.064)	(.077)
Age: Second Oldest Quartile	582	−0.08	−0.07	−0.09	−0.11
		(.059)	(.052)	(.058)	(.068)
Age: Oldest Quartile	481	0.00	0.00	0.01	0.01
		(.060)	(.071)	(.058)	(.072)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Figure B.2: Figure 2 for Study 1, Wave 1 only.

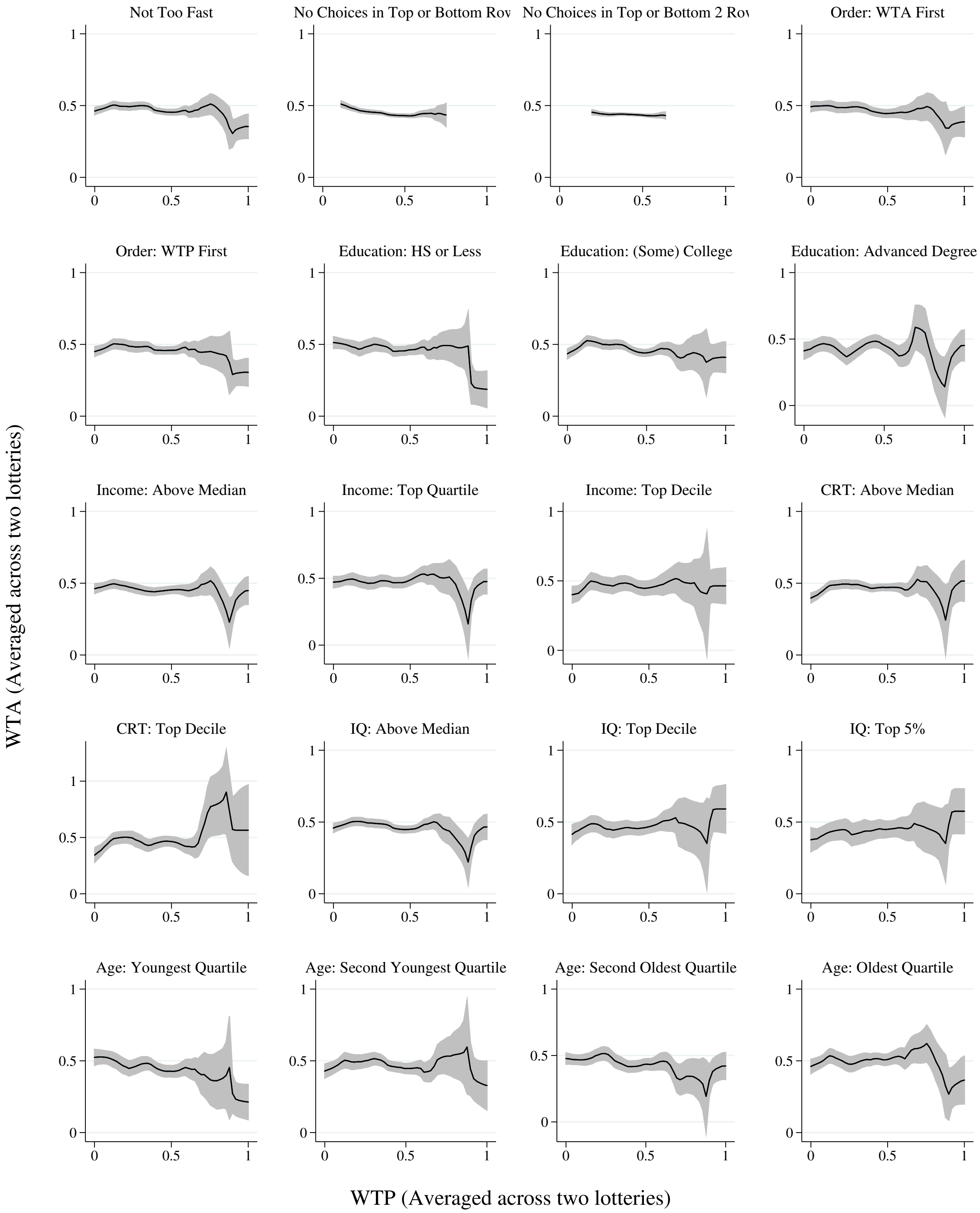


Figure B.3: Figure 1 for Study 1, Wave 2 only.

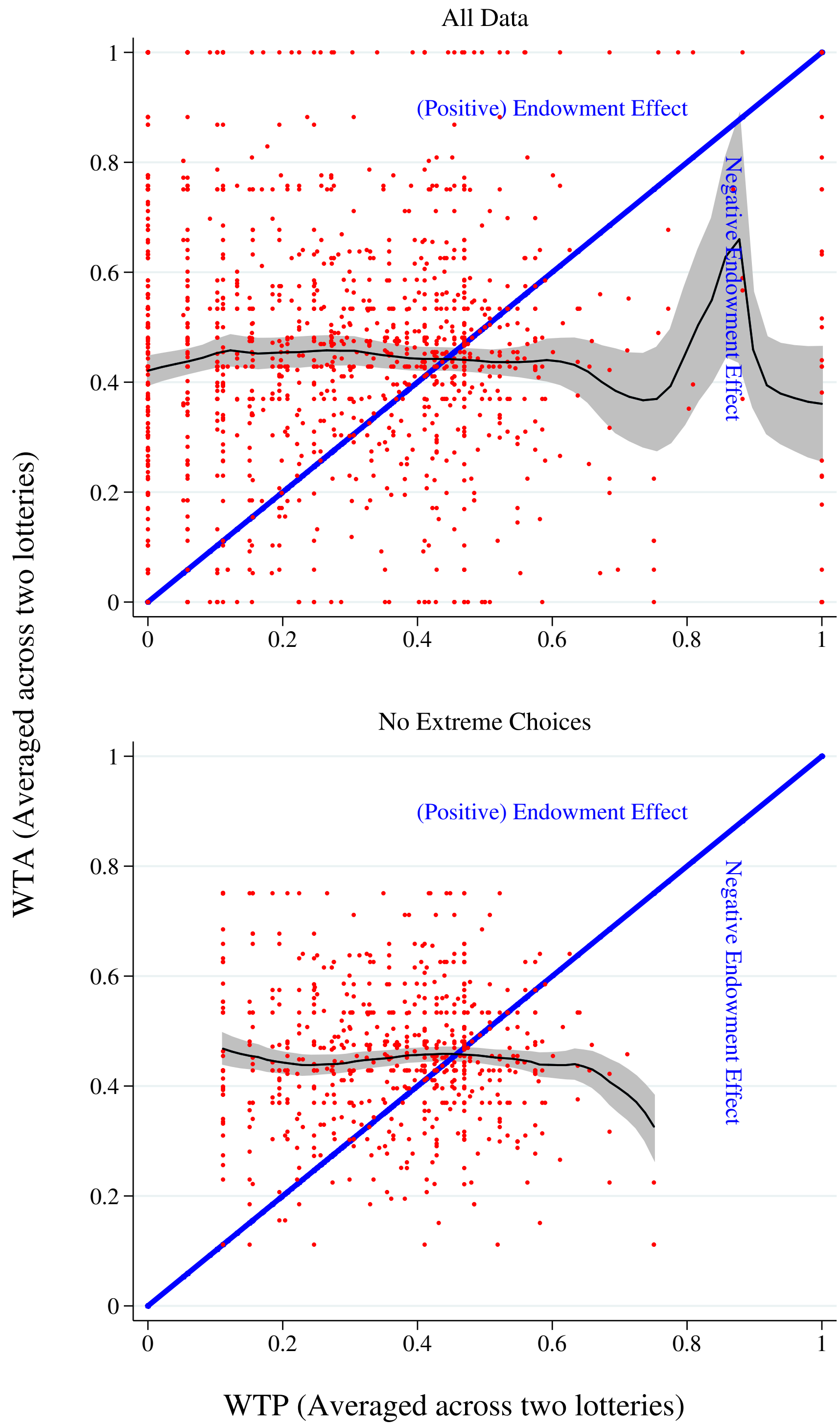


Table B.2: Table 4 for Study 1, Wave 2 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	1,465	−0.01	−0.02	−0.02	−0.02
		(.050)	(.049)	(.053)	(.063)
Not Too Fast	1,319	0.00	−0.02	−0.01	−0.01
		(.052)	(.052)	(.057)	(.069)
No Dominated Choices	1,465	−0.01	−0.02	−0.02	−0.02
		(.050)	(.049)	(.053)	(.063)
No Switches in Top or Bottom Two Rows	717	0.05	−0.13***	−0.05	−0.06
		(.058)	(.051)	(.052)	(.073)
No Switches in Top or Bottom Three Rows	483	0.10	−0.04	0.02	0.04
		(.080)	(.061)	(.070)	(.11)
Question Order: WTA First	754	0.00	0.07	0.04	0.05
		(.069)	(.069)	(.075)	(.088)
Question Order: WTP First	711	−0.03	−0.12*	−0.08	−0.09
		(.071)	(.064)	(.074)	(.088)
Education: HS or Less	525	−0.02	−0.06	−0.05	−0.05
		(.071)	(.075)	(.077)	(.088)
Education: Some College	774	0.00	−0.01	0.00	0.00
		(.075)	(.072)	(.083)	(.10)
Education: Advanced Degree	166	−0.06	0.08	0.04	0.04
		(.089)	(.090)	(.087)	(.11)
Income: Above Median	751	0.06	0.13**	0.11*	0.13*
		(.064)	(.061)	(.064)	(.076)
Income: Top Quartile	389	0.05	0.12*	0.11	0.14
		(.087)	(.073)	(.076)	(.094)
Income: Top Decile	203	0.02	0.12	0.07	0.08
		(.126)	(.106)	(.113)	(.138)
CRT: Above Median	702	0.12**	0.12**	0.15***	0.19***
		(.058)	(.053)	(.054)	(.068)
CRT: Top Decile	190	0.18*	0.04	0.13	0.18
		(.11)	(.12)	(.11)	(.16)
IQ: Above Median	855	−0.01	0.02	0.01	0.02
		(.067)	(.064)	(.072)	(.090)
IQ: Top Decile	208	0.20**	0.29**	0.30**	0.36**
		(.098)	(.13)	(.12)	(.14)
IQ: Top 5%	95	0.30**	0.23**	0.29**	0.35**
		(.142)	(.114)	(.13)	(.17)
Age: Youngest Quartile	252	−0.05	−0.11	−0.09	−0.11
		(.12)	(.11)	(.13)	(.16)
Age: Second Youngest Quartile	389	0.04	0.06	0.07	0.08
		(.11)	(.11)	(.12)	(.13)
Age: Second Oldest Quartile	442	−0.03	−0.04	−0.04	−0.05
		(.083)	(.076)	(.084)	(.10)
Age: Oldest Quartile	382	−0.02	−0.01	0.00	0.00
		(.073)	(.069)	(.071)	(.087)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Figure B.4: Figure 2 for Study 1, Wave 2 only.

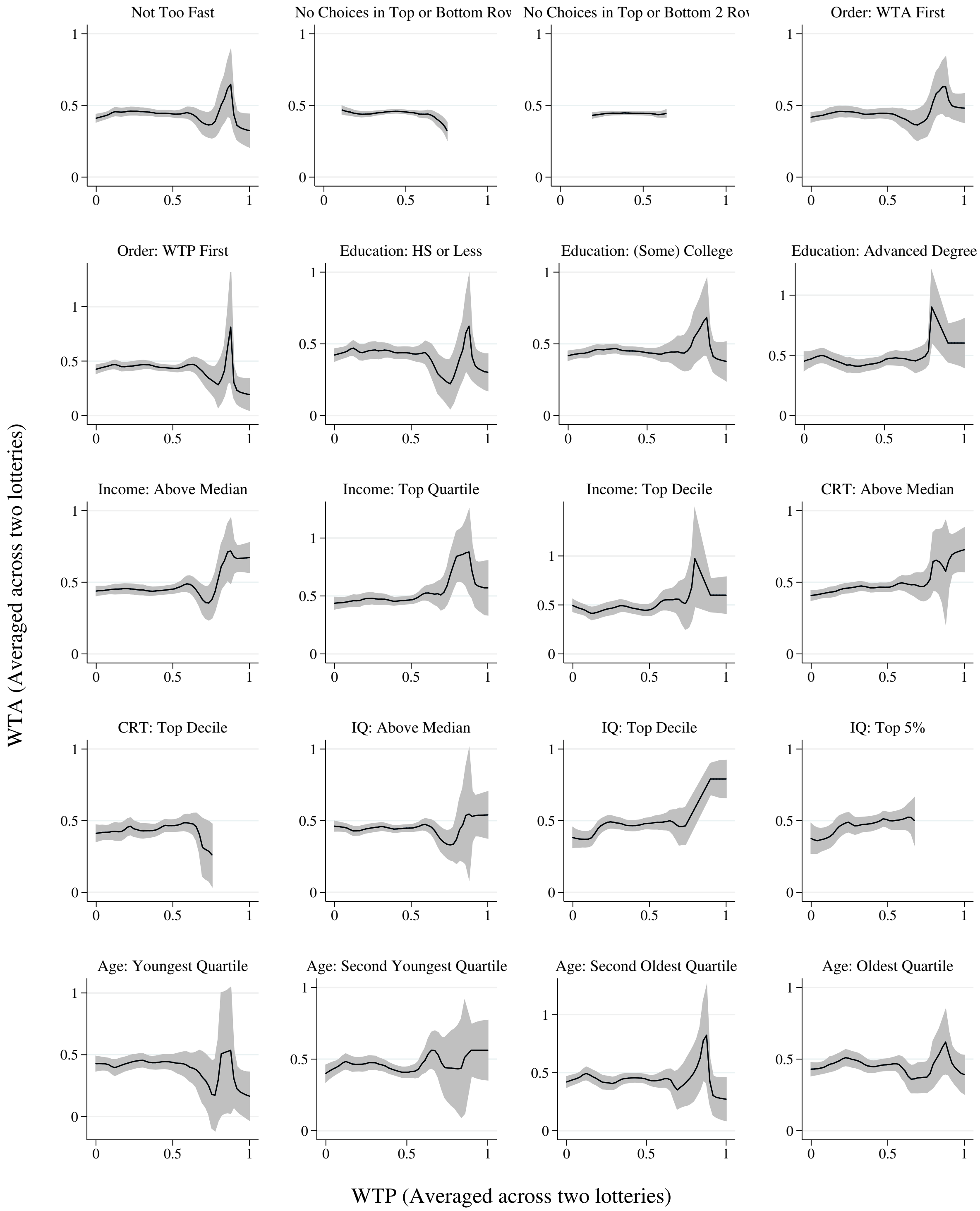


Figure B.5: Figure 1 for Study 2 only.

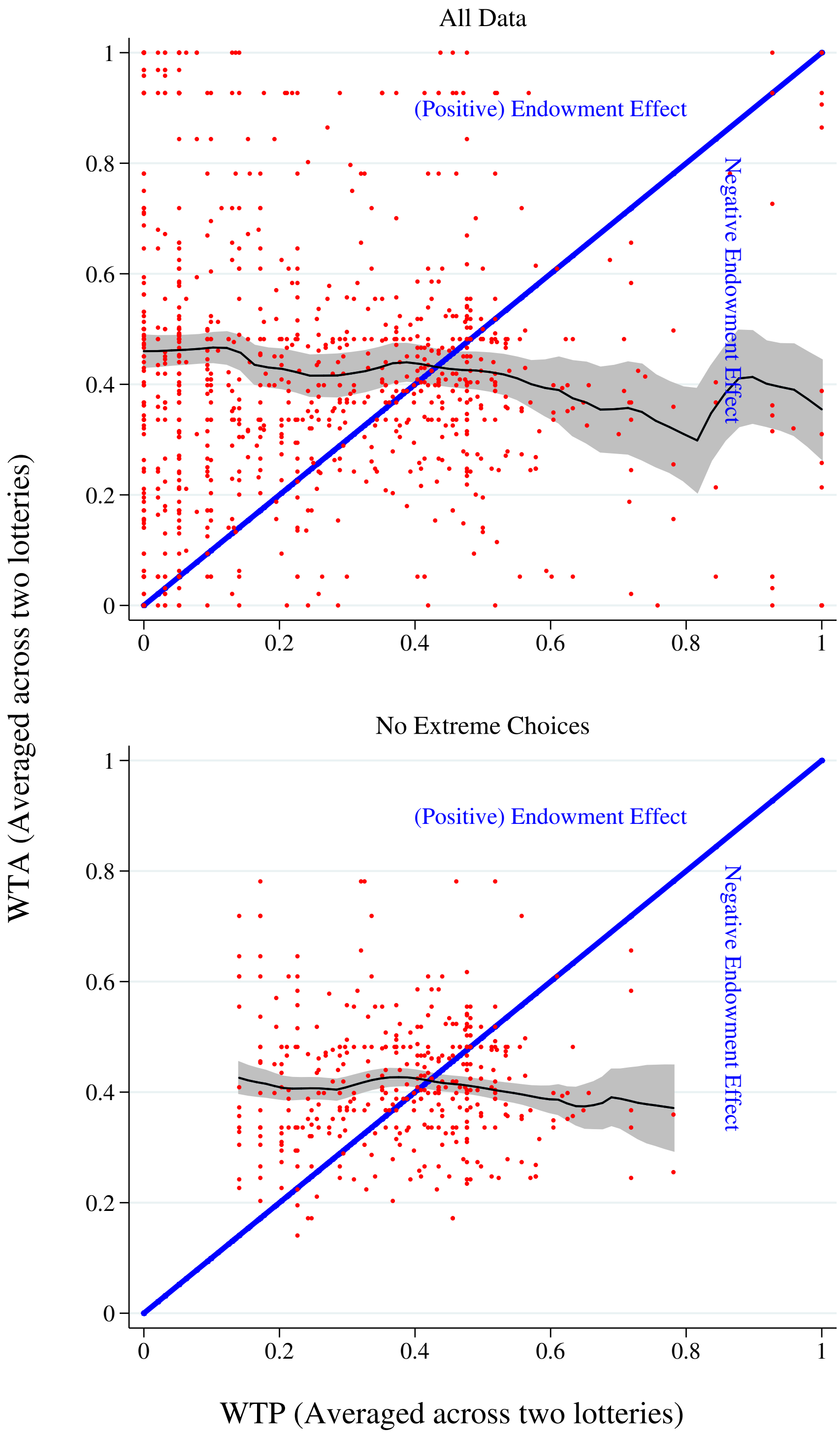
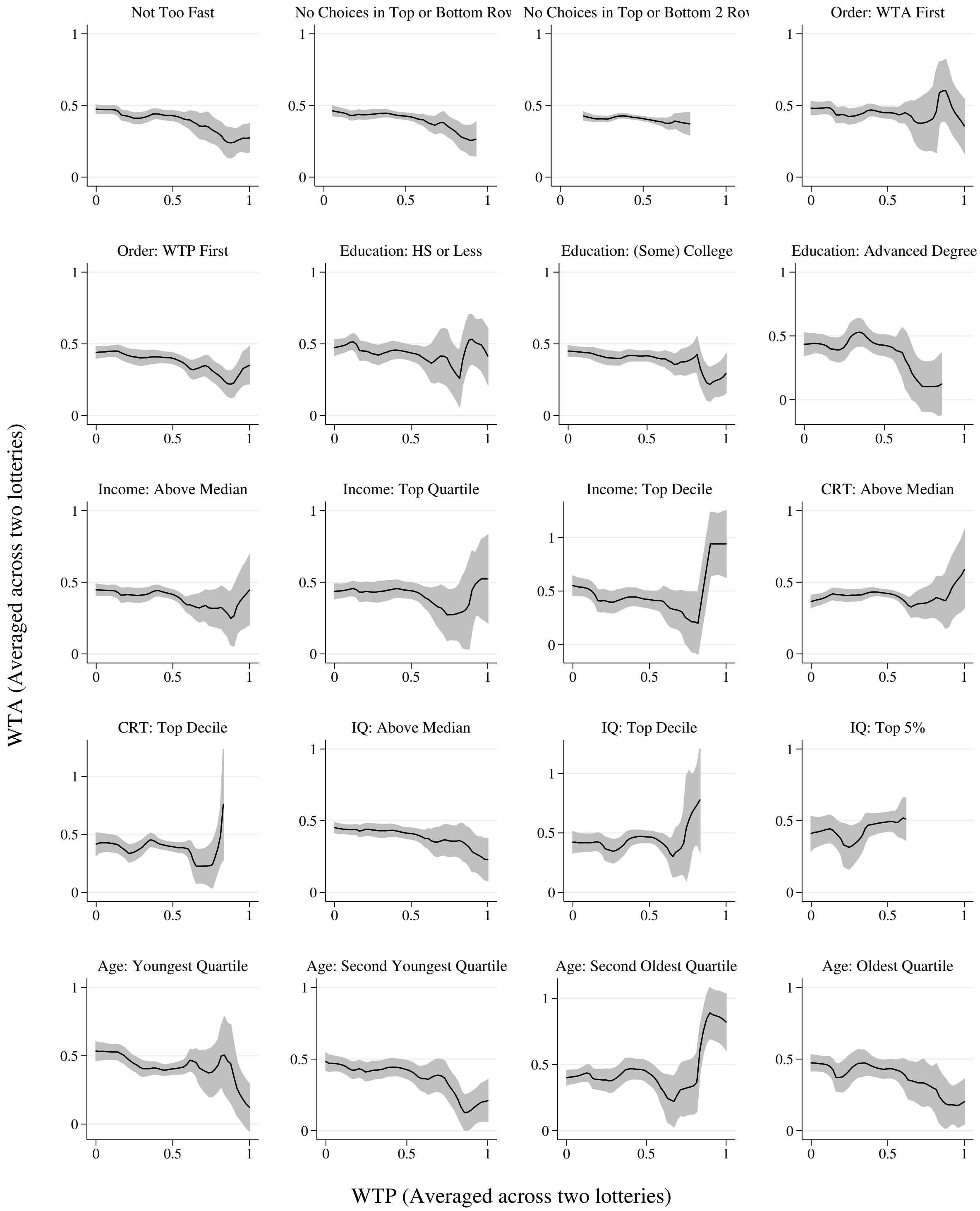


Table B.3: Table 4 for Study 2 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	1,000	−0.09*	−0.06	−0.09	−0.11
		(.051)	(.055)	(.057)	(.068)
Not Too Fast	900	−0.12***	−0.09*	−0.13***	−0.16***
		(.045)	(.050)	(.049)	(.059)
No Dominated Choices	672	−0.10*	−0.13**	−0.13**	−0.15**
		(.057)	(.059)	(.057)	(.067)
No Switches in Top or Bottom Two Rows	672	−0.10*	−0.13**	−0.13**	−0.15***
		(.058)	(.059)	(.057)	(.066)
No Switches in Top or Bottom Three Rows	399	−0.03	−0.11	−0.07	−0.10
		(.098)	(.071)	(.075)	(.11)
Question Order: WTA First	490	−0.06	−0.04	−0.06	−0.06
		(.084)	(.088)	(.091)	(.10)
Question Order: WTP First	510	−0.13**	−0.09	−0.13**	−0.17**
		(.055)	(.060)	(.059)	(.075)
Education: HS or Less	363	−0.08	−0.00	−0.06	−0.07
		(.087)	(.094)	(.10)	(.12)
Education: Some College	520	−0.09	−0.10*	−0.11*	−0.13*
		(.059)	(.062)	(.062)	(.072)
Education: Advanced Degree	117	−0.12	−0.18*	−0.15	−0.18
		(.12)	(.10)	(.11)	(.14)
Income: Above Median	516	−0.04	−0.11*	−0.07	−0.08
		(.062)	(.061)	(.063)	(.075)
Income: Top Quartile	273	0.02	−0.08	−0.01	−0.02
		(.082)	(.079)	(.083)	(.10)
Income: Top Decile	124	−0.08	−0.07	−0.08	−0.10
		(.13)	(.11)	(.13)	(.15)
CRT: Above Median	458	0.08	0.07	0.09	0.11
		(.068)	(.071)	(.071)	(.088)
CRT: Top Decile	99	−0.07	−0.05	−0.05	−0.06
		(.13)	(.14)	(.14)	(.16)
IQ: Above Median	577	−0.09	−0.09	−0.10*	−0.11*
		(.057)	(.059)	(.058)	(.065)
IQ: Top Decile	105	0.10	0.05	0.07	0.08
		(.12)	(.12)	(.12)	(.13)
IQ: Top 5%	49	0.26*	0.17	0.23	0.25
		(.15)	(.15)	(.15)	(.17)
Age: Youngest Quartile	218	−0.27***	−0.25***	−0.30***	−0.39***
		(.082)	(.094)	(.089)	(.11)
Age: Second Youngest Quartile	262	−0.20***	−0.19**	−0.21**	−0.25**
		(.076)	(.086)	(.085)	(.099)
Age: Second Oldest Quartile	263	0.14	0.21*	0.17	0.19
		(.11)	(.11)	(.12)	(.13)
Age: Oldest Quartile	257	−0.11	−0.11	−0.13	−0.16
		(.087)	(.083)	(.088)	(.11)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

Figure B.6: Figure 2 for Study 2 only.



C Plott and Zeiler (2005) Training Rounds

Table C.1: Data from Plott and Zeiler Training Rounds

Study	Group (N)	Lottery	Correlation	WTA < WTP
Plott and Zeiler (2005)	1 (38)	$0.3 * -0.1 \oplus 0.7 * 0.8$	0.15 (.17)	20%
		$0.3 * 1 \oplus 0.7 * 11$	0.26 (.16)	20%
		$0.4 * 1 \oplus 0.6 * 6$	0.53*** (.14)	8%
		$0.5 * -3 \oplus 0.5 * 9$	0.11 (.17)	17%
		$0.7 * 1 \oplus 0.3 * 8$	0.21 (.16)	17%
	2 (36)	$0.3 * 1 \oplus 0.7 * 8$	0.39*** (.16)	29%
		$0.5 * -3 \oplus 0.5 * 9$	0.61*** (.14)	32%
		$0.6 * 1 \oplus 0.4 * 6$	0.20 (.17)	13%
		$0.7 * -0.1 \oplus 0.3 * 0.8$	0.69*** (.13)	21%
		$0.7 * 1 \oplus 0.3 * 11$	0.55*** (.14)	8%

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level. Correlations with standard errors in parentheses.

D Screenshots and Analysis of MPL Ordering

Descriptions of the WTA and WTP questions, as drawn from our design documents, are shown in the text. Here, we display screenshots of the WTA and WTP questions from Study 2. Complete design documents are available at hss.caltech.edu/∼snowberg/wep.html. We also explore the consequences of the different ordering of the MPLs, including varying the order experimentally.

Figure D.1: WTA, Lottery 1.

YouGov

For this question, you are given a lottery ticket that has a **50% chance** of paying you **9,000 points,** and a **50% chance** of paying you **1,000 points**.

You have two options for this lottery ticket:

1. Keep it or
2. Sell it for a certain amount of points (for example, 3,000 points)

For each row in the table below, which option would you prefer?

✓	The lottery ticket	or	Sell it for 0 points
	The lottery ticket	or	Sell it for 1,000 points
	The lottery ticket	or	Sell it for 2,000 points
	The lottery ticket	or	Sell it for 2,500 points
	The lottery ticket	or	Sell it for 3,000 points
	The lottery ticket	or	Sell it for 3,250 points
	The lottery ticket	or	Sell it for 3,500 points
	The lottery ticket	or	Sell it for 3,750 points
	The lottery ticket	or	Sell it for 4,000 points
	The lottery ticket	or	Sell it for 4,250 points
	The lottery ticket	or	Sell it for 4,500 points
	The lottery ticket	or	Sell it for 4,750 points
	The lottery ticket	or	Sell it for 5,000 points
	The lottery ticket	or	Sell it for 5,250 points
	The lottery ticket	or	Sell it for 5,500 points
	The lottery ticket	or	Sell it for 6,000 points
	The lottery ticket	or	Sell it for 7,000 points
	The lottery ticket	or	Sell it for 8,000 points
	The lottery ticket	or	Sell it for 9,000 points
	The lottery ticket	or	✓ Sell it for 10,000 points

Reset Autofill

Review the instructions

Figure D.2: WTA, Lottery 2.

YouGov

For this question, you are given a lottery ticket that has a **50% chance** of paying you **8,000 points**, and a **50% chance** of paying you **2,000 points**.

You have two options for this lottery:

1. Keep it
2. Sell it for a certain amount of points (for example, 3,000 points)

For each row in the table below, which option would you prefer?

[x] The lottery ticket	or	[ ] Sell it for 1,500 points
[ ] The lottery ticket	or	[ ] Sell it for 2,000 points
[ ] The lottery ticket	or	[ ] Sell it for 2,500 points
[ ] The lottery ticket	or	[ ] Sell it for 3,000 points
[ ] The lottery ticket	or	[ ] Sell it for 3,250 points
[ ] The lottery ticket	or	[ ] Sell it for 3,500 points
[ ] The lottery ticket	or	[ ] Sell it for 3,750 points
[ ] The lottery ticket	or	[ ] Sell it for 4,000 points
[ ] The lottery ticket	or	[ ] Sell it for 4,250 points
[x] The lottery ticket	or	[ ] Sell it for 4,500 points
[ ] The lottery ticket	or	[x] Sell it for 4,750 points
[ ] The lottery ticket	or	[ ] Sell it for 5,000 points
[ ] The lottery ticket	or	[ ] Sell it for 5,250 points
[ ] The lottery ticket	or	[ ] Sell it for 5,500 points
[ ] The lottery ticket	or	[ ] Sell it for 6,000 points
[ ] The lottery ticket	or	[ ] Sell it for 7,000 points
[ ] The lottery ticket	or	[ ] Sell it for 8,000 points
[ ] The lottery ticket	or	[x] Sell it for 9,000 points

Reset Autofill

Review the instructions

Figure D.3: WTP, Lottery 1.

YouGov

For this question, you **have been given 10,000 points**. You will be offered the opportunity to exchange some of these points for a lottery ticket. This lottery ticket has a **50% chance** of paying you **9,000 points**, and a **50% chance** of paying **1,000 points**.

For example, if you choose to pay 2,000 points for a lottery ticket, and this question is chosen for payment, you will:

- Pay 2,000 points for the lottery ticket
- Keep 8,000 points for yourself
- Earn whatever proceeds you get from the lottery ticket (if any)

For each row in the table below, which option would you prefer?

☑ Keep 10,000 points	or	☐ Buy the lottery ticket for 10,000 points and keep the remaining 0 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 9,000 points and keep the remaining 1,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 8,000 points and keep the remaining 2,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 7,000 points and keep the remaining 3,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 6,000 points and keep the remaining 4,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 5,500 points and keep the remaining 4,500 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 5,250 points and keep the remaining 4,750 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 5,000 points and keep the remaining 5,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 4,750 points and keep the remaining 5,250 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 4,500 points and keep the remaining 5,500 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 4,250 points and keep the remaining 5,750 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 4,000 points and keep the remaining 6,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 3,750 points and keep the remaining 6,250 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 3,500 points and keep the remaining 6,500 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 3,250 points and keep the remaining 6,750 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 3,000 points and keep the remaining 7,000 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 2,500 points and keep the remaining 7,500 points
☐ Keep 10,000 points	or	☐ Buy the lottery ticket for 2,000 points and

Examining the screenshots above shows that if a user always clicked, say, the third from

Figure D.4: WTP, Lottery 2.

YouGov

For this question, you **have been given 9,000 points**. You will be offered the opportunity to exchange some of these points for a lottery ticket. This lottery ticket has a **50% chance** of paying you **8,000 points**, and a **50% chance** of paying **2,000 points**.

For example, if you choose to pay 3,000 points for a lottery ticket, and this question is chosen for payment, you will:

- Pay 3,000 points for the lottery ticket
- Keep 6,000 points for yourself
- Earn whatever proceeds you get from the lottery ticket (if any)

For each row in the table below, which option would you prefer?

☑ Keep 9,000 points	or	☐ Buy the lottery ticket for 9,000 points and keep the remaining 0 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 8,000 points and keep the remaining 1,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 7,000 points and keep the remaining 2,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 6,000 points and keep the remaining 3,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 5,500 points and keep the remaining 3,500 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 5,250 points and keep the remaining 3,750 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 5,000 points and keep the remaining 4,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 4,750 points and keep the remaining 4,250 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 4,500 points and keep the remaining 4,500 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 4,250 points and keep the remaining 4,750 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 4,000 points and keep the remaining 5,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 3,750 points and keep the remaining 5,250 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 3,500 points and keep the remaining 5,500 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 3,250 points and keep the remaining 5,750 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 3,000 points and keep the remaining 6,000 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 2,500 points and keep the remaining 6,500 points
☐ Keep 9,000 points	or	☐ Buy the lottery ticket for 2,000 points and keep the remaining 7,000 points
☐ Keep 9,000 points	or	☑ Buy the lottery ticket for 1,500 points and

the top box in each MPL it would induce a negative correlation between WTA/P. This is not a particular concern due to the fact that results are robust to excluding those that give extreme answers, as these people are the most likely to follow such a pattern. However, here we conduct more extensive checks that control explicitly for the average position a participant chooses on non-risk related MPLs, including analyzing the data of an additional large scale study that reversed the order of the MPL for WTP.

In particular, as the questions about WTA and WTP are far apart in the study, any issue caused by people picking different points on an MPL would be more likely due to a tendency, rather than an explicit position chosen on each MPL. Thus, if we control for such a tendency, this should get rid of spurious correlation (or lack thereof) created by the ordering of the MPLs. To do so, we need a measure of this tendency. This is easily obtained from the other MPLs in our studies, especially Study 2, which contains a number of MPLs that measure something other than risk attitudes.[1] We use, as a control for this tendency, the location of the switching point on six MPLs unrelated to risk in Study 2: two regarding time preference, and four regarding social preferences (two in the advantageous domain and two in the disadvantageous domain). For each of these MPLs we identify the first row in which the participant clicked the right hand side the MPL (if individuals never switched they are assigned a value of the last row number plus one).

We control for a general tendency to select answers in higher or lower spots on an MPL in several ways in Table D.1. All columns examine the correlation between the Average WTA and Average WTP measures from Study 2 (where we have the most controls for MPL position). The first column shows the unconditional correlation from a standardized regression. The next four columns enter information about other choices in various ways. The second column includes the first three principal components of the switching rows. The first principal component is essentially an average, and the others contain more information

[1] Risk measures are correlated with WTA and WTP for lotteries—see Section 6. Thus, including the MPL positions on those measures would add error to the average position variables, as some of the variation in that variable will represent real correlations between risk measures and WTA/WTP.

about choices in these other MPLs. Together, the first three principal components capturing 85% of the variation in MPL switching point location. The third column breaks each of these three components into deciles, and then enters a dummy variable for each decile. This is 30 dummy variables in all. The fourth column breaks the first principal component into 100 percentile bins, and enters a dummy for each. Both of these allow for a more non-parametric dependence of the correlation on average choice. The final column enters a dummy variable for each possible switching position in each of the six MPLs. Across all columns and both panels the pattern is clear: the partial correlation barely moves no matter how we try to control for the average position a participant takes on other MPLs.

We can look at this issue one other way. Those that are most wedded to a given position on the MPL will have a lower standard deviation of switching points. If switching points are just random, with a person-specific parameter deciding where on the MPL they chose to switch, then we should see a positive relationship between the standard deviation of switching points and the correlation between WTA and WTP.

Figure D.5 shows the correlation between WTA and WTP as a function of the standard deviation of MPL switching points. To produce the figure, we generate a variable for person i that describes their contribution to the correlation in their percentile p as

$$\frac{(\mathrm{WTA}_i - \overline{\mathrm{WTA}}_p)(\mathrm{WTP}_i - \overline{\mathrm{WTP}}_p)}{(\mathrm{Var}[\mathrm{WTA}_p]\mathrm{Var}[\mathrm{WTP}_p])^{\frac{1}{2}}}.$$

This can then be plotted, non-parametrically, versus the percentile of the standard deviation. The black lines indicate the non-parametric plot, and the grey bar indicates the 95% confidence intervals.

The first panel of Figure D.5 does not smooth the correlations across percentiles. As such, there does not seem to be an apparent pattern. Therefore, in the second panel, we smooth the non-parametric plot. As can be seen, those that have very little variation in their MPL switching point do, indeed, exhibit a negative correlation between WTA and

Table D.1: Partial correlations controlling for average MPL switching position on non-risk questions.

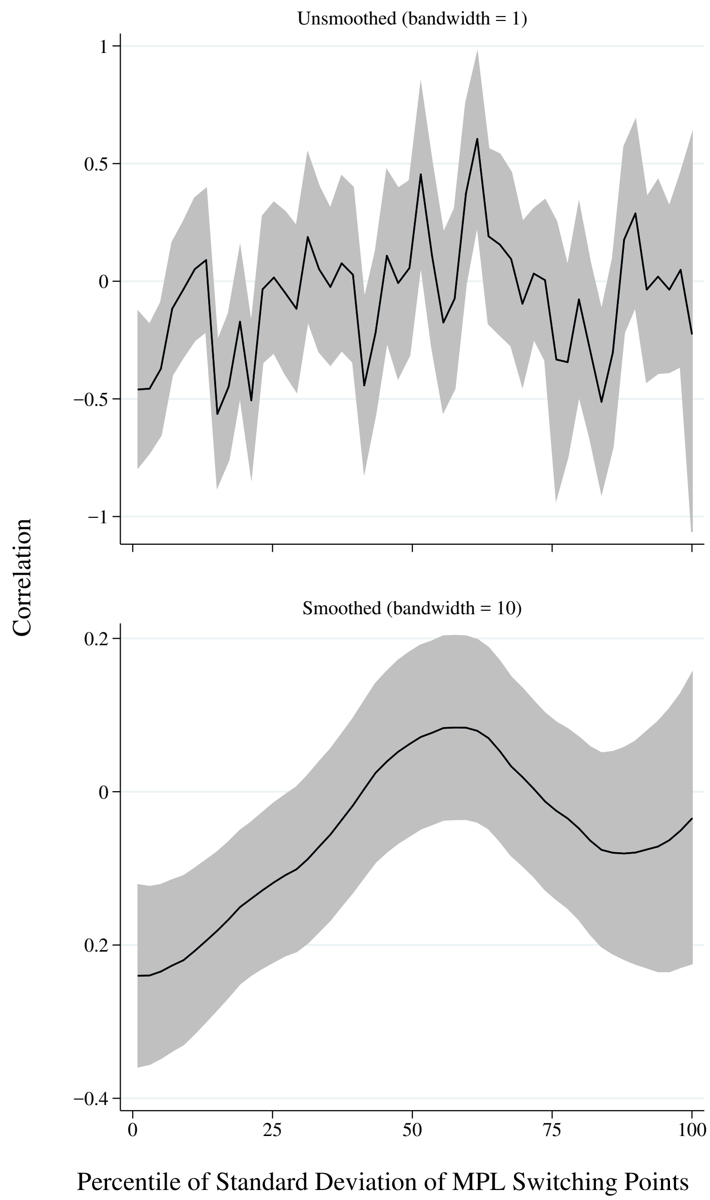
Dependent Variable: Average WTP					
Average WTA	−0.09 (.057)	−0.04 (.052)	−0.04 (.047)	−0.05 (.046)	−0.05 (.037)
Three Principal Components		Y			
Deciles of First 3 Principal Components			Y		
Percentiles of First Principal Component				Y	
Indicators for Switching Point in six Questions					Y
Dependent Variable: Average WTA					
Average WTP	-0.09 (.059)	-0.04 (.056)	-0.05 (.051)	-0.06 (.049)	-0.05 (.040)
Three Principal Components		Y			
Deciles of First 3 Principal Components			Y		
Percentiles of First Principal Component				Y	
Indicators for Switching Point in six Questions					Y

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with standard errors in parentheses.

WTP. However, above the 25th percentile, there is a non-monotonic relationship between the correlation of WTA/WTP and the standard deviation of MPL switching points. Indeed, the non-parametric curve never exceeds 0.07, and the 95% confidence interval never exceeds 0.2. The average correlation above the 25th percentile is 0.00. Thus, any effect of MPL ordering on our results is likely to be quite small.

As a final test, as part of another study, a subset of the coauthors experimentally varied whether or not WTP was elicited using the standard ordering, featured throughout this

Figure D.5: Correlation as a function of standard deviation of switching points.



paper, or the opposite ordering. This study was administered to a non-representative subset of the British population (Carvalho et al., 2017). Analysis of this data, in the style of Table 2, is shown in Table D.2.

There are a number of salient features of the data. First, the correlations using the standard ordering are about 0.1 to 0.15 higher than found in our study. As discussed in Section 5, this difference is likely due to a different population being studied. Second, using

Table D.2: Reversing the order of the WTP MPL.

		Correlation between WTA and WTP				Correlation within Type	
	N	Lottery 1	Lottery 2	Averages	ORIV	WTA	WTP
Standard Order	1,037	0.05 (.035)	0.03 (.038)	0.05 (.036)	0.06 (.044)	0.69*** (.030)	0.72*** (.031)
Reverse, Order (WTP)	953	0.15*** (.039)	0.18*** (.039)	0.18*** (.037)	0.22*** (.046)	0.74*** (.029)	0.64*** (.041)
All Data	1,990	0.10*** (.027)	0.10*** (.027)	0.11*** (.026)	0.13*** (.031)	0.71*** (.021)	0.69*** (.025)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with bootstrapped standard errors from 10,000 simulations in parentheses.

the alternate ordering increases correlations by 0.1 to 0.15. Finally, as there does not seem to be a "correct" ordering for these MPLs, the row gives what is essentially an average of the first two rows. A final point worth noting is the distribution of endowment effects is virtually unchanged by reversing the order of the WTP MPL. Thus, had we altered the ordering of the WTP MPL in our studies, it is likely that the correlation we observed would be much closer to zero, while other results would remain virtually unchanged.

However, the effect of reversing the MPL order seems to not just change answers, it seems to change the composition of the people completing the survey. Although each ordering was administered to 50% of the sample population, about 10% fewer people completed the alternate version. This seems to be because people found the alternate version so confusing they dropped out of the study altogether. An additional piece of evidence is the much lower correlation between WTP measures when using the alternative ordering. This suggests that an increase of 0.15 in the correlations due to changing the order of the MPL is an upper bound, as some of this change is likely due to a change in the composition of the people completing the survey.

E Histogram of the Endowment Effect

Figure E.1: Histogram of the Endowment Effect for WTA/P data in Figure 1