CESifo WORKING PAPERS

6492
2017

May 2017

Willingness-To-Pay and Willingness-To-Accept are Probably Less Correlated than You Think

Jonathan Chapman, Pietro Ortoleva, Erik Snowberg, Colin Camerer, Mark Dean

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6492*
Category 13: Behavioural Economics

Willingness-To-Pay and Willingness-To-Accept are Probably Less Correlated than You Think

Abstract

Willingness to pay (WTP) and willingness to accept (WTA) a monetary amount for a lottery should be closely related. In data from an incentivized survey of a representative sample of 3,000 U.S. adults, we find that WTA and WTP for a lottery are, at best, weakly correlated. Across all respondents, the correlation is slightly negative. For the subgroups that we examine, the correlation is almost always small, typically statistically insignificant, and often negative. The exception is respondents who score highly on a within-study IQ test, where the correlation is around 0.2. A meta-study of similar lab experiments with university students also shows a correlation of around 0.15-0.2. While poorly related to each other, our measures of WTA and WTP are strongly related to different measures of risk aversion, and relatively stable across time. These various patterns allow us to show that this lack of relationship between WTA and WTP is compatible with existing theories, such as Prospect Theory and Stochastic Reference Dependence, only under very specific, and unlikely, correlational structures between parameters. We suggest a simpler formalization.

Keywords: willingness to pay, willingness to accept, endowment effect, loss aversion, risk aversion.

Jonathan Chapman
NYUAD / Abu Dhabi / UAE
jchapman@nyu.edu
www.jnchapman.com

Pietro Ortoleva
Columbia / New York / NY / USA
pietro.ortoleva@columbia.edu
www.columbia.edu/~po2205/

Erik Snowberg
Caltech / Pasadena / CA / USA
erik.snowberg@ubc.ca
hss.caltech.edu/~snowberg/

Colin Camerer
Caltech / Pasadena / CA / USA
camerer@hss.caltech.edu
hss.caltech.edu/~camerer/

Mark Dean
Columbia / New York / NY / USA
mark.dean@columbia.edu
www.columbia.edu/~md3405/

May 15, 2017
We thank Douglas Bernheim, Benedetto De Martino, Graham Loomes, PeterWakker, MichealWoodford, and the participants of seminars and conferences for their useful comments and suggestions. Evan Friedman and Khanh Ngoc Han Huynh provided excellent 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] A large experimental literature finds this is not the case: WTA is usually larger than WTP, a phenomenon called the *endowment effect*.[2] This finding can be rationalized by theories involving reference-dependent preferences, in particular, Prospect Theory (Kahneman and Tversky, 1979). In these theories, the value of the object depends on whether someone possesses it.

A common feature of both standard and prospect theory is that WTA and WTP should be related. We find, instead, that WTA and WTP for lottery tickets are, at best, weakly correlated. The correlation between WTA and WTP is almost always small, typically insignificant, and often negative for the subgroups we examine. This is despite the fact that there is an endowment effect (WTA > WTP) for a majority of participants. Our results come from three large, incentivized, and representative surveys of U.S. adults totaling 3,000 people. Each of these surveys 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. 2015). This does not change the results. Nor does excluding respondents that give extreme answers. Nor do a number of other refinements. The only statistically robust correlation (of 0.2–0.3) we observe is for those in the top 5% of an in-survey IQ test. Extending this to the top 10% reduces the correlation by two-thirds, and renders it statistically insignificant.

Our first survey consisted of 2,000 respondents. This was designed to explore the relationships between a wide variety of behavioral and political measures over time (Camerer et al., 2017a). Our second survey re-interviewed the respondents six months later. Of the original 2,000 respondents, 1,466 completed the second survey. The surprising relationship between WTA and WTP, and a few other factors, encouraged us to run a second study of 1,000 independent respondents (Camerer et al., 2017b). Together, these data allow us to doc-

[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) for reviews.

ument relevant facts precisely, even for small sub-groups. Additionally, repeated elicitation from the same people allows us to document that WTA, WTP, and the endowment effect for the two lottery tickets are relatively stable over six months. The cross-time correlation is around 0.3, comparable to risk-aversion and time preferences. All of our elicitations use the Multiple Price List (MPL) technique.

To understand the relationship of our results with the literature, we examine the correlation of WTA and WTP in data from existing experiments that collected within-subject measures. There are a very small number of studies that satisfy this criteria, and none of them examine the correlation between WTA and WTP. The three studies we identify had five groups of participants between them, with a total of 365 participants. The weighted average correlation of these pairs is around 0.15–0.2, depending on the treatment of those expressing dominated values for the lotteries. This is consistent with the correlation we observe in our data for similar respondents—those with very high IQs. Notably, all three of these studies use the BDM method of Becker et al. (1964) to elicit valuations, suggesting that our results are not simply an artifact of the elicitation technique we use.

As our measures of WTA and WTP concern lottery tickets, we investigate the relation between these measures and risk aversion. Our second study contained six additional measures of risk aversion. These measures can be broken into two groups: those with a fixed lottery on the left-hand-side of the MPL (like WTA), and those involving a variable lottery. We observe that WTA is strongly correlated with the risk measures featuring a fixed lottery. Moreover, these measures are highly correlated with each other, but poorly correlated with those with a variable lottery and WTP. Poor correlations between risk measures have been observed in many experiments.[3] However, our results suggest there is a structure to this pattern of correlations.

[3]See Isaac and James (2000); Berg et al. (2005); Bruner (2009); Deck et al. (2010); Harbaugh et al. (2010); Deck et al. (2013); Nielsen et al. (2013); Loomes and Pogrebna (2014); Lönnqvist et al. (2015); Crosetto and Filippin (2016); He et al. (2016). For a review of the relevant literature in psychology, see Weber and Johnson (2008). Gillen et al. (2015) suggest that many of these findings may be due to measurement error, and we adopt their techniques to rule this out in our case.

This pattern of correlations can be easily structured by viewing the left-hand-side of the MPL as an endowment, as promoted by Sprenger (2015).[4] Our data provides the first direct evidence for this view by showing how the fixed element of the MPL generates measures that are strongly correlated with those obtained by giving an explicit reference point: the WTA and WTP questions.

The patterns we discover are difficult to reconcile with existing theories. There are two ways to reconcile them with Prospect Theory. First, by specifying a specific correlation between loss aversion and risk aversion. Unfortunately, the correlation needed to rationalize these findings seems to be the opposite of that found in preliminary data. Second, by specifying that people treat buying and selling in different ways, breaking the clear link between these activities and gains and losses frames in standard instantiations of Prospect Theory. Stochastic Reference Dependence (Köszegi and Rabin, 2006) fares better, but also relies on a specific correlational structure between risk and loss aversion. This specific pattern must hold for every subgroup we examine, and for every lottery ticket. While this is possible, it seems unlikely.

We also explore a simple theory that could generate the endowment effect as well as our other observations. We posit that WTA and WTP for lotteries represent the outcomes of two different risk-preference mechanisms. In the population, these preferences are independently distributed, but not from identical distributions. This is consistent with findings in neuroscience that different parts of the brain are involved with the evaluation of selling and buying (De Martino et al., 2009). The distribution of WTA has a mean greater than that of WTP, generating the endowment effect for most people. But, for some people, WTP will be greater than WTA (25–30% in our data). These processes also manifest themselves in other risk elicitations. How they manifest depends on whether the framing causes participants to view the choice as buying a lottery ticket or selling one.

[4] See also Bleichrodt et al. (2001), and the many references therein.

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 respondents. 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 match them, on observables, to members of the 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. We use these weights throughout the paper. Unweighted results are in Appendix A.

All surveys were incentivized. That is, respondents were paid based on outcomes associated with their choices. Their choices were used to elicit a large number of attitudes, including WTA and WTP (see Camerer et al., 2017a,b, for more details on these other questions). Two of these choices were selected for payment after the respondent completed the entire survey.[5]

As our study was run online, WTA and WTP could not be elicited for physical goods. Instead, we elicited these quantities for lottery tickets—although we could not physically give the ticket to respondents. However, the literature shows that the endowment effect is still present when using abstract lotteries rather than physical goods (see, for example, Isoni et al., 2011).

All outcomes were expressed in points. This is an internal YouGov currency used to pay participants. It can be converted to US dollars, or prizes, using the approximate rate of $0.001 per point. The average payment to respondents was around $9 (9,000 points). The survey took participants between 45 minutes and an hour. This amount is quite high by internet survey standards, and represents a rate of pay approximately three times the

[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). However, even if respondents fail to reduce compound lotteries, this is still incentive compatible (Segal, 1990). 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).

average rate for YouGov surveys.

The ordering of many of the questions, including WTA and WTP, was randomized. 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 1,466 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. When combining these studies, we only use the first wave of the first study to ensure independence of the observations. This gives us a total population of 3,000 participants, with two observations for approximately half of them.

2.1 Measuring WTA and WTP

Each survey 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 6). Using two lottery tickets allows us to correct for measurement error, as discussed in Section 2.2. A qualitative question was asked between the elicitations of WTA, and also between the elicitations of WTP. Approximately half of the subjects 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).

Each elicitation is performed using a multiple price list (MPL) (Holt and Laury, 2002).[6] An MPL consists of a table with two columns of outcomes. In each row, the respondent is asked asked to make a choice between the outcomes in the columns. One column contains the

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

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. Respondents 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 this point, the respondent should also choose the latter option.[7] To increase respondents' understanding, the first and the last row of the MPL always involved a dominated option—for example, 200 points instead of a lottery that pays 200 or 0 with equal probability—with the undominated option pre-selected.[8]

The elicitation of WTA had the following form:[9]

> 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).

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

The lottery ticket had one of two payment schemes. Table 1 describes the lotteries used in each study. Both waves of Study 1 used the same lotteries. In Study 2, neither lottery included the possibility of receiving 0. This was included as a robustness check. Note that all lotteries are mean preserving spreads of each other.

[7]Respondents who do not understand the question may switch back and forth between columns multiple times. However, in our implementation, respondents 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 limits our ability to study other types of behaviors that may be indicated by multiple switches, but those are not our focus here or elsewhere.

In practice, very few instances of multiple switches were observed. This is due to extensive training before the respondent answered their first MPL.

[8]We used the average value of the varying option over the two rows where the respondent's choices switched columns as the value for the lottery ticket. Using the minimum or maximum value does not alter results.

[9]This is taken from the design documents. These, and screenshots of the surveys, may be found at http://people.hss.caltech.edu/ snowberg/wep.html.

Table 1: Lotteries Used

	Lottery 1	Lottery 2
Study 1, Wave 1 & 2	50% 10,000; 50% 0	50% 8,000; 50% 2,000
Study 2	50% 9,000; 50% 1,000	50% 8,000; 50% 2,000

For WTP, the same lottery tickets were used, but respondents 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 then 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", where x changes with the row.

In order to identify subjects that chose dominated options without carefully reading the question, an additional design change was made to Study 2. In particular, we included *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 respondents. 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 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. If a respondent 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, 81% of those alerted to a dominated choice did not change it. We present some results below which exclude those who chose a dominated option.

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 will reduce the bias in estimated coefficients by approximately $1/\sqrt{2}$. This will also tend to reduce standard errors, which, coupled with our large sample size, should increase confidence in our results.

Second, we can use the ORIV technique (Gillen et al., 2015). This adapts an errors-in-variables instrumental variables (IV) approach to the specific form of our data. The main difficulties 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 estimates:

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

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

clustering the standard errors by participant. 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. (2015) shows that this produces consistent estimates of Corr[WTA*, WTP*].

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 negligible. That is, observing a high willingness to pay for a lottery ticket conveys very little information about willingness to accept.This is true for all three of our surveys. Second, the 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. Once again, this is true across all of our surveys.

Table 2: Simple analysis of 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)	-0.10^{**} (.047)	0.71^{***} (.023)	0.74^{***} (.028)
Study 1, Wave 2	-0.01 (.050)	-0.02 (.049)	-0.02 (.054)	-0.02 (.069)	0.67^{***} (.034)	0.79^{***} (.024)
Study 2	-0.09^{*} (.051)	-0.06 (.056)	-0.09 (.058)	-0.11 (.071)	0.70^{***} (.036)	0.75^{***} (.040)

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

The remaining two columns of Table 2 take steps to reduce concerns about the role of measurement error in our results. First, we normalize, and then average, the measures of WTA and WTP.[10] We then examine the correlation between the two resultant measures. The magnitudes of the correlations grow slightly more negative, consistent with a reduction in attenuation bias.

Next, we implement the ORIV procedure described in Section 2.2. As ORIV produces consistent measures of correlations, this reduces attenuation bias even further (to zero, in large samples). Once again, the negative correlations grow in magnitude. However, the statistical significance of the correlations goes *down*. This occurs because, like all IV procedures, ORIV increases standard errors. Thus, we show estimates based both on averages and ORIV so the reader can be assured that statistical insignificance is not an artifact of larger standard errors.

[10]Normalizing 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 respondents 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%	23%	16%	61%
Study 2	28%	12%	59%	24%	14%	62%

3.1 The Endowment Effect

Our data exhibit an endowment effect (WTA > WTP) in-line with existing laboratory data. Across the studies, the average WTA was 89% of the expected value of the lottery, with a median of 83% and standard deviation of 31%. The mean WTP was slightly lower at 69%, with a median of 63% and standard deviation of 27%. The WTP also has a higher proportion of subjects reporting extremely low values—36% of subjects reported a WTP of less than half of the expected value, compared to 14% for WTA. As this may present a concern for some readers, we will return to this point in Section 3.3 and Section 4.

As shown in Table 3, the endowment effect is consistent across surveys. Approximately 60% of respondents exhibit an endowment effect, and a further 10% exhibit no endowment effect. These figures are in line with previous studies. Although, as noted in the introduction, there are very few studies that elicit WTA and WTP from the same participant. The final 30% of respondents exhibit a negative endowment effect. This is difficult, though not impossible, to reconcile with prospect theory. Doing so relies on being loss loving (rather than loss averse).

The existence of a negative endowment effect is not likely to be due solely to measurement error. The correlation between the two different measures of the endowment effect (corresponding to each of the two lottery tickets) is around 0.75, which is quite high. 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 measured 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 others. Thus, there is likely a bit more noise among respondents with a negative endowment effect, but only a bit.

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 respondents, 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 1,466 of the original 2,000 respondents 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 0.27 and 0.24 across surveys. For WTP, the correlations are 0.28 and 0.24.[11] Using ORIV produces correlation estimates of 0.36 for WTP, and 0.35 for WTP. All of these correlations are statistically significantly different from zero, but not from each other.

The stability of WTA and WTP are 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 0.34 and 0.38 respectively. A third was measured only once per survey, and has a correlation of 0.32 across time. This is similar to the stability of measures of risk preferences over about the same amount of time in Caltech undergrads (Gillen et al., 2015).[12] Study 2 also two measures of ambiguity aversion (ORIV correlation 0.22 within-subject, across time), time preferences (ORIV correlation 0.33), and amount given in the dictator game (ORIV correlation 0.49).

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. Second, the endowment effect has a (within-participant) correlation of 0.33 across time. Third, we can examine the correlation for WTA and WTP on the subset of people who gave

[11]Recall from Table 1 that both waves employed the same lotteries. Unweighted correlations are 0.27 and 0.24 for WTA, and 0.28 and 0.24 for WTP.

[12]This study did not elicit WTA and WTP.

the same valuations in both Wave 1 and Wave 2. For the first lottery, this restricts the sample to 96 participants.[13] The correlation between WTA and WTP for these respondents is 0.11 ($p = 0.27$) for the first lottery, and 0.02 ($p = 0.84$) and 0.03 ($p = 0.77$) for the second lottery in the first and second waves, respectively two waves. For the second lottery, the sample is 115 participants. The correlation between WTA and WTP for these participants is -0.02 ($p = 0.78$) in the second lottery, and 0.03 ($p = 0.71$) and 0.01 ($p = 0.89$) 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 these 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. This is done in the first panel of Figure 1. Here we show all of our data on average WTA and WTP for both studies.[14] 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.

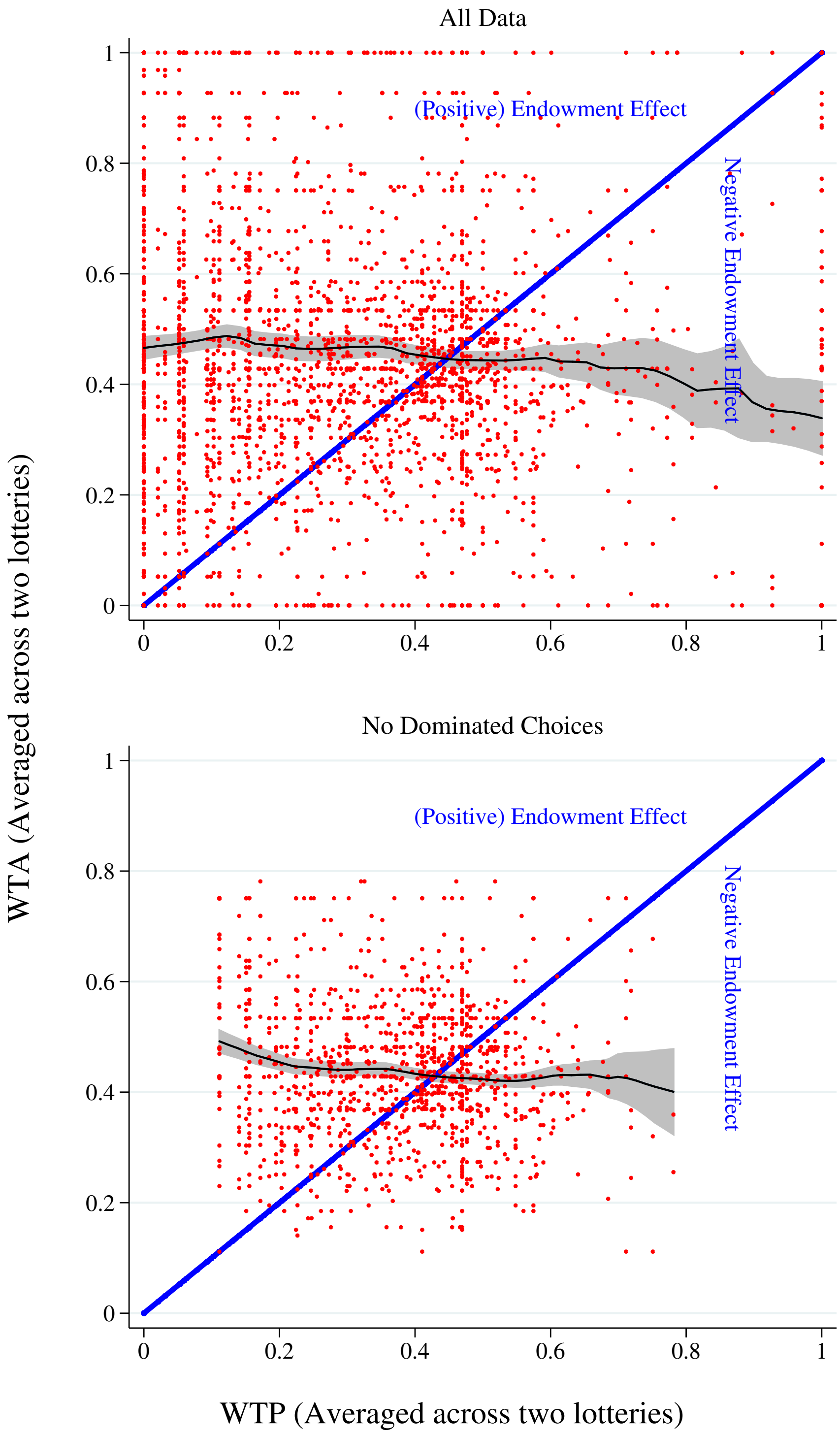
New patterns emerge from the figure as well. First, to the extent that there is any correlation between WTA and WTP, it is due to participants with above risk-neutral values expressed for WTP. As this describes only ∼25% of participants, this negative relationship is weak, as indicated by the expanding standard errors. Second, there is wide variation in both WTA and WTP across the spectrum.

This wide variation has an apparent downside as well: many participants give extremely high or extremely low values for WTA and / or WTP. Thus, in the second panel of Figure

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

[14] This excludes Wave 2 of Study 1, as these are repeated observations of respondents from Wave 1. All results are shown separately by survey in Appendix B.

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



1, we remove all subjects 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 of the data in order to assess whether our finding is an artifact of a particular sample, or more general.

4 Subgroups

One may believe that a correlation between WTA and WTP would be restored once we focus on specific subgroups: for example, subjects with higher IQ, or higher education, or subjects 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 3000 independent observations.[15]

Most of the subgroups in the table need no explanation. However, a few do. Similar to our approach in Figure 1, we tried a few different ways of removing "extreme" choices. Two are displayed in the table: one removes respondents whose switching point was just before the first selectable item or just after the last; the second removes subjects if their switching point was in the top three or bottom three rows. In addition, "Not too fast" removes the 10% of respondents who completed the survey in the least amount of time.

In almost all subgroups, correlations are small in magnitude and / or negative. However, those with high CRT or IQ scores have correlations between WTA and WTP in the range of 0.1–0.2. In both studies, we administered the Cognitive Reflection Test of Frederick (2005). Around 50% of respondents got one or more of the three questions correct, representing all those above the median. Approximately 10% got all three questions correct, which is the

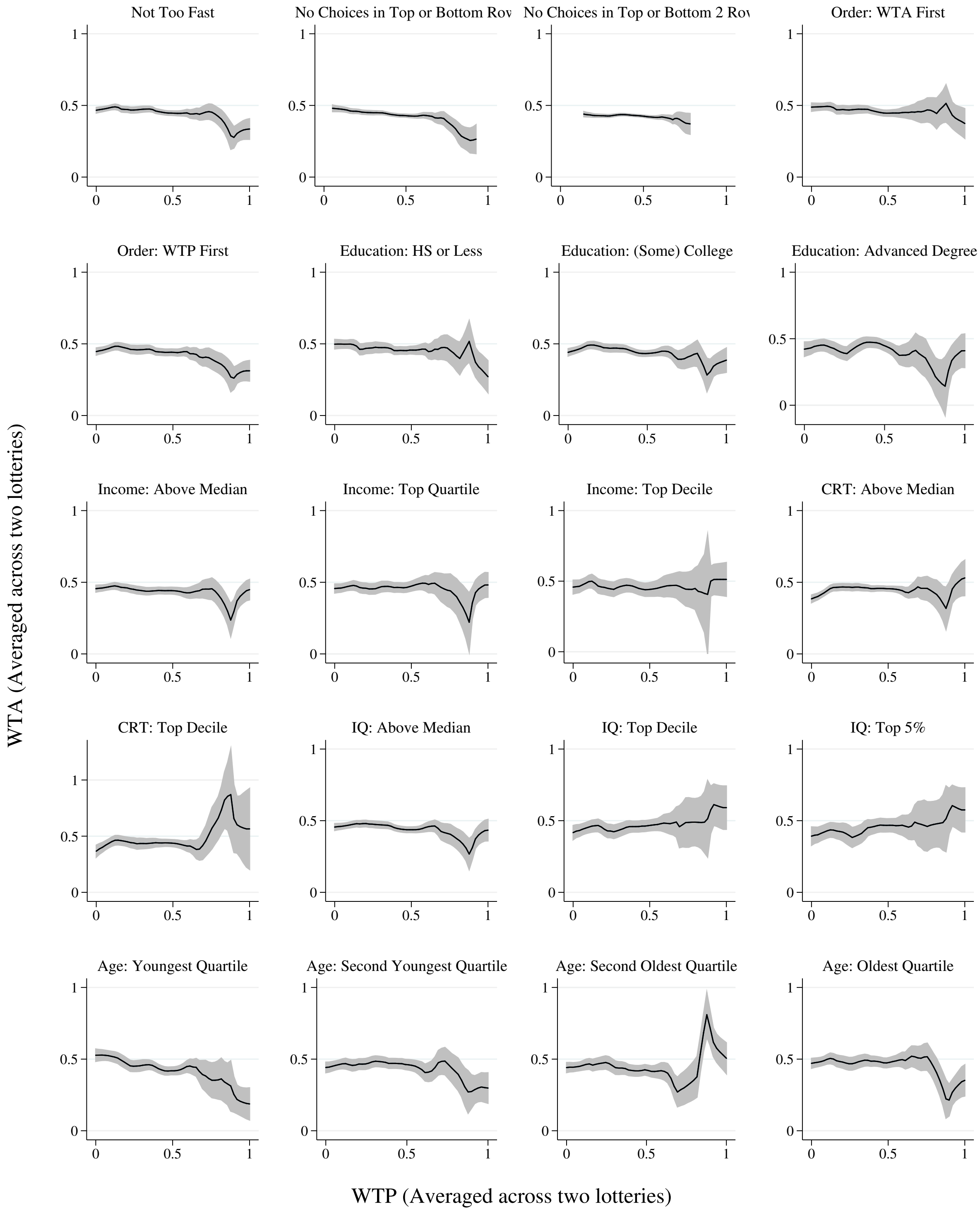
[15] Wave 2 consists of the same respondents as Wave 1, hence the observations are not independent. For results by survey, see Appendix B.

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)	(.032)	(.031)	(.039)
No Dominated Choices	1,322	−0.09*	−0.13***	−0.15***	−0.19***
		(.052)	(.042)	(.044)	(.063)
Not Too Fast	2,700	−0.06*	−0.06**	−0.07**	−0.10**
		(.031)	(.031)	(.031)	(.039)
No Switches in Top or Bottom Two Rows	1,595	−0.11**	−0.14***	−0.14***	−0.19***
		(.043)	(.043)	(.042)	(.053)
No Switches in Top or Bottom Three Rows	1,028	−0.05	−0.05	−0.05	−0.11
		(.058)	(.050)	(.056)	(.084)
Order: WTA First	1,501	−0.06	−0.06	−0.07	−0.09
		(.042)	(.046)	(.043)	(.056)
Order: WTP First	1,499	−0.08*	−0.07*	−0.09**	−0.13**
		(.043)	(.043)	(.044)	(.055)
Education: HS or Less	1,128	−0.11**	−0.05	−0.11**	−0.15**
		(.049)	(.056)	(.053)	(.068)
Education: Some College	1,538	−0.03	−0.06*	−0.05	−0.07
		(.040)	(.036)	(.039)	(.049)
Education: Advanced Degree	334	−0.01	−0.09	−0.04	−0.06
		(.074)	(.061)	(.069)	(.091)
Income: Above Median	1,520	−0.03	−0.06	−0.04	−0.06
		(.037)	(.040)	(.038)	(.047)
Income: Top Quartile	813	−0.01	−0.01	0.01	0.00
		(.054)	(.058)	(.054)	(.068)
Income: Top Decile	399	0.06	−0.04	0.01	0.02
		(.072)	(.068)	(.073)	(.092)
CRT: Above Median	1,371	0.11***	0.05	0.09**	0.12**
		(.043)	(.041)	(.043)	(.055)
CRT: Top Decile	288	0.10	0.11	0.10	0.11
		(.085)	(.084)	(.085)	(.11)
IQ: Above Median	1,694	−0.03	−0.06*	−0.05	−0.07
		(.038)	(.035)	(.037)	(.047)
IQ: Top Decile	337	0.10	0.07	0.09	0.11
		(.090)	(.076)	(.087)	(.11)
IQ: Top 5%	146	0.28***	0.13	0.20**	0.28**
		(.093)	(.10)	(.10)	(.13)
Age: Youngest Quartile	611	−0.20***	−0.23***	−0.26***	−0.33***
		(.069)	(.062)	(.068)	(.085)
Age: Second Youngest Quartile	798	−0.03	−0.04	−0.03	−0.05
		(.056)	(.056)	(.053)	(.067)
Age: Second Oldest Quartile	783	−0.02	0.04	0.00	−0.01
		(.058)	(.068)	(.062)	(.079)
Age: Oldest Quartile	808	−0.04	−0.05	−0.04	−0.07
		(.049)	(.053)	(.049)	(.063)

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

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



upper decile. We also administered an in-survey IQ test to all respondents. This IQ 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. Answering all six questions correctly was only achieved by 5% of respondents.[16] The top decile represents five or more correct answers, and above the median is two or more correct answers.

The correlations for those in the upper decile of CRT and IQ scores is in the range of 0.1–0.2. While the correlation in this group is still modest, it is worth noting as most most experimental economics studies are conducted in laboratories at elite universities where it is likely that participants have high IQs. Thus, this is likely to be the most relevant subgroup to compare with the prior literature.[17]

All told, 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 respondents in terms of IQ reduces the correlation to around 0.1. However, as there are 88 correlations in the table, it is perhaps not surprising that some turn out to be statistically significant.

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.

5 Relation with Measures of Risk Aversion

Thus far, our results suggests that WTA and WTP for lottery tickets are relatively stable, but independent. As these are values for lotteries, we examine how WTA and WTP relate

[16] Consistent with this, none of the coauthors of this study were able to get all six questions correctly on the first try.

[17] Of course, these participants are also likely to be young. Young people in our study exhibit moderate negative correlations. The young / high IQ group is too small to draw meaningful inferences about.

to a number of measures of risk aversion. The results are consistent with the *frame*—the way in which a question is asked—being particularly important in risk attitude elicitation. In particular, whether the framing can be said to be one of "buying" or "selling" seems to dictate the pattern of correlations. This section uses data from Study 2 only, as it contained a richer set of risk-attitude measures. We focus on correlations corrected for measurement error using ORIV.

The different measures of risk aversion fall into two broad categories: those where the MPL features a fixed lottery, and those where it features a variable lottery. The former category contains four measures that vary in terms of the randomization device in the question, and the domain—gains, losses, or both—over which risk is measured.[18]

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

Gain: A lottery where one payoff was a small gain, 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, 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: Here 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).

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

Table 5: ORIV Correlations between WTA/P and Risk Measures, Study 2

			Fixed Lottery					Variable Lottery	
		WTA	Urn	Gain	Mixed	Loss	WTP	FM	2L
	Urn	0.66***					−0.07		
		(.048)					(.067)		
	Gain	0.66***	0.65***				−0.04		
Fixed		(.056)	(.058)				(.071)		
Lottery	Mix	0.58***	0.51***	0.59***			−0.19***		
		(.058)	(.057)	(.058)			(.070)		
	Loss	0.27***	0.26***	0.39***	0.65***		−0.31***		
		(.076)	(.066)	(.070)	(.062)		(.068)		
	FM	0.03	0.05	0.09	−0.13*	−0.18**	0.45***		
Variable		(.070)	(.066)	(.069)	(.070)	(.075)	(.045)		
Lottery	2L	−0.12*	−0.17***	−0.13*	−0.21***	−0.15***	0.28***	0.40***	
		(.070)	(.063)	(.071)	(.073)	(.077)	(.060)	(.062)	
Qualitative		0.24***	0.18***	0.18**	0.17**	−0.05	0.15**	0.15**	0.13*
		(.063)	(.059)	(.077)	(.070)	(.089)	(.064)	(.063)	(.066)

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

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 are (re-)coded so that a higher value corresponds to less risk aversion.

The ORIV correlations between these measures, as well as between these measures and WTA and WTP, are displayed in Table 5. A first pattern of note is that many of these measures of risk aversion are only loosely related to each other. This is consistent with a large literature on risk attitude elicitation.[19]

Despite imperfect correlations, there are distinct groupings of higher and lower (or negative) correlations. Those in the fixed lottery group 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. The variable lottery measures are generally highly correlated with WTP, but not WTA. The only measure that appears to be (weakly) correlated across the board is the qualitative measure.

This grouping is intuitive if the fixed element of an MPL as acting as a reference point, as promoted by Sprenger (2015).[20] WTA is strongly related to measures of risk aversion with a fixed lottery, in which the agent is asked to "price" the lottery by comparing it with varying monetary amounts. It is as if the respondent were endowed with that lottery—indeed she can "keep" it quite easily—and thus, it is hardly surprising that her behavior here mimics her behavior when she is actually endowed with the lottery, as in the WTA question. By contrast, in questions with variable lotteries, the fixed option is an amount of money—that the respondent can keep by choosing left all the time—-and she is asked to choose a lottery

[19]For more recent results from the lab or the field, see Isaac and James (2000); Berg et al. (2005); Bruner (2009); Deck et al. (2010); Harbaugh et al. (2010); Deck et al. (2013); Nielsen et al. (2013); Loomes and Pogrebna (2014); Lönnqvist et al. (2015); Crosetto and Filippin (2016); He et al. (2016). See Weber and Johnson (2008) for a summary of the psychology literature on this topic. This is also naturally related to preference reversals (Grether and Plott, 1979).

[20]See also Bleichrodt et al. (2001), and the many references therein.

for which she is willing to renounce the money. Thus, following the same logic, this should be similar to WTP. While this perspective makes intuitive sense, this is the first direct evidence it is likely to hold. In particular, we are the first to show that MPLs where the respondent is "endowed" with a lottery through its appearance on the left-hand-side produce results consistent with those where the endowment is clearly stated in the question itself.

What is possibly less intuitive, and an empirical contribution, is that these two groups—the WTA and fixed lottery measures, and the WTP and variable lottery measures—are at best weakly related to each other.[21] Taken together, these results suggest that our data is structured by whether the question is framed in terms of buying or selling.

6 Prior Studies

The endowment effect has been extensively studied by both economists and psychologists. This wealth of former studies provides an opportunity to examine whether our results are consistent with prior data, or due to the use of representative samples. We show that both play a role. In existing studies the correlation between WTA and WTP for lotteries is around 0.15–0.2. This is greater than the correlation of ~ 0 observed in our data. However, it is consistent with the correlation we find for high-IQ people, who likely resemble the university students that populate existing studies.

Unfortunately, despite the abundance of prior studies, only six contain incentivized measures of both WTA and WTP for lotteries for the same people—that is, they use a within-subject design.[22] Of these six studies, we could only recover data from three: Plott and

[21] Camerer et al. (2017b) discusses the overall correlational structure of our data, and performs principal component analyses. Unsurprisingly, WTA and the fixed lottery measures load on one component, and WTP and the variable lottery measures load on another. Appendix C reports the relationship between WTA and WTP and other non-risk attitude measures in our data.

[22] 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). Tunçel and Hammitt (2014) conducted a similar search and found five studies that fit our criteria. The sixth we found was Fehr et al. (2015), a replication of Plott and Zeiler (2005), written after Tunçel and Hammitt (2014) was published.

Zeiler (2005), Isoni et al. (2011), and Fehr et al. (2015).[23] None of these published studies examined the correlation between WTA and WTP.

The three studies with available data are related. Plott and Zeiler (2005) claim the endowment effect is an artifact of the method used in most studies to elicit WTA and WTP—the BDM method of Becker et al. (1964). In particular, the paper notes that training respondents in how to use the BDM mechanism results in a substantially lower, and possibly zero, endowment effect across lab participants. However, the training for these final WTA and WTP elicitations was conducted by pricing lottery tickets through BDM. These training data show an endowment effect. This is consistent with the thesis of the paper, as these data are precisely the low-quality data that Plott and Zeiler argue creates spurious endowment effects. Isoni et al. (2011), in an extremely careful study, show that an endowment effect for lotteries exists even in data from participants who are well-trained with the BDM mechanism. However, they note several differences between their data and that of Plott and Zeiler (2005), likely stemming from the fact that the former is much higher quality. Fehr et al. (2015) also seek to replicate Plott and Zeiler (2005).

Each of the three papers contains measures of WTA and WTP for several lotteries. Two contain two distinct groups of participants who price different lotteries. All contain lotteries over gains. Two also contain lotteries that have both a gain and a loss.

We calculate correlations for each lottery over gains in each group in each study, and present the results in Table 6. We first examine these correlations for all participants. Next, following the robustness checks conducted with our own data, we exclude a participant in a particular lottery when that participant gave a dominated price for either WTA or WTP. This has two effects. First, it eliminates outlying answers. This reduces standard errors, and also tends to increase estimated correlations, due to the narrowing of possible values. Second, it reduces sample size, increasing standard errors. On net, standard errors increase somewhat, and, depending on the number of observations eliminated, correlations vary quite

[23] We contacted the authors of Harless (1989), Eisenberger and Weber (1995), and Borges and Knetsch (1998). All informed us that their data was no longer available.

Table 6: The correlation between WTA and WTP for lotteries over gains is limited in prior studies.

Study	Group (N)	Lottery	All Data	Excluding Dominated Choices	Excl. those w/ Dom. Choices
Plott and Zeiler (2005)	1 (38)	$0.3*1 \oplus 0.7*11$	0.26 (.16)	0.22 (.18)	0.28 (.19)
		$0.4*1 \oplus 0.6*6$	0.53*** (.14)	0.48*** (.15)	0.46** (.18)
		$0.7*1 \oplus 0.3*8$	0.21 (.16)	0.07 (.18)	0.13 (.20)
	2 (36)	$0.3*1 \oplus 0.7*8$	0.39*** (.16)	0.39*** (.16)	0.32 (.20)
		$0.6*1 \oplus 0.4*6$	0.20 (.17)	0.11 (.19)	−0.16 (.21)
		$0.7*1 \oplus 0.3*11$	0.55*** (.14)	0.16 (.19)	−0.09 (.21)
Isoni et al. (2011)	1 (100)	$0.3*1 \oplus 0.7*4$	0.01 (.10)	0.11 (.11)	0.08 (.14)
		$0.5*1.5 \oplus 0.5*3.5$	0.03 (.10)	0.17 (.12)	0.21 (.14)
		$0.6*1 \oplus 0.4*3$	0.20** (.10)	0.42*** (.10)	0.36*** (.13)
		$0.7*0.1 \oplus 0.3*0.8$	0.03 (.10)	−0.12 (.12)	−0.02 (.14)
		$0.7*1 \oplus 0.3*5$	0.10 (.10)	0.20 (.11)	0.29*** (.13)
Fehr et al. (2015)	1 (95)	$0.3*1 \oplus 0.7*8$	0.15 (.10)	0.33*** (.11)	0.15 (.15)
		$0.5*1 \oplus 0.5*1.5$	0.26** (.10)	0.19 (.15)	0.10 (.16)
		$0.6*1 \oplus 0.4*6$	0.20* (.10)	0.21** (.11)	−0.07 (.16)
		$0.7*1 \oplus 0.3*11$	0.11 (.10)	0.26** (.10)	0.31** (.15)
	2 (96)	$0.5*1 \oplus 0.5*1.5$	0.15 (.10)	0.49** (.18)	0.49** (.18)

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

a bit. A prime example of this is the second group in the study of Fehr et al. (2015). Only 27 out of the 96 participants gave non-dominated answers. When focusing on this subsample, the correlation between WTA and WTP grows substantially. The final set of correlations comes from eliminating all subjects that ever made a dominated choice.

The data from Isoni et al. (2011) and Fehr et al. (2015) paint different pictures. The former indicate little or no correlation between WTA and WTP, while the latter indicate a correlation between 0.15 and 0.25. The two different groups in Plott and Zeiler (2005) show a similar pattern of disagreement. Indeed, it was the small sample size, and contradictions in Plott and Zeiler (2005) (around the existence of the endowment effect, rather than the correlation between WTA and WTP) that lead Isoni et al. (2011) to pursue their study. It is unclear what is driving these differences. Taking an average—weighted by number of participants—of correlations across all studies indicates a correlation of 0.15–0.2 (depending on how dominated values are treated).

While this correlation is somewhat larger than we observe in our data, it is still quite a bit smaller than one might expect. Moreover, it is more or less in line with what we observe for those in the top 5% of IQ in our sample. Given that these studies were conducted at elite universities, this may be a consistent pattern across our data and prior data.

The finding here is less definitive than for our broader sample. However, it is worth noting that many of the lotteries—between 5 and 7 out of 15, depending on the treatment of dominated choices—show statistically insignificant coefficients. Thus, our results are not out of line with prior studies, and differences may be due to changes in subject populations or the structure of the lotteries being examined in these other studies. Moreover, to the extent that prior studies also show signs of a small or zero correlation, it shows that the difference in elicitation methodology—MPL versus BDM—is unlikely to play too much of a role.

Next, in Table 7, we present a similar analysis of lotteries that contain both gains and losses. While there are many fewer ob, and most are concentrated in the relatively small samples of Plott and Zeiler (2005), they do suggest a more robust pattern of correlation

Table 7: WTA and WTP for lotteries with gains and losses exhibit stronger correlations in prior studies.

Study	Group (N)	Lottery	All Data	Excluding Dominated Choices	Excl. those w/ Dom. Choices
Plott and Zeiler (2005)	1 (38)	$0.3 * -0.1 \oplus 0.7 * 0.8$	0.15*** (.17)	0.48*** (.16)	0.43** (.18)
		$0.5 * -3 \oplus 0.5 * 9$	0.11 (.17)	0.11 (.16)	−0.02 (.20)
	2 (36)	$0.5 * -3 \oplus 0.5 * 9$	0.61*** (.14)	0.61*** (.14)	0.37 (.20)
		$0.7 * -0.1 \oplus 0.3 * 0.8$	0.69*** (.13)	0.61*** (.14)	0.63*** (.17)
Fehr et al. (2015)	1 (95)	$0.5 * -3 \oplus 0.5 * 9$	0.35*** (.10)	0.36*** (.10)	0.48*** (.14)
		$0.7 * -0.1 \oplus 0.3 * 0.8$	0.22** (.10)	0.53*** (.11)	0.48*** (.14)

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

between WTA and WTP. We are hesitant to make too much of this, however, it is an intriguing difference that may be useful in discriminating between theoretical explanations for our findings.

7 Relation to Theory

Before discussing extant theories of the endowment effect, and how our findings apply to those models, it is worth reviewing what we have found. First, WTA and WTP appear to be uncorrelated, or possibly, slightly negatively correlated. This is true across a large number of subgroups of the data. Second, a substantial minority of respondents exhibit a negative endowment effect. 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 the buying or selling frame of a question is important. Fifth, and finally, our results are quite similar to existing studies that contain within-person measures of WTA and WTP.

None of these results imply that a given person's WTA and WTP should be independent across different goods or lotteries. That is, we expect that people who have a higher WTP for A than B would tend to maintain that ordering in WTA. For example, most people are likely to have a much higher WTA and WTP for a Ferrari than a used bicycle. Even for the same object, there are likely to be many cases where WTA and WTP would be highly correlated. This would tend to occur when some people have no value for an object, while others value it highly. For example, a ticket for a broadway show will have a much higher WTA and WTP for someone who lives in New York City than someone who lives in Timbuktu.

The theories we consider below are Prospect Theory, Stochastic Reference Dependence, and a simple theory of our own construction. As we describe below, the first two can only explain our findings with some difficulty. The third fits our data well, by design, and suggests questions for future research.

Throughout what follows, we assume that the lotteries under consideration have a high h and low l payoff, each equally likely. Each respondent i has a utility over gains given by a strictly increasing function $u_i(\cdot)$. We suppress subscripts, as these are not necessary for the analysis. The expected utility of a lottery is given by $EU = (u(l) + u(h))/2$.

7.1 Prospect Theory

Under Prospect Theory, selling a lottery for x is modeled as giving value $u(x) - \lambda EU$ where λ denotes loss aversion.[24] On the other hand, when buying a lottery, a monetary outlay is

[24] We do not consider probability weighting here. Adding this does not qualitatively change the discussion. Note that in our setting $\lambda < 2\frac{u(h)}{u(l)+u(h)}$, otherwise WTA will exceed the maximum payoff of the lottery, meaning the person prefers first order stochastically dominated options.

not *usually* encoded as a loss (Tversky and Kahneman, 1991, p. 1055), leading to:

$$\text{WTA}_{PT} = u^{-1}(\lambda EU)\,, \quad \text{and} \quad \text{WTP}_{PT} = u^{-1}(EU)\,.$$

Note that WTP_{PT} is simply the certainty equivalent of the lottery. Thus, it is decreasing in risk aversion. WTA_{PT} has a similar relationship with risk aversion, but is increasing in loss aversion. As both quantities depend on EU, WTP_{PT} and WTA_{PT} will be positively correlated if loss aversion λ is fixed, or if risk aversion and loss aversion are uncorrelated. In order to rationalize a small or zero correlation between WTP_{PT} and WTA_{PT}, a decrease in risk aversion—which leads to an increase in EU—must be offset by a decrease in loss aversion λ. Thus, a zero correlation between WTP_{PT} and WTA_{PT} implies a precise positive correlation between risk aversion and loss aversion. Unfortunately, preliminary evidence suggests that these aversions are negatively correlated, or uncorrelated.[25] This makes it difficult to reconcile our findings with the standard formulation of Prospect Theory.

However, as noted (but discarded) by the literature, it may be the case that respondents encode the payment for a lottery as a loss (Kahneman et al., 1990; Tversky and Kahneman, 1991). In this case, WTA_{PT} remains the same, but willingness to pay is now given by:

$$\text{WTP}'_{PT} = u^{-1}(EU/\lambda)\,.$$

As EU/λ and λEU always move in opposite directions, WTP'_{PT} and WTA_{PT} will be negatively correlated if the correlation between risk aversion and loss aversion is moderate or large, whatever the sign. Smaller correlations lead to zero or even small positive correlations between WTP'_{PT} and WTA_{PT}.

However, other scholars have discarded this formulation because it disagrees with other aggregate evidence (Kahneman et al., 1990; Tversky and Kahneman, 1991). Thus, we see one other way to rationalize our results, as well as others in the literature, using Prospect Theory.

[25] For example, Dean and Ortoleva (2015) report a negative correlation between the two variables.

In particular, the zero correlation may be the result of some people encoding spending money as a loss, while others do not. Thus, heterogeneity in how people encode the "buying" and "selling" frames, as gains and losses becomes important in rationalizing our findings, just as in Section 5. We return to this idea in Section 8.

7.2 Stochastic Reference Dependence

As we elicit WTP and WTA over lotteries, it is natural to consider our results in light of the theory of Stochastic Reference Dependence (Köszegi and Rabin, 2006). In this model, utility is composed of two parts: reference-free and reference-dependent. The reference-free part is computed using standard Expected Utility. The reference-dependent part is the difference between the utility obtained and the utility of the reference point, multiplied by loss aversion λ in cases of losses. As the reference point is stochastic, the reference-dependent part takes the expectation over all possible values of the reference point.

Under this model, keeping a lottery has value $EU + (0/2 + (u(h) - u(l))/2) + (\lambda(u(l) - u(h))/2 + 0/2)$. Selling the lottery $x \in (l, h)$ yields $u(x) + \frac{1}{2}(u(x) - u(l)) + \frac{1}{2}(\lambda(u(x) - u(h)))$. Setting them equal, we obtain

$$\text{WTA}_{SRD} = u^{-1}(EU) = \text{WTP}_{PT}.$$

As with WTP_{PT}, this is decreasing in risk aversion. The lack of dependence on λ has been observed more generally elsewhere (Sprenger, 2015).

WTP depends on the respondent's endowment e. The utility of not buying the ticket is $u(e)$. The utility for buying the lottery for $y \in (l, h)$ is $(u(e - y + h) + u(e - y + l))/2 + (u(e - y + h) - u(e))/2 + \lambda(u(e - y + l) - u(e))/2$. Assuming $e = h$, as in our experiment, and setting these expressions equal implies that WTP_{SRD} satisfies:

$$u(h)(3 + \lambda) = 2u(2h - \text{WTP}_{SRD}) + u(h + l - \text{WTP}_{SRD})(1 + \lambda).$$

WTP_{SRD} is decreasing in both loss aversion λ and risk aversion.[26]

Following similar logic to that developed above for Prospect Theory, WTP_{SRD} and WTA_{SRD} will be positively correlated if λ is fixed. However, as WTP_{SRD} is decreasing in λ, this implies that a negative correlation between risk and loss aversion is necessary to rationalize small correlations between WTP_{SRD} and WTA_{SRD}. As mentioned above, a negative relation is supported by preliminary empirical evidence. But note that for the correlation between WTA and WTP to be small, the relation and variance of the curvature of the utility and of λ must be exactly calibrated, not just for the sample as a whole, but for every subsample we examine.

7.3 Incomplete Preferences

Finally, we consider incomplete preferences (Masatlioglu and Ok, 2014). In this model, a respondent may not know exactly how much she values an object, only that it is, say, between \$5 and \$10. Coupled with a status quo bias—she only moves away from her current situation if she knows the alternative is better—this generates the endowment effect. In particular, if she sees the object being sold for \$5 or less she will buy it, and if someone offers her \$10 or more for it she will sell. As this model has little additional structure, it is trivially compatible with any correlation between WTA and WTP. The only feature of our data it is incompatible with are respondents exhibiting a negative endowment effect.

We propose a formalization of our results that is compatible with the model described above, but adds more structure. In particular, we posit that respondents are characterized by *two* risk aversion parameters: one related to buying, and one to selling. These parameters are independently, but not identically, distributed. Additionally, we assume that the distribution

[26] The latter is easy to see using linear utility, as suggested by Köszegi and Rabin (2006). To see this more generally, assume u is continuous. To show that WTP_{SRD} is decreasing in λ, note that if λ increases and WTP_{SRD} were to stay equal, the LHS increases by $u(h)/2$ and the RHS by $\frac{1}{2}u(h+l-\text{WTP}_{SRD})$. But note that $h > h+l-\text{WTP}_{SRD}$ as long as $h > \text{WTP}_{SRD} > l$, and thus $u(h) > u(h+l-\text{WTP}_{SRD})$; but then, if WTP_{SRD} did not change, the LHS would increase more than the RHS, violating the equality. Thus, WTP_{SRD} changes with λ; in particular, it must adjust to increase the RHS further—which is obtained by a smaller WTP_{SRD}. Thus WTP_{SRD} is decreasing in λ. The fact that WTP_{SRD} is decreasing in the curvature of u follows from standard arguments on risk aversion.

of risk preferences for buying first order stochastically dominates that for selling.[27] This formalization is conceptually related to the many models that suggest the presence of different utility functions used in different frames—for example, certainty vs. uncertainty, time vs. risk, etc.(Andreoni and Sprenger, 2010, 2012; Cerreia-Vioglio et al., 2015).

The advantage of this formalization is that it assumes the two independent values are generated by two independent parameters—and not by a highly specific relation between two underlying ones. It is the most direct rationalization of our data. It is compatible with an average endowment effect, but allows for a minority of respondents to have a negative endowment effect—rationalizing the choices of a sizable minority of participants (typically around 30%). By associating these two parameters with other choices under risk (as described in Section 5) it accounts for much of the heterogeneity in risk parameters observed in those choices, and the correlational structure between them.

A final advantage is that this formalization presents a descriptively accurate account of the psychology and neuroscience of decision making. Psychology has long noted domain differences in risk preferences across domains, and has argued that these preferences are distinct, and not generally related (see Weber and Johnson, 2008, for a summary). More specifically, different parts of the brain are activated when people estimate their WTP and WTA. (De Martino et al., 2009). These areas are distinct from those thought to be responsible for loss aversion.[28]

8 Discussion

The above suggests two plausible accounts compatible with our results and those in the literature. The first relies on Prospect Theory, but some people encode a loss when paying

[27]Formally, each person is characterized by two parameters: α_{buy} and α_{sell}, where $\alpha_{\text{buy}} \sim F_{\text{buy}}$ and $\alpha_{\text{sell}} \sim F_{\text{sell}}$, with $F_{\text{buy}} \perp\!\!\!\perp F_{\text{sell}}$. An endowment effect, on average, implies the means of the two distributions differ. Indeed we can assume that F_{buy} first order stochastically dominates F_{sell}: most respondents are more risk averse when they buy than when they sell.

[28]WTP is processed by the the medial orbitofrontal cortex—m-OFC, while WTA is processed by a more lateral portion of OFC—the l-OFC. Plassmann et al. (2007) also documents activity in the m-OFC related to the WTP for food by hungry subjects.

money for an object. The second suggests that people have distinct risk parameters governing buying and selling. These parameters are independently, but not identically, distributed.

Both accounts have drawbacks. The first suggests some individuals should be treated differently, but does not provide guidance on how to discern who should be treated one way or the other. Thus, it breaks the theoretically clean link between reference points and question structure. The latter theory is problematic as well. It does not link the endowment effect with loss aversion, which is one of the main benefits of both Prospect Theory and Stochastic Reference Dependence.

Another seeming drawback of the latter account—it requires two utility functions–sis illusory. The endowment effect cannot be accommodated by standard theory. Thus, to explain it with a utility-like theory, *any* theory will have, in effect, two utility functions. In Prospect Theory, or Stochastic Reference Dependence, these two functions are fused together at the reference point, resulting in loss aversion. In our account, these functions are not forced to meet.

It is worth noting that both of these accounts distinguish buying and selling from gains and losses. Intuitively, those who exhibit an endowment effect care a lot about getting the best deal when buying and selling, that is, they do not want to be a "sucker." Those who exhibit a negative endowment effect have a lower WTA and a higher WTP to "close the deal." This plausible intuition is coarse, and possibly inaccurate. However, it is difficult to tell given the current state of knowledge about buying and selling. This is unfortunate, as buying and selling are essential to economic activity.

References

Andreoni, James and Charles Sprenger, "Certain and Uncertain Utility: The Allais Paradox and Five Decision Theory Phenomena," 2010. University of California San Diego. *mimeo.*

_ **and** _ , "Risk Preferences Are Not Time Preferences," *American Economic Review*, 2012, *102* (7), 3357–76.

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.

Berg, J, J Dickhaut, and K 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.

Bleichrodt, Han, Jose Luis Pinto, and Peter P. Wakker, "Making Descriptive use of Prospect Theory to Improve the Prescriptive use of Expected Utility," *Management Science*, 2001, *47* (11), 1498–1514.

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.

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 Chapaman, Pietro Ortoleva, and Erik Snowberg**, "Time Stability of Econographics," 2017. Caltech, *mimeo.*

_ **, Mark Dean, Jonathan Chapaman, Pietro Ortoleva, and Erik Snowberg**, "Econographics: Factors of Economic Decision Making," 2017. Caltech, *mimeo.*

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

Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva, "Cautious Expected Utility and the Certainty Effect," *Econometrica*, 2015, *83* (2), 693–728.

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," 2015. 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.*

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques and Applications from the Caltech Cohort Study," 2015. California Institute of Technology, *mimeo.*

Grether, David M. and Charles R. Plott, "Economic Theory of Choice and the Preference Reversal Phenomenon," *The American Economic Review*, 1979, *69* (4), 623–638.

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 and the Role of Ambiguity: An Artefactual Field Experiment from China," 2016. ETH Zurich, *mimeo.*

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.

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.

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

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.

Masatlioglu, Yusufcan and Efe A. Ok, "A Canonical Choice Model with Initial Endowment," *Review of Economic Studies*, 2014, *81* (2), 851–883.

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.

Plassmann, Hilke, John O'Doherty, and Antonio Rangel, "Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions," *Journal of Neuroscience*, 2007, *27* (37), 9984–9988.

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.

Segal, Uzi, "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 1990, *58* (2), 349–377.

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.

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.

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^{**} (.035)	0.69^{***} (.020)	0.74^{***} (.021)
Study 1, Wave 2	-0.01 (.032)	0.02 (.031)	0.01 (.033)	0.00 (.043)	0.66^{***} (.024)	0.75^{***} (.023)
Study 2	-0.08^{**} (.035)	-0.04 (.036)	-0.07^{**} (.037)	-0.09^{**} (.044)	0.70^{***} (.031)	0.78^{***} (.028)

Notes: ***, **, * denote statistical significance at the 1%, 5%, and 10% level, with standard errors 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%	13%	60%	25%	12%	63%	

Figure A.1: Figure 1, unweighted.

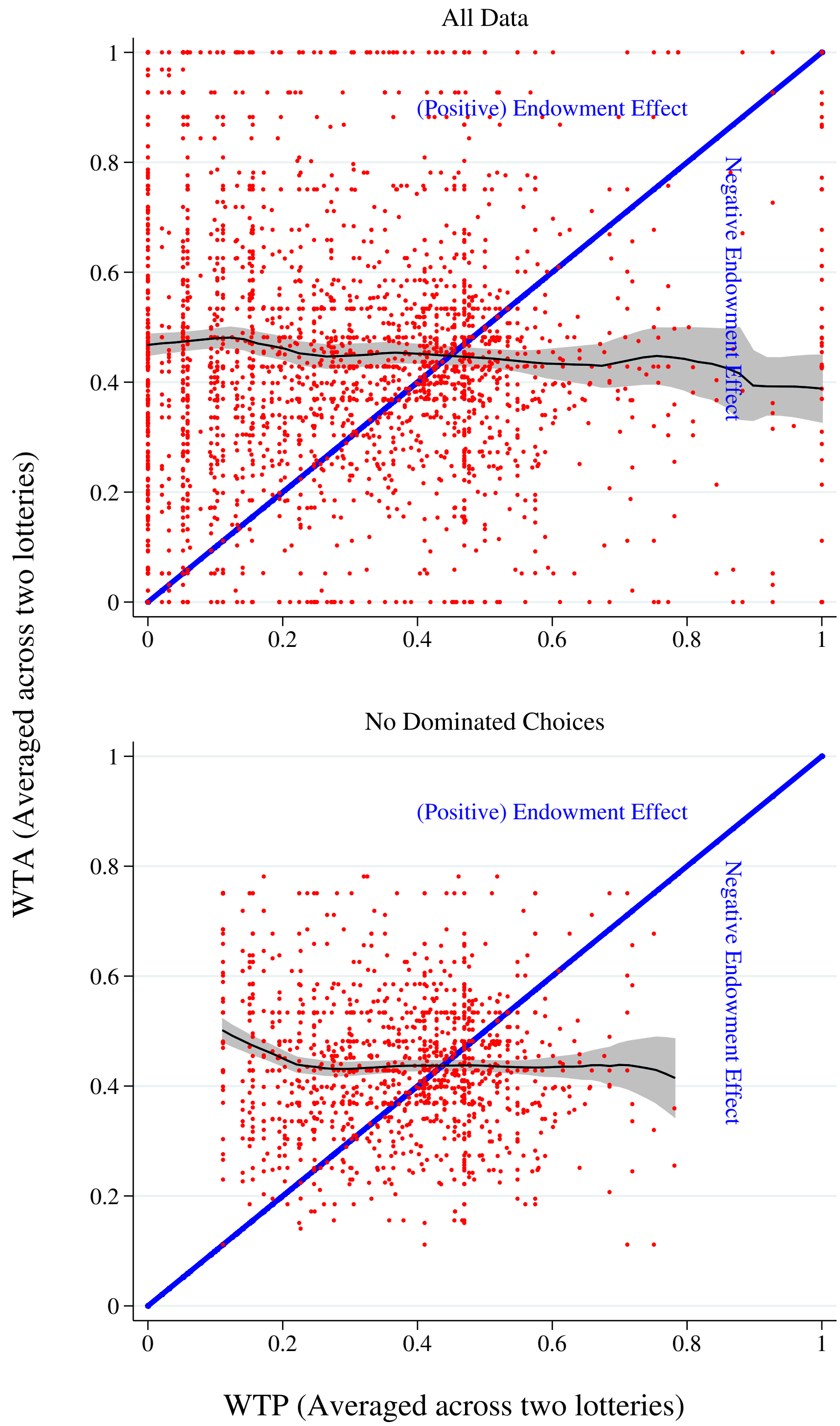


Table A.3: Table 4, unweighted.

	N	Lottery 1	Lottery 2	Average	ORIV
All	3,000	−0.06***	−0.05**	−0.06***	−0.09***
		(.021)	(.021)	(.022)	(.027)
No Dominated Choices	1,322	−0.10***	−0.06**	−0.12***	−0.16***
		(.032)	(.030)	(.032)	(.046)
Not Too Fast	2,700	−0.03	−0.03	−0.04*	−0.05*
		(.022)	(.022)	(.022)	(.029)
No Switches in Top or Bottom Two Rows	1,595	−0.08**	−0.08**	−0.08**	−0.12***
		(.031)	(.033)	(.032)	(.041)
No Switches in Top or Bottom Three Rows	1,028	−0.01	0.02	0.01	−0.01
		(.036)	(.036)	(.037)	(.056)
Order: WTA First	1,501	−0.07**	−0.04	−0.06**	−0.08**
		(.030)	(.030)	(.030)	(.038)
Order: WTP First	1,499	−0.06*	−0.06*	−0.07**	−0.09**
		(.030)	(.029)	(.031)	(.039)
Education: HS or Less	1,128	−0.11***	−0.06*	−0.11***	−0.15***
		(.033)	(.035)	(.034)	(.044)
Education: Some College	1,538	−0.03	−0.03	−0.03	−0.04
		(.031)	(.029)	(.031)	(.039)
Education: Advanced Degree	334	−0.00	−0.04	−0.02	−0.05
		(.068)	(.060)	(.065)	(.085)
Income: Above Median	1,520	−0.00	−0.03	−0.01	−0.02
		(.031)	(.030)	(.031)	(.040)
Income: Top Quartile	813	0.04	0.02	0.05	0.05
		(.043)	(.040)	(.042)	(.055)
Income: Top Decile	399	0.09	0.04	0.07	0.09
		(.062)	(.059)	(.063)	(.080)
CRT: Above Median	1,371	0.08**	0.06*	0.07**	0.09**
		(.033)	(.033)	(.034)	(.043)
CRT: Top Decile	288	0.10	0.11	0.10	0.11
		(.078)	(.075)	(.079)	(.097)
IQ: Above Median	1,694	−0.01	−0.02	−0.01	−0.02
		(.029)	(.027)	(.028)	(.036)
IQ: Top Decile	337	0.12*	0.10*	0.12**	0.15*
		(.061)	(.060)	(.061)	(.078)
IQ: Top 5%	146	0.25***	0.15*	0.21**	0.28**
		(.091)	(.088)	(.086)	(.12)
Age: Youngest Quartile	611	−0.07	−0.06	−0.08*	−0.11*
		(.048)	(.046)	(.049)	(.063)
Age: Second Youngest Quartile	798	−0.07	−0.05	−0.07	−0.09*
		(.041)	(.043)	(.042)	(.052)
Age: Second Oldest Quartile	783	−0.07	−0.01	−0.04	−0.06
		(.042)	(.042)	(.042)	(.053)
Age: Oldest Quartile	808	−0.05	−0.07*	−0.07*	−0.09*
		(.040)	(.039)	(.041)	(.053)

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

Figure A.2: Figure 2, unweighted

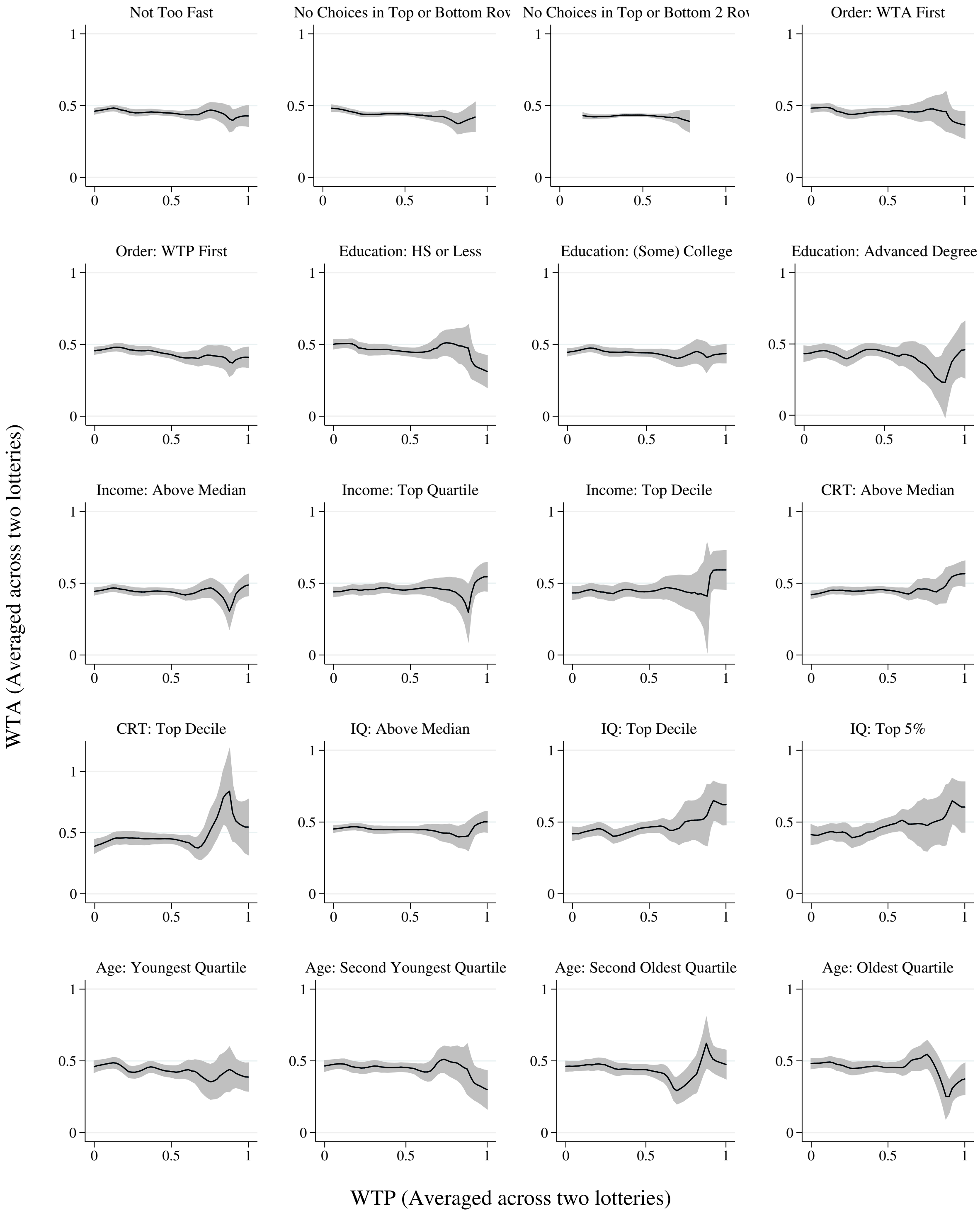


Table A.4: Table 5, unweighted.

			Fixed Lottery					Variable Lottery	
		WTA	Urn	Gain	Mixed	Loss	WTP	FM	2L
Fixed Lottery	Urn	−0.61***					0.05		
		(.042)					(.043)		
	Gain	−0.62***	0.68***				0.05		
		(.044)	(.037)				(.046)		
	Mix	−0.51***	0.47***	0.55***			0.23***		
		(.038)	(.036)	(.035)			(.042)		
	Loss	−0.30***	0.26***	0.36***	0.68***		0.34***		
		(.043)	(.042)	(.044)	(.035)		(.048)		
Variable Lottery	FM	0.02	0.03	0.05	−0.16***	−0.18***	−0.38***		
		(.047)	(.045)	(.047)	(.042)	(.043)	(.036)		
	2L	0.12**	−0.17***	−0.16***	−0.27***	−0.23***	−0.27***	0.42***	
		(.047)	(.044)	(.047)	(.042)	(.044)	(.041)	(.041)	
Qualitative		−0.16***	0.17***	0.18***	0.16***	0.01	−0.11***	0.08*	0.04
		(.045)	(.039)	(.046)	(.041)	(.050)	(.041)	(.042)	(.045)

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

B Results for each Study / Wave

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

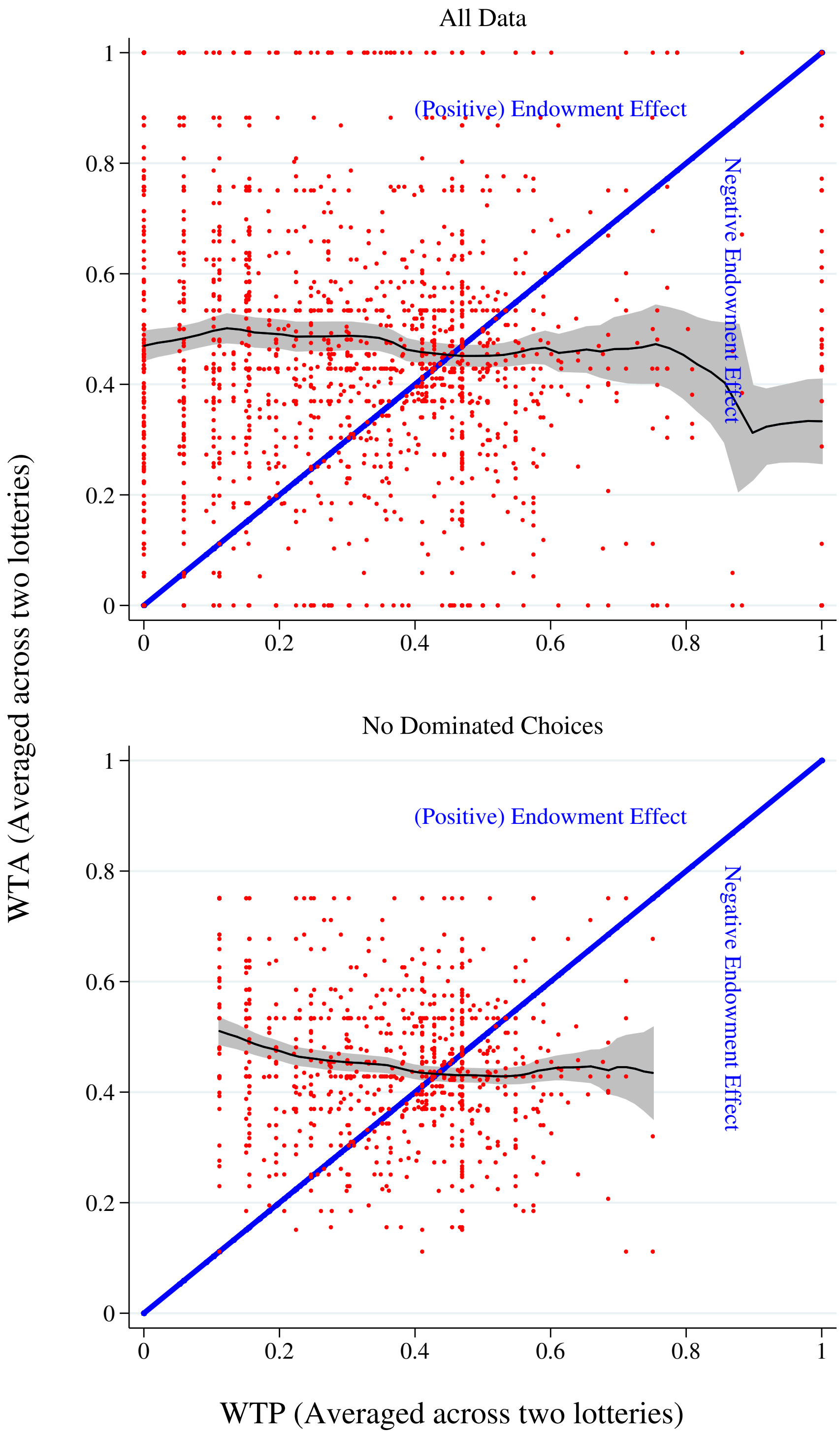


Table B.1: Table 4 for Study 1, Wave 1 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	2,000	−0.06*	−0.06*	−0.08**	−0.10**
		(.037)	(.037)	(.037)	(.047)
No Dominated Choices	923	−0.11**	−0.14***	−0.17***	−0.24***
		(.061)	(.051)	(.053)	(.077)
Not Too Fast	1,800	−0.03	−0.04	−0.05	−0.06
		(.040)	(.040)	(.039)	(.051)
No Switches in Top or Bottom Two Rows	923	−0.11**	−0.14***	−0.17***	−0.24***
		(.061)	(.051)	(.053)	(.077)
No Switches in Top or Bottom Three Rows	629	−0.06	0.00	−0.07	−0.14
		(.063)	(.070)	(.078)	(.12)
Order: WTA First	1,011	−0.06	−0.07	−0.08*	−0.10*
		(.046)	(.042)	(.044)	(.057)
Order: WTP First	989	−0.06	−0.06	−0.08	−0.10
		(.057)	(.059)	(.057)	(.073)
Education: HS or Less	765	−0.13**	−0.09	−0.13**	−0.18**
		(.058)	(.065)	(.059)	(.073)
Education: Some College	1,018	−0.01	−0.04	−0.03	−0.03
		(.050)	(.044)	(.049)	(.063)
Education: Advanced Degree	217	0.05	−0.02	0.00	0.02
		(.092)	(.076)	(.088)	(.12)
Income: Above Median	1,004	−0.04	−0.03	−0.04	−0.05
		(.046)	(.053)	(.046)	(.059)
Income: Top Quartile	540	−0.03	0.04	0.01	0.00
		(.070)	(.078)	(.069)	(.088)
Income: Top Decile	275	0.12	0.00	0.06	0.11
		(.087)	(.080)	(.087)	(.11)
CRT: Above Median	913	0.12**	0.04	0.08	0.12**
		(.053)	(.049)	(.052)	(.069)
CRT: Top Decile	189	0.16	0.21**	0.16	0.19
		(.099)	(.098)	(.10)	(.13)
IQ: Above Median	1,117	−0.01	−0.05	−0.04	−0.04
		(.049)	(.043)	(.047)	(.063)
IQ: Top Decile	232	0.10	0.09	0.09	0.12
		(.11)	(.096)	(.11)	(.15)
IQ: Top 5%	97	0.30***	0.16	0.20	0.34*
		(.11)	(.13)	(.12)	(.19)
Age: Youngest Quartile	374	−0.17*	−0.19**	−0.23**	−0.28**
		(.092)	(.086)	(.093)	(.12)
Age: Second Youngest Quartile	563	0.01	0.00	0.00	0.00
		(.071)	(.069)	(.065)	(.083)
Age: Second Oldest Quartile	582	−0.08	−0.07	−0.09	−0.12
		(.059)	(.052)	(.057)	(.073)
Age: Oldest Quartile	481	0.00	0.00	0.01	0.01
		(.062)	(.073)	(.059)	(.076)

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

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

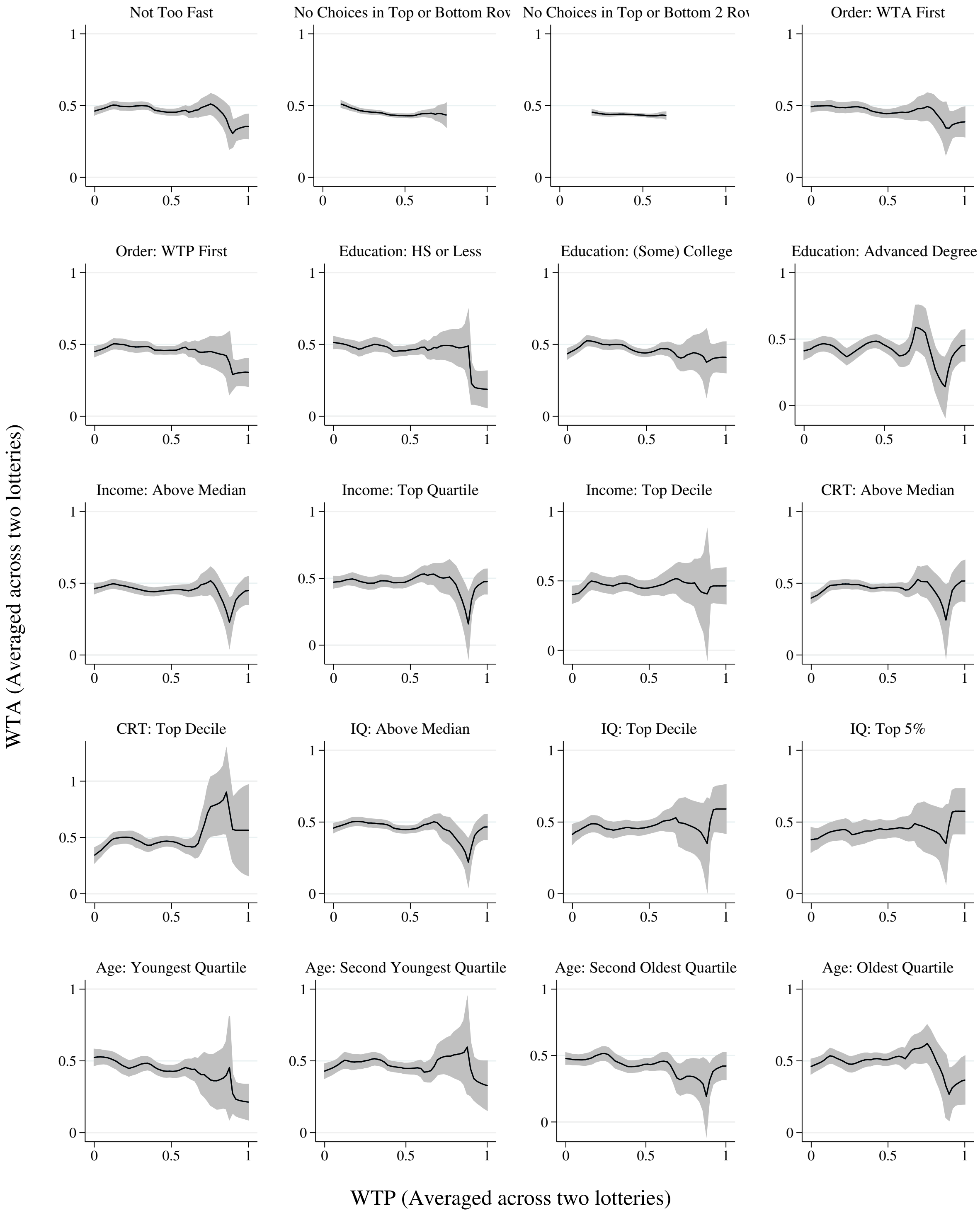


Figure B.3: Figure 1 for Study 1, Wave 2 only.

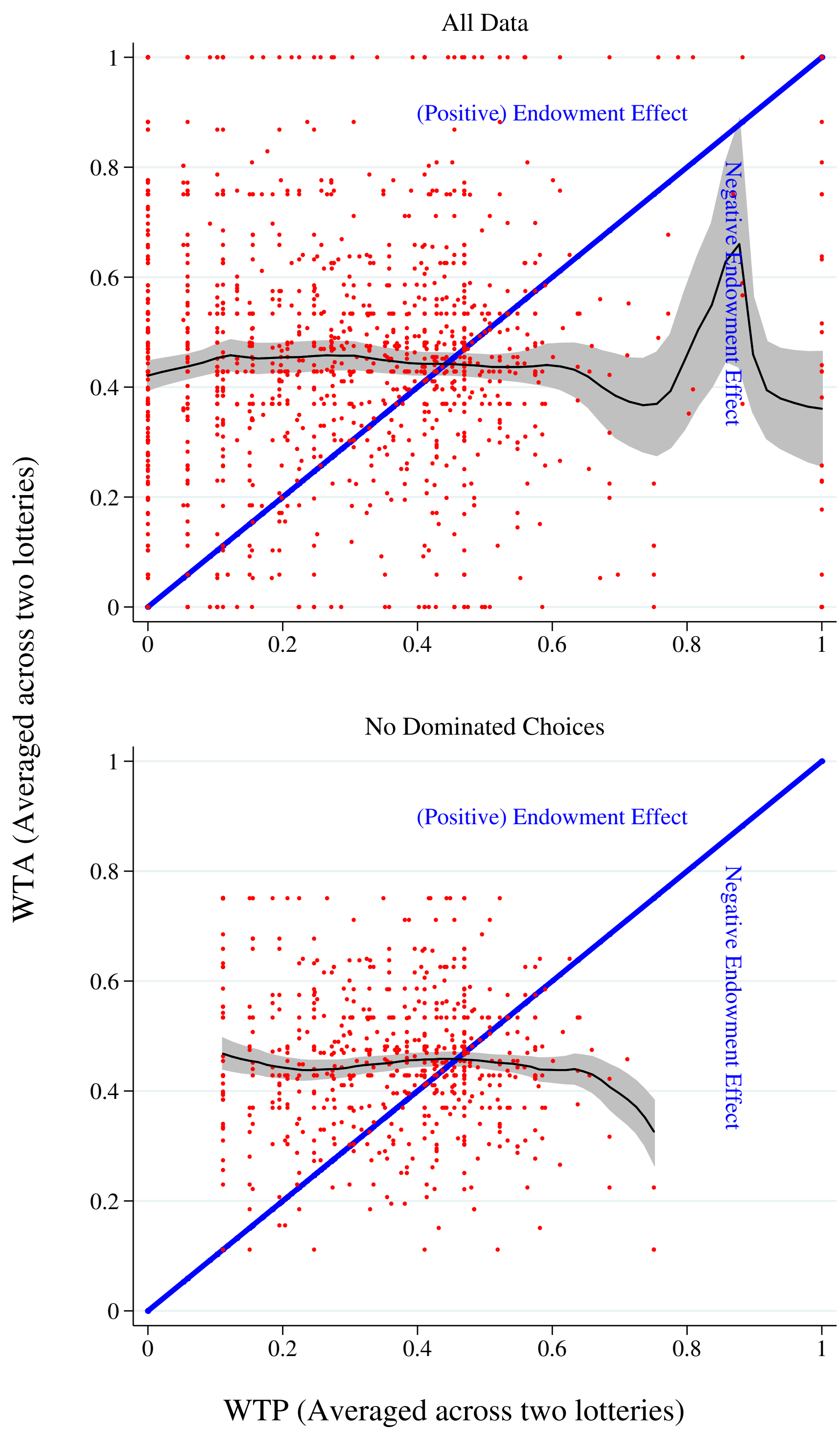


Table B.2: Table 4 for Study 1, Wave 2 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	1,466	−0.01	−0.02	−0.02	−0.02
		(.050)	(.049)	(.054)	(.069)
No Dominated Choices	718	0.05	−0.13**	−0.05	−0.02
		(.058)	(.051)	(.052)	(.083)
Not Too Fast	1,320	0.00	−0.02	−0.01	−0.01
		(.053)	(.052)	(.057)	(.075)
No Switches in Top or Bottom Two Rows	718	0.05	−0.13**	−0.05	−0.02
		(.058)	(.051)	(.052)	(.083)
No Switches in Top or Bottom Three Rows	484	0.10	−0.04	0.02	0.08
		(.083)	(.061)	(.070)	(.12)
Order: WTA First	754	0.00	0.07	0.04	0.04
		(.070)	(.070)	(.075)	(.097)
Order: WTP First	712	−0.03	−0.12**	−0.08	−0.08
		(.072)	(.063)	(.073)	(.095)
Education: HS or Less	525	−0.02	−0.06	−0.05	−0.05
		(.072)	(.076)	(.077)	(.095)
Education: Some College	775	0.00	−0.01	0.00	0.00
		(.077)	(.074)	(.085)	(.11)
Education: Advanced Degree	166	−0.06	0.08	0.04	0.02
		(.089)	(.089)	(.085)	(.11)
Income: Above Median	752	0.06	0.13**	0.11*	0.13
		(.063)	(.059)	(.063)	(.082)
Income: Top Quartile	389	0.05	0.12*	0.11	0.13
		(.087)	(.072)	(.077)	(.10)
Income: Top Decile	203	0.02	0.12	0.07	0.07
		(.13)	(.11)	(.11)	(.15)
CRT: Above Median	703	0.12**	0.12**	0.15***	0.20***
		(.056)	(.053)	(.054)	(.075)
CRT: Top Decile	190	0.18*	0.04	0.13	0.23
		(.11)	(.12)	(.11)	(.16)
IQ: Above Median	856	−0.01	0.02	0.01	0.01
		(.068)	(.065)	(.073)	(.098)
IQ: Top Decile	209	0.19**	0.29**	0.30***	0.37***
		(.096)	(.12)	(.10)	(.12)
IQ: Top 5%	96	0.30**	0.23**	0.29**	0.38**
		(.14)	(.11)	(.13)	(.17)
Age: Youngest Quartile	252	−0.05	−0.11	−0.09	−0.11
		(.12)	(.11)	(.13)	(.17)
Age: Second Youngest Quartile	390	0.04	0.06	0.07	0.09
		(.11)	(.11)	(.12)	(.15)
Age: Second Oldest Quartile	442	−0.03	−0.04	−0.04	−0.05
		(.083)	(.078)	(.084)	(.11)
Age: Oldest Quartile	382	−0.02	−0.01	0.00	−0.01
		(.073)	(.069)	(.072)	(.096)

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

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

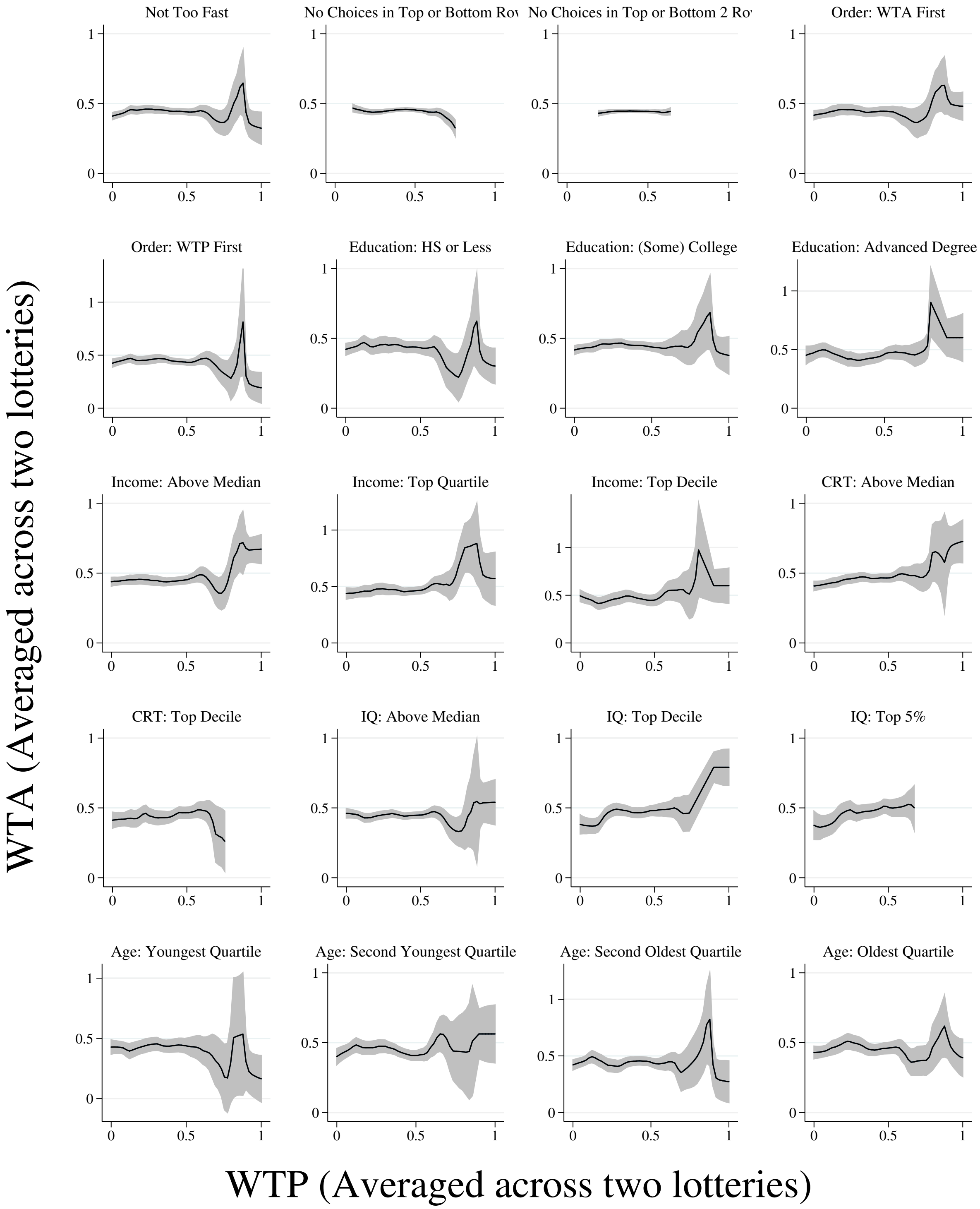


Figure B.5: Figure 1 for Study 2 only.

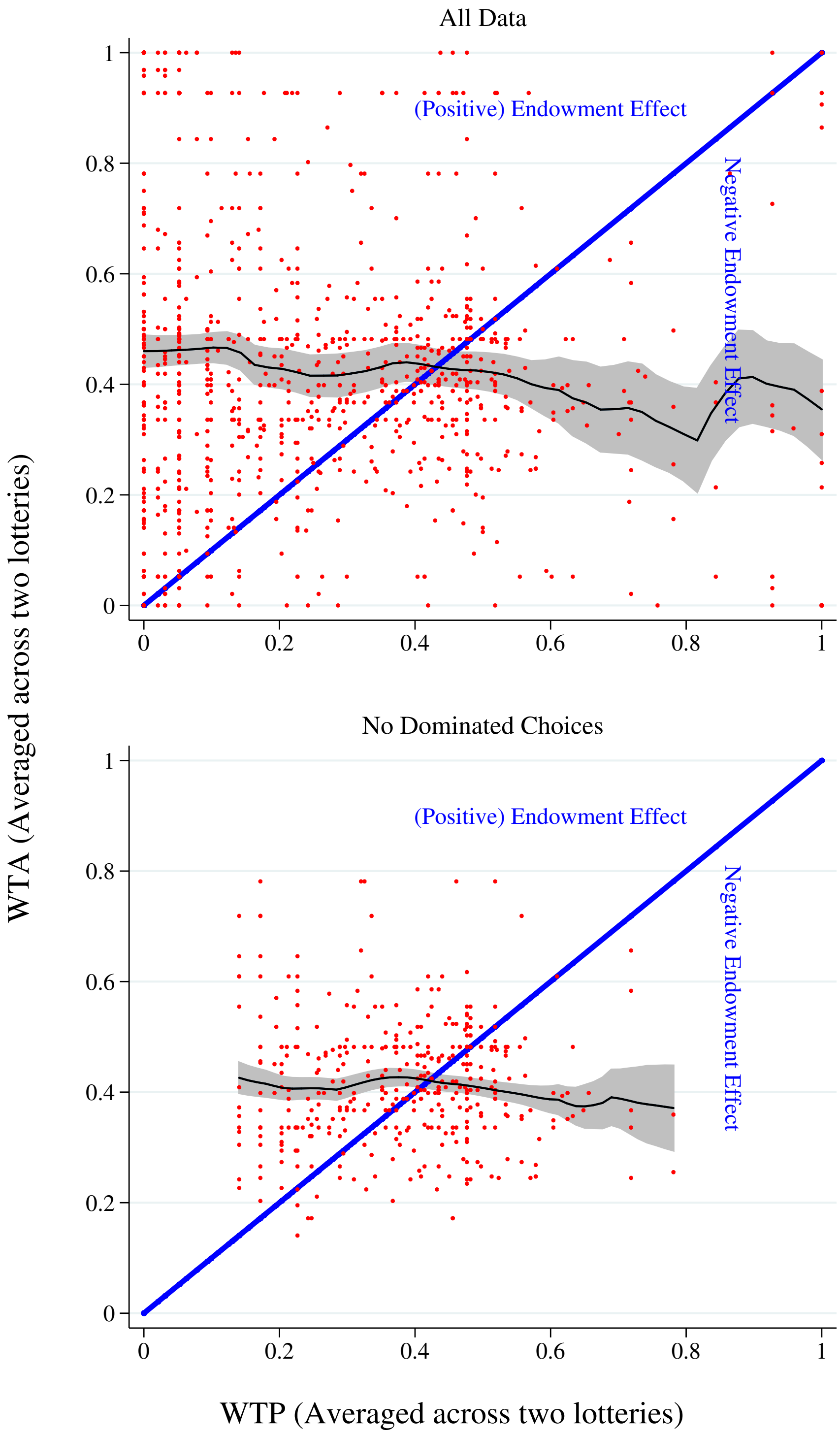
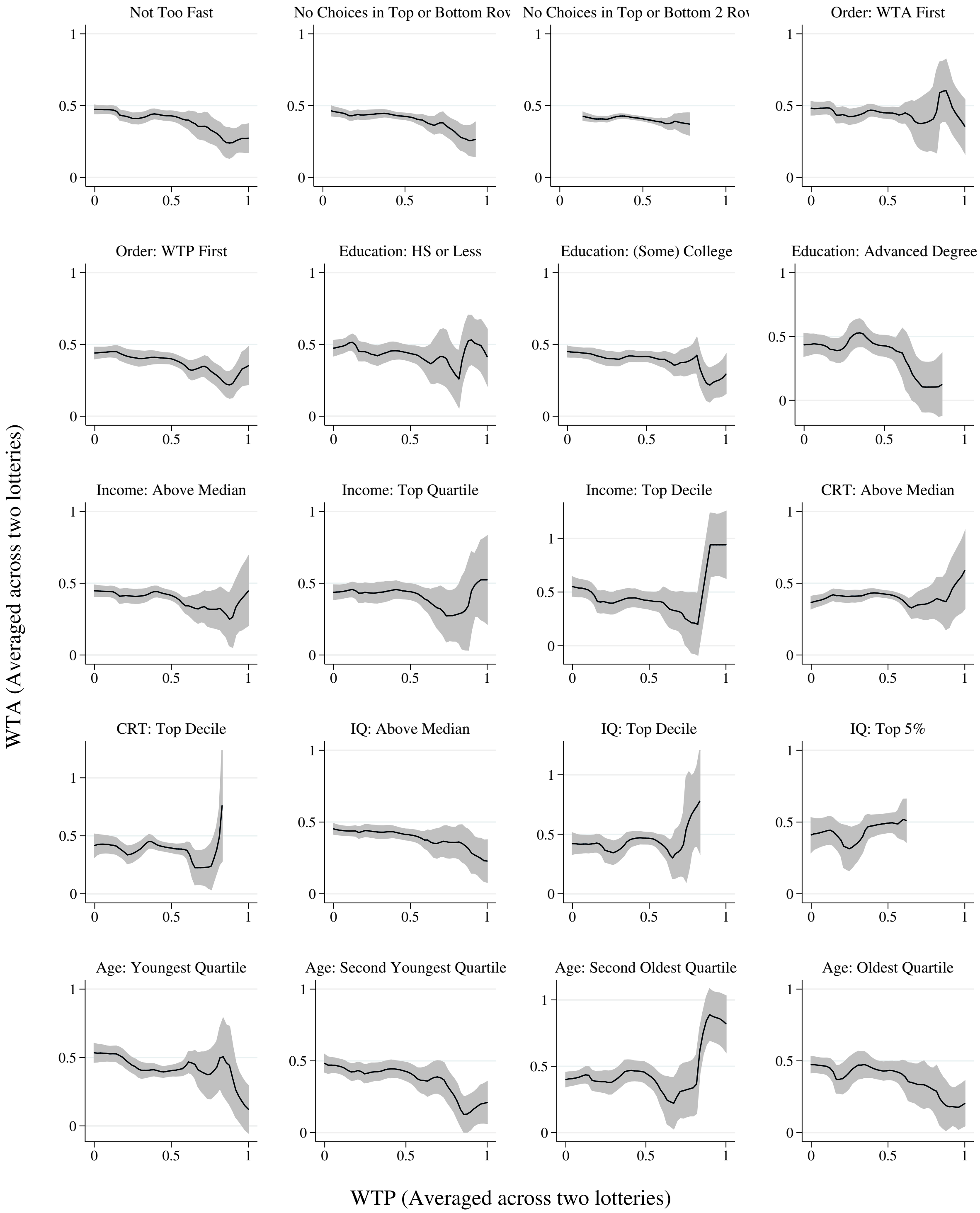


Table B.3: Table 4 for Study 2 only.

	N	Lottery 1	Lottery 2	Average	ORIV
All	1,000	−0.09*	−0.06	−0.09	−0.11
		(.051)	(.056)	(.058)	(.071)
No Dominated Choices	399	−0.03	−0.11	−0.07	−0.09
		(.10)	(.071)	(.075)	(.12)
Not Too Fast	900	−0.12***	−0.09*	−0.13***	−0.16***
		(.045)	(.049)	(.048)	(.060)
No Switches in Top or Bottom Two Rows	672	−0.10*	−0.13**	−0.13**	−0.15**
		(.058)	(.060)	(.057)	(.068)
No Switches in Top or Bottom Three Rows	399	−0.03	−0.11	−0.07	−0.09
		(.10)	(.071)	(.075)	(.12)
Order: WTA First	490	−0.06	−0.04	−0.06	−0.07
		(.085)	(.090)	(.093)	(.11)
Order: WTP First	510	−0.13**	−0.09	−0.13**	−0.18**
		(.055)	(.060)	(.060)	(.077)
Education: HS or Less	363	−0.08	−0.00	−0.06	−0.08
		(.087)	(.098)	(.10)	(.13)
Education: Some College	520	−0.09	−0.10*	−0.11*	−0.13*
		(.059)	(.061)	(.062)	(.074)
Education: Advanced Degree	117	−0.12	−0.18*	−0.15	−0.18
		(.120)	(.097)	(.11)	(.14)
Income: Above Median	516	−0.04	−0.11*	−0.07	−0.08
		(.062)	(.062)	(.064)	(.077)
Income: Top Quartile	273	0.02	−0.08	−0.01	−0.01
		(.083)	(.078)	(.082)	(.10)
Income: Top Decile	124	−0.08	−0.07	−0.08	−0.10
		(.12)	(.11)	(.12)	(.14)
CRT: Above Median	458	0.08	0.07	0.09	0.12
		(.069)	(.071)	(.072)	(.090)
CRT: Top Decile	99	−0.07	−0.05	−0.05	−0.06
		(.13)	(.14)	(.13)	(.16)
IQ: Above Median	577	−0.09	−0.09	−0.10*	−0.12*
		(.057)	(.059)	(.058)	(.067)
IQ: Top Decile	105	0.10	0.05	0.07	0.09
		(.12)	(.12)	(.12)	(.14)
IQ: Top 5%	49	0.26*	0.17	0.23	0.27*
		(.15)	(.14)	(.14)	(.16)
Age: Youngest Quartile	218	−0.27***	−0.25**	−0.30***	−0.40***
		(.083)	(.096)	(.087)	(.11)
Age: Second Youngest Quartile	262	−0.20***	−0.19**	−0.21**	−0.25**
		(.077)	(.088)	(.085)	(.10)
Age: Second Oldest Quartile	263	0.14	0.21*	0.17	0.19
		(.10)	(.11)	(.11)	(.13)
Age: Oldest Quartile	257	−0.11	−0.11	−0.13	−0.17
		(.087)	(.084)	(.089)	(.11)

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

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



C Relation to Other Variables, Study 2

Table C.1 extends this analysis to the other variables collected in Study 2. While none of the correlations are particular large in magnitude, some are significantly related with WTA and WTP. When one of these two has a statistically significant relationship with an attitude, the other tends not to, suggesting, once again, that these questions are measuring something, just not the same thing. Results on the relationship between all of the variables measured in Study 2 are discussed at length in Camerer et al. (2017b).

D MPL Ordering

Examining the design documents and screenshots found at http://people.hss.caltech.edu/ snowberg/wep.html shows that if a user always clicked, say, the third from 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, who are the most likely to follow such a pattern. However, here we conduct more extensive checks that control explicitly for the average position a respondent chooses on non-risk related MPLs.

In particular, as the questions about WTA and WTP are far apart in the survey, then 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 control for such a tendency, we need a measure of it. 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] To control for this tendency, we use 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 respondent clicked the right hand side the MPL (if individuals never switched they are assigned a value of the last row number plus one).

To control for a general tendency to select answers in higher or lower spots on an MPL, this information is entered in a number of 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

[1] Risk measures are correlated with WTA and WTP for lotteries—see Section 5. Thus, including the MPL positions on those measures would add error to the average position variables as some of that variable will represent real correlations between risk measures and WTA / WTP.

choices in various ways. The second column includes the first three principal components of the switching rows linearly. The first principal component is essentially an average, and the others contain more information about choices in these other MPLs. Together the first three principal components capturing 85% of the variation in MPL switching point location for these six choices. The third 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 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 correlation barely moves no matter how we try to control for the average position a respondent 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.1 shows the correlation between WTA and WTP as a function of the standard deviation of MPL switching points. To generate 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.1 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 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.

A final point is worth considering here. The prior studies, discussed in Section 6, produce results consistent with ours. All of those studies use BDM rather than MPL. Thus, it is unlikely that the particular design details are driving our results. To the extent they are, this would apply to a broader class of valuation elicitation methods.

Table C.1: ORIV correlations with other measures

	WTA	WTP
Common Ratio	-0.14^{**}	-0.17^{***}
	(.068)	(.060)
IQ	-0.05	0.07
	(.049)	(.054)
CRT	-0.11^{**}	0.06
	(.043)	(.053)
Patience	0.15^{**}	-0.05
	(.061)	(.063)
Strategic Sophistication	-0.13^{*}	0.01
	(.078)	(.084)
Reciprocity	-0.02	-0.04
	(.060)	(.064)
Punishment	0.01	0.03
	(.078)	(.069)
Altruism	-0.13^{**}	0.02
	(.061)	(.064)
Trust	-0.03	0.06
	(.054)	(.060)
Qualitative Trust	0.16^{**}	0.04
	(.063)	(.068)
Qualitative Trust	0.05	0.12^{**}
(binary Measure)	(.058)	(.055)
Qualitative Altruism	0.02	0.11^{**}
	(.056)	(.049)
Qualitative Reciprocity	-0.01	0.02
	(.061)	(.057)
Qualitative Punishment	-0.07	0.14^{**}
	(.069)	(.059)
Qualitative Patience	-0.01	-0.06
	(.061)	(.066)
Age	-0.03	-0.11^{**}
	(.049)	(.046)
Male	0.01	0.03
	(.050)	(.049)
Income	-0.03	-0.08
	(.052)	(.047)
Education	-0.10^{**}	0.04
	(.048)	(.053)

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

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

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

Figure D.1: Correlation as a function of standard deviation of switching points.