Reexamining Coherent Arbitrariness for the Evaluation of Common Goods and Simple Lotteries[§]

Drew Fudenberg[*], David K. Levine[**] and Zacharias Maniadis[***]

Abstract

We reexamined the effects of an anchoring manipulation on the valuation of common market goods that was introduced in Ariely, Lowenstein and Prelec (2003) in lab setting. We found much weaker anchoring effects. We performed the same manipulation on the evaluation of binary lotteries, and we found no anchoring effects at all. This suggests limits on the robustness of strong anchoring effects.

[§] We thank Drazen Prelec for many helpful conversations, and for sharing the experimental instructions and data from Ariely, Lowenstein and Prelec (2003). We are also grateful to Tore Ellingsen and Oscar Bergman for sharing their instructions and data with us. Zacharias Maniadis would like to thank Nikolaos Georgantzis, Fabio Tufano and Charlie Plott for useful comments. We are also grateful to Kathy Zeiler and for sharing the lottery data from Plott and Zeiler (2005). NSF grants 0646816 and SES-08-51315 provided financial support.

[*] Harvard University

[**] Washington University of St Louis

[***] Bocconi University, Department of Decision Sciences, and Dondena Center

1. Introduction

Decades of evidence from individual-choice experiments have documented that subjects do not always satisfy the assumptions of consumer theory and of expected utility theory. Some departures from the standard model can be accommodated while maintaining the assumption that people do have some sort of stable preferences. A more troubling issue is raised by experiments on phenomena such as anchoring effects, which suggest that many people do not make decisions using a fixed or slowly changing set of preferences. A defender of the standard models might argue that these arbitrary preferences do not apply to most everyday consumer decisions, and that an average person, with enough experience will have relatively fixed preferences for common market goods. A related question is whether decisions about uncertain prospects (lotteries) exhibit similar arbitrariness: do people have enough experience with choice over lotteries to have stable preferences over them?

Ariely, Lowenstein and Prelec (2003, henceforth referred to as ALP) used a simple anchoring manipulation to provide evidence that people do not have constant preferences for common consumer goods. They first generated a personal random number between 0 and 99 for each subject, asking them to turn the last two digits of their Social Security Number (SSN) into a dollar price. They then asked subjects whether they would buy each consumer good for this price. Subsequently, they elicited subjects' Willingness to Pay (WTP) for the goods using the incentive-compatible Becker-DeGroot-Marschak (1964) (BDM) mechanism. The basic evidence, in favor of the notion that preferences are "arbitrary", was that subjects' stated WTP for the goods was significantly correlated with their personal, randomly generated, number.

These results are troubling for economic theory. First of all, they are very strong. ALP write: "The effect is even more striking when examining the valuations by quintiles of the social security number distribution […] the valuations of the top quintile subjects were typically greater by a factor of *three*. For example, subjects with SSN in the top quintile were willing to pay $56 on average for a cordless computer keyboard, compared with $16 on average for subjects with bottom quintile numbers. Evidently, these subjects did not have, or were unable to retrieve personal values for ordinary products." Similarly, the average stated WTP for a bottle of average wine, for subjects whose SSN-based anchor price was in the highest quintile, was $27.9, more than three times as high as the analogous average WTP for subjects, whose SSN-based anchor price was in the lowest quintile. Also, ALP present evidence that these results seem to be robust across many dimensions: repetition, different levels of simplicity of the "hedonic experience" (consuming a good is one such experience), different subject pools, different levels of monetary incentives, and markets vs. individual choice. We believe that the ALP experiments are important and influential enough that they deserve to be replicated and tested for robustness.

We conducted three sets of experiments, in order to examine whether preferences for common market goods and for simple lotteries are "arbitrary", in the sense of ALP. Prior research, reviewed below, shows that both gambles and goods might be subject to anchoring effects. However, the relative magnitude of these effects has not been addressed. Are preferences for lotteries or common goods more arbitrary? It is difficult to predict *ex ante*: on the one hand, every lottery has intrinsic elements that can be used as a self-generated anchor for valuation, such as expected value. For example, in Johnson and Schkade (1989) about 40% of subjects use expected value as their valuation. On the other hand, lotteries are not sold in the market in the same way that common goods are. Hence, people may lack enough experience to form a constant valuation for them.

We repeat the part of ALP (2003) which deals with common market goods, but we find very weak anchoring effects with either WTP or WTA. The Pearson correlation coefficients between the anchor and stated valuation are generally much lower than in ALP (2003). We also used the anchoring manipulation of ALP for six simple lotteries, and also did not find anchoring effects on the valuation of these lotteries.

There are several possible reasons why our findings are different than ALP. Because ALP's experiment was conducted in a classroom, we had to deviate in several minor respects from the ALP environment. However, if the anchoring effects were as robust as evinced by the full set of experiments of ALP, these differences should not importantly change the results. We did observe that the goods which seem the most difficult to evaluate (because of unfamiliarity) appear to be relatively more prone to anchoring (but still correlations are not strong).

2. Literature Review

Several types of systematic phenomena, incompatible with constant preferences over final states, have been documented, although we should point out that most experiments in this literature use small gambles and multi-attribute goods, both of which make violations of stable preferences more likely.

If constant underlying preferences drive choices and inform valuation tasks, then, given a choice set, people should select the option for which they have the highest monetary valuation. However, a large literature on preference reversals for lotteries indicates that this might not be so. In particular, consider two binary bets, one of which offers a high probability of earning a small amount (while the other prize is zero), and the other bet offers a large amount with a low probability (and zero with the complementary probability). These have been called the "P-bet" and the "$-bet", respectively. In many experiments, subjects tend to choose the P-bet in binary choice tasks, but assign a higher value to the $-bet.[1]

One of the main explanations offered for the preference reversal phenomenon is called "response mode compatibility" (Slovic, Griffin and Tversky, 1991). According to this explanation, when stating valuation, subjects are more likely to anchor on the possible winning amount. Since adjustment from the anchor is insufficient, the stated valuation is close to the possible winning amount. On the other hand, when a choice mode is implemented, people anchor first on the probability of winning, and then insufficiently adjust for the possible winning amounts. Tversky and Thaler (1990, p. 209) write: "This phenomenon, or cluster of phenomena, challenges the traditional assumption

[1] The phenomenon was first discovered in the 60's and 70's by psychologists. Slovic and Lichtenstein (1983) provide a small survey, emphasizing the robustness of the phenomenon of preference reversals in these early experiments. Grether and Plott (1979) review the psychology literature and propose several possible confounds from the economist's point of view. When they correct for these confounds, the preference reversal phenomenon remains. Pommerehne, Schneider and Zweifel (1982) further increase the stakes and provide learning opportunities, but still observe sizable rates of reversals. Cox and Grether (1996) find that repetition in market environments tends to significantly reduce the rate of reversals.

that the decision maker has a fixed preference order that is captured accurately by any reliable elicitation procedure."

A similar preference reversal phenomenon has been documented in Certainty Equivalence vs. Probability Equivalence methods for eliciting indifference among options. In Hershey and Shoemaker (1985) subjects are first asked to state a sure amount, the certainty equivalent (CE), such that they are indifferent between receiving this amount, and receiving a binary gamble with a 50% chance of a fixed positive amount of money, and a 50% chance of zero. A week after this experiment, the same subject is asked to adjust the probability of winning the fixed amount in order to be indifferent between the resulting bet and an amount equal to the CE that they had stated in the first session of the experiment. The latter method is called probability equivalence (PE). Most subjects demand more than the 50% probability in the second stage. A similar bias, where subjects tend to demand an excessively good lottery, in order to be indifferent in stage two, was observed by carrying out the experiment in the opposite order (PE/CE).[2]

There is also a large class of experimental results, inconsistent with the assumption of constant preferences over final states, but largely consistent with the notion that people evaluate outcomes according to deviations from a reference point. With respect to standard consumer theory, several experiments have shown that the valuation for common goods might depend on whether it is measured as WTP or as WTA (see Kahneman, Knetsch and Thaler, 1991).

If the monetary valuation of goods and gambles is based on constant underlying preferences, they should be independent of irrelevant prior "anchoring" questions. Johnson and Schkade (1989) showed that asking subjects whether they would prefer a fixed amount (anchor), to a given gamble, affected the stated certainty equivalents, which were elicited subsequently. Chapman and Johnson (1999) had subjects generate a random anchor from their SSN, and found very strong anchoring effects on the stated WTA for selling a purely hypothetical lottery. Simonson and Drolet (2004) used the same SSN-based anchor as ALP, and found no anchoring effects on the stated WTA, and weak

[2] Johnson and Schkade (1989) increased the possible initial probability levels, and found that the second-stage bias in the CE/PE experiments depends negatively on the initial probability of winning the large amount.

anchoring effects on WTP, for four common consumer goods. Bergman et al. (2010) also used the design of ALP for six common goods, and found anchoring effects, but of smaller magnitude than in ALP.

3. Experiment 1: Common Market Goods

All of our experiments were paper-and-pencil as was the case in ALP. Our first (control) experiment reexamined the effects of random anchors on subjects' valuation for common market goods. The experiment took place in the California Social Science Laboratory (CASSEL) at UCLA, in August of 2009, and subjects were UCLA students. Subjects were recruited using the standard procedure of the lab, which is sending an announcement to a large email list. A total of 84 subjects participated,[3] and each of the three sessions had 26-28 subjects. The average cash payment per subject was 13.8 US dollars. Moreover, subjects received goods of total retail value equal to $298, which implies that the average value of earned goods per subject was about $3.5.[4]

Our subject pool consisted almost exclusively of undergraduate students, and the goods were chosen to be of interest for such subjects. The goods were an academic planner, a cordless keyboard, a financial calculator, a designer book, a pack of quality chocolates and a cordless mouse. The average retail price of the goods was $51.7. When the experiment started, subjects were asked to press the "Enter" key, and generate a random number between 0 and 99 (we used the program "Excel"). Subjects were then asked to copy this number six times in their answer sheet. After subjects had done so, we started reading the instructions loudly, so that everyone could hear.

For each good, subjects had to answer two questions. The first (anchor) question asked them whether they would be willing to sell the good, for an amount of money equal

[3] Although 84 subjects participated, five subjects were excluded from the data analysis since they could not qualify as observations. One subject failed to write a WTA for any of the goods, and another subject failed to write a WTA for one of the six goods. These subjects left the space allocated to the second question for each good (the WTA question), blank. One subject failed to write the same anchor number for the six goods. So, this subject formed six different anchoring questions, instead of a common one, for all the six goods. One subject wrote a random number equal to $100, which was not in the range of possible random numbers (0-99). Finally, one subject failed to answer any of the anchoring questions (the first question for each of the goods), leaving the space blank. So, our actual sample size consists of 79 subjects.

[4] Each of the participants who were not chosen received a participation fee of 13 dollars. Chosen persons also received this fee, in addition to being assigned one good.

to the personal random number that they had written down before the start of the experiment. In the second question, subjects were asked to state the minimum price for which they would be willing to sell the good. They were instructed to act as if the good was theirs to keep, since there was a positive probability that they would truly own it. We explained that three subjects would be randomly drawn. Each of the three chosen persons would be assigned one of the goods (a different good for each of three subjects drawn). For each of them, one of her answers (corresponding to her assigned good) would be carried out for real.[5]

If she were the first person randomly drawn, her answer to the first question would be consequential. So, if the subject stated that she wanted to sell the good for an amount equal to her random anchor, she did so. Otherwise, she kept the good, and she received it at the end of the experiment. If she were the person randomly drawn second or third, her answer to the second question (for their assigned good) would matter. We explained that in this case, a price would be randomly drawn from a matrix containing 100 values. The values of the matrix were drawn before the experiment, from a triangular distribution in [0,100], with a mode equal to zero. Subjects were shown the matrix on the screen. If the stated WTA was higher than the random price, the subject kept the good, and if not, she received a monetary amount equal to the random price, and she sold the good. We also explained, using short examples, why the BDM mechanism gives them an incentive to truthfully reveal their preferences.[6]

We explained that all uncertainty would be resolved using physical devices. The random price would be chosen, among the 100 possible prices, by throwing a 10-sided

[5] This process emulates the procedures of ALP (2003).

[6] We based this explanation on the instructions of Plott and Zeiler (2005). In particular, we said: "Please, leave your instructions aside for a while. With the second question, we want to know how much the opportunity to have each item is worth to you. This method simply gives you an incentive to state you true minimum you would be willing to accept for selling each item. Why is your best strategy to write the minimum you would be willing to accept? Because you will not receive the amount you ask for. Instead, if you receive anything, you will receive the random price. Let's try to understand this with an example. Imagine that I own the item. Say I believe that this item is worth 500$ for me. What happens if I falsely say that the minimum I would be willing to accept is more than 500? Assume I say 600$. If the random price is, say, 530, I do not sell the item. But, had I said 500, I could receive 530 for an item that I think is only worth 500. I lose out. What happens if I say less than 500? For example, I say only 400. If the random price is, say, 420, then I have to sell my item. I lose out, because I have to sell for 420 an item that I think is worth 500."

dice twice. The three goods would be determined by throwing a six-sided dice three times. The three chosen subjects would be chosen using a bingo cage with balls numbered according to subjects' participant numbers.

After we had finished explaining the instructions, we showed the six goods to subjects. We displayed each good in clear view, and a photograph of each good was also projected on a screen. After showing each good, we gave them time to complete both answers. After they finished, they were asked to copy their responses to a second answer sheet. After we picked up the second answer sheet, we resolved all uncertainty as explained, and concluded the experiment.

Given the results of ALP, our prior expectation was to find a positive relationship between the personal random number and the stated valuation. Therefore, in all our experiments, the statistical tests we shall use we will be one-tailed tests. For example, we shall reject the null hypothesis of a null Pearson correlation between the personal random number and the WTA - at a certain significance level - if we get a high enough positive value of the sample correlation.

Table 1 presents the main results. For each of the five quintiles of the distribution of the personal random number, we report the mean and the median WTA (the medians are in the parentheses). We also report the Pearson correlation coefficients, to understand whether there is a linear relationship between the WTA and the personal random number, and in order to allow comparisons with ALP. As seen in the table, the Pearson correlation coefficients between the random anchor and stated WTA were relatively small and only one of them was significant. We also report the Spearman correlation coefficient, which does not depend on parametric assumptions, and we find very similar values. The Pearson correlation coefficients in ALP's study range from 0.319 to 0.516, and they are all highly statistically significant. To get further evidence for the importance of anchoring, we test whether the distribution of stated WTA for subjects in the lowest quintile (0-19) of random numbers differs from the analogous distribution of WTA for subjects in the highest quintile (80-99). We believe that comparing behavior in the two extreme quintiles is the most favorable test for finding anchoring effects. Table 2 illustrates the results of the test. We use the nonparametric Wilcoxon rank-sum test, and for five of the six goods, we are unable to reject the null hypothesis of homogeneity at the

10% significance level. However, for the financial calculator, the hypothesis is rejected at the 2% significance level, which provides evidence that the distribution of WTA in the high quintile stochastically dominates the distribution in the low quintile. Hence, the Wilcoxon rank-sum test provides evidence in favor of anchoring effects for only one of the six goods.

A look at the data reveals more interesting features. The financial calculator and the designer book are relatively sophisticated goods, and were perhaps unfamiliar to our undergraduate student subjects. Hence, one would expect that these goods may have the most ambiguous value for our subjects. In fact, the correlation coefficients for these goods are higher.

Quintile (# Obs.)	**Academic Planner**	**Cordless Keyboard**	**Financial Calculator**	**Design Book**	**Milk Chocol.**	**Cordless Mouse**
0-19 (13)	7 (*5*)	38.85 (*35*)	10.23 (*10*)	11.30 (*10*)	5.62 (*4*)	22.07 (*25*)
20-39 (16)	12.44 (*10*)	50.19(*47.5*)	20.37 (*20*)	14.75 (*15*)	17.88 (*7*)	42.88 (*23*)
40-59 (16)	11.50 (*10*)	51.94 (*50*)	20.94 (*15.5*)	18.12(*17.5*)	8.47(*5.25*)	24.56 (*20*)
60-79 (20)	10.05 (*8*)	38 (*30*)	18.78 (*11*)	16.5 (*13.5*)	7.47 (*5*)	19.95 (*15*)
80-99 (24)	7.64 (*7.5*)	47.28 (*30*)	31.92 (*23.5*)	15.57 (*15*)	7.36 (*6.5*)	26.85 (*20*)
All data	9.90 (*9*)	45.07 (*40*)	20.46 (*15*)	15.46 (*15*)	9.46(*6*)	27.10 (*20*)
Pearson	-0.047	-0.022	0.207	0.134	-0.096	-0.108
p-value	>0.5	>0.5	0.034	0.12	>0.5	>0.5
Spearman	-0.0458	-0.0590	0.2075	0.1261	0.0704	-0.1334
p-value	>0.5	>0.5	0.033	0.134	0.266	>0.5

Table 1. Experiment 1, average and median WTA per quintile of anchor's distribution, and correlations.

Item	Academic Planner	Cordless Keyboard	Financial Calculator	Design Book	Milk Chocolates	Cordless Mouse
z	-0.587	-0.685	-3.008	-1.424	-1.342	-0.171
p-value	0.557	0.493	0.003	0.154	0.180	0.864

Table 2. Experiment 1, Wilcoxon test of homogeneity of the distributions of WTA for subjects in the first and fifth quintile of the range of random numbers.

4. Experiment 2: Lotteries

In our second experiment, six lotteries, with a range of expected values relatively similar to the range of the market prices of the six goods, were used. The experiment also took place in the CASSEL lab at UCLA, during November of 2008 and August of 2009. A total of 110 subjects participated,[7] in four sessions of 26-28 subjects each. Subjects were UCLA students, and the great majority was undergraduate students. The average total earnings per subject were equal to $21.7.[8]

The six gambles are shown in Table 2. The procedures and the instructions were similar to Experiment 1, with lotteries instead of goods.[9] In each experimental session,

[7] We exclude from the data a person who did not state a WTA for one of the lotteries, leaving the space allocated for the WTA question blank. We also exclude a participant who stated a WTA equal to $300 for the first lottery, and a WTA $400 for the sixth lottery. Since the maximum prize of any lottery is $100, we felt that it is impossible that someone could actually have such a WTA for the lotteries. Therefore, our final sample size is 108. It is worth emphasizing that we have not excluded a subject whose WTA for the chocolates and the cordless mouse was $100 and $300, respectively, because it is not a priori impossible for someone to have such WTA for these common goods.

[8] All subjects received a participation fee of $16. The three subjects who were randomly chosen to receive a lottery could earn money on top of that.

[9] The only difference was that in this experiment we further explained to subjects how they should answer the questions. In particular, after presenting the two questions, we drew attention away from the written instructions, and told subjects: "Please go back to the last two pages of your instructions but please do not write anything yet. Please have a look at the lotteries and the questions you are asked to answer. The first question for each lottery is easy: do you prefer to have the opportunity to play this lottery or the specified amount of money? An example will illustrate what you should answer to the second question. Imagine that I own the right to play a particular lottery. Say this sample lottery ticket. How do I know what amount is the minimum I 'd be willing to accept for the lottery? First of all, notice that, if I keep it, this lottery ticket gives me the opportunity to earn some money with a certain probability. I am not willing to give it up without getting some money. The question is, what is the minimum I would accept? Start with 100$. Would I be willing to give up my lottery in exchange for 100? If so, decrease the amount to 95. If I am willing to accept 95 in exchange for the lottery, then decrease further. I keep decreasing until I come to the amount that makes me indifferent between keeping the lottery and getting the money. Your objective is to

three subjects would be drawn at random, and each would get to own one lottery for real, and would either keep it or sell it, depending on the answers given to the experimental questions. The lottery was carried out for real, using a dice. Subjects, who chose to keep them, by indicating a WTA higher than the random price, would receive the prize drawn from the lottery. Otherwise, they simply received the random price from selling the lottery. All of this was made common knowledge to subjects.

In this experiment, we did not observe significant anchoring effects. As Table 3 shows, the Pearson and Spearman correlation coefficients between the personal random anchor and the stated WTA were not statistically significant for any of the lotteries. The results of the Wilcoxon rank-sum test for this experiment are illustrated in Table 4. For all six lotteries, this test fails to reject the hypothesis of homogeneity of the distributions of stated WTA, for subjects in the two extreme quintiles of the distribution of personal random numbers.

The lack of anchoring in our experiment might be due to the simplicity of the lotteries we considered. Chapman and Johnson (1999) used a SSN-based anchor, but for the complicated binary lottery (287$, 17%; 18.5$, 83%), and found a much higher correlation coefficient (r=0.45).

Another factor that may be relevant is that lotteries like Lottery 3, with a very small probability of winning a very large amount, may be difficult to evaluate. For example, in Grether and Plott (1979), the lottery couple with the highest incidence of preference reversals was the couple L1=(4$, 32/36; -0.5$, 4/36) L2=(40$, 4/36; -1$, 32/36). L2 had the largest possible prize among all lotteries considered in their experiments. In the same vein, Butler and Loomes and (2007) argue that subjects' behavior in preference reversal experiments can be understood if one assumes that they have imprecise preferences. Consider binary lotteries, for example. Fixing a specific binary lottery, Butler and Loomes argue that people have an "imprecise preference set" of binary lotteries, which corresponds to the fixed lottery. For any element of this set, they are unsure whether they prefer it to the fixed lottery. Using the previous arguments by

think like this for all six lotteries. You should think carefully before answering these questions because these questions are not purely hypothetical. There is a chance you will actually own each lottery, and each of your decisions may be carried out for real money."
This part was also based on the instructions by Plott and Zeiler (2005).

McCrimmon and Smith (1986), they claim that the more dissimilar from certainty a bet is, the wider should the "set of imprecise preferences" be. Consequently, the lack of strong anchoring for Lotteries 3-6 might be due the fact that the range of possible prizes was narrow, and the probability of winning the large amount relatively high.

Quintile (# Obs.)	**0.5, \$100; 0.5, \$0**	**0.25, \$100; 0.75, \$0**	**0.1, \$100; 0.9, \$0**	**0.75, \$60; 0.25, \$20**	**0.5, \$60; 0.5, \$20**	**0.9, \$60; 0.1, \$20**
0-19 (18)	42.11 (*40*)	24.67 (*20*)	14.67 (*10*)	41.11 (*40*)	34.39 (*35*)	49.39 (*52*)
20-39 (13)	49.69 (*50*)	26.62 (*25*)	15.77 (*11*)	44.69 (*45*)	37 (*40*)	48.69 (*50*)
40-59 (22)	51.54 (*50*)	31.82 (*32.5*)	23.55 (*20*)	46.05 (*45*)	37.86 (*40*)	55.32 (*55*)
60-79 (21)	41.67 (*50*)	22.19 (*25*)	10.23 (*10*)	39.67 (*45*)	34.67 (*40*)	49.52 (*55*)
80-99 (34)	50.02 (*50*)	31.45 (*27*)	21.15(*13.5*)	43.15(*42.5*)	34.18 (*30*)	51.38(*52.5*)
All data	47.35 (*50*)	28.01 (*25*)	17.78 (*10*)	42.90 (*45*)	35.40(*39.5*)	51.17 (*55*)
Pearson	0.057	0.082	0.071	-0.010	-0.076	0.032
p-value	0.278	0.199	0.23	>0.5	>0.5	0.37
Spearm.	0.119	0.090	0.034	-0.026	-0.104	0.025
p-value	0.110	0.176	0.362	>0.5	>0.5	0.399

Table 3. Experiment 2, average and median WTA per quintile of anchor's distribution, and correlations.

Lottery	**0.5, \$100; 0.5, \$0**	**0.25, \$100; 0.75, \$0**	**0.1, \$100; 0.9, \$0**	**0.75, \$60; 0.25, \$20**	**0.5, \$60; 0.5, \$20**	**0.9, \$60; 0.1, \$20**
z	-1.577	-1.312	-1.125	-0.371	0.304	-0.485
p-value	0.115	0.189	0.261	0.711	0.761	0.628

Table 4. Experiment 2, Wilcoxon test of homogeneity of the distributions of WTA for subjects in the first and fifth quintile of the range of random numbers.

5. Experiment 3: Willingness to Pay

In order to examine whether the differences of our results from ALP's were driven by the fact that we elicited WTA, whereas ALP had elicited WTP, we performed an additional treatment. This treatment also took place at the CASSEL lab at UCLA, in April of 2010, and the participants were UCLA students. A total of 79 subjects participated, in three sessions of 26 or 27 students.[10] The average payment per subject was equal to $21.8.[11]

The sessions were almost identical as those of Experiment 1. The main difference was that subjects were asked to state the most they would be willing to pay for the goods.[12] Hence, the subjects that would be chosen at random would not own the good, but they would have the chance to buy it. To conform to the standards of the CASSEL lab, a large participation fee ($93) was offered to the three chosen subjects, to make sure they can buy each good, and in order to prevent subjects from losing money in the experiment.

The results are illustrated in Table 3. The data show non-significant anchoring effects, and the both the Pearson and the Spearman correlations are very close to zero. The results of the Wilcoxon rank-sum test,[13] shown in Table 6, also confirm the lack of anchoring effects. The results also indicate the presence of a possible WTP-WTA gap, but this seems to be a relatively small one, in the order of 20%-30%.

6. General Discussion

There were several differences in our experimental design relative to ALP (2003), who elicited subjects' WTP for six common goods: Belgian chocolates, a wireless keyboard, a design book, a bottle of rare wine, a bottle of cheap wine, and a cordless trackball. They performed the experiment in a class with 55 MBA students, and decisions were consequential for only six subjects, drawn at random, one for each good. We deviated in several aspects from ALP: we conducted the experiments in a lab environment; we changed some of the goods, since they would be given to undergraduate

[10] One subject failed to answer any of the six anchoring questions and was excluded for the analysis. Hence, our final data set consists of 78 subjects.
[11] The subjects who were not chosen received a participation fee of $13.
[12] Four of the six goods were the same as in Experiment 1, while two goods were not the same, but very similar. A Class Organizer was used instead of an Academic Planner, and a very similar Cordless Mouse was used instead of the one used in Experiment 1.
[13] As before, this test is for equality of the distributions of stated WTP, between the groups of subjects in the first and fifth quintiles of the random number's distribution.

students,[14] and we used part of the instructions of Plott and Zeiler (2005), in order to explain the BDM mechanism, and help subjects understand the concept of WTA for lotteries. The reason for this last change in procedure is that the subjects in ALP were taught about the BDM mechanism and WTA as part of a classroom lecture, which makes the exact form of their instructions difficult to replicate. Moreover, in two recent papers, Plott and Zeiler (2005, 2007) showed that the complicated nature of the BDM mechanism makes it difficult to understand. They also showed that lengthy instructions and a long practice process were instrumental for ameliorating this problem. The use of these explanations might have played a role in the difference of our results. Subjects' misconceptions might inflate anchoring effects, if subjects are inclined to search cues from the environment in order to respond to complicated problems: the anchor might play the role of such a cue.

Quintile (# Obs.)	**Class Organiz.**	**Cordless Keyboard**	**Milk Chocol.**	**Cordless Mouse**	**Design Book**	**Financial Calculator**
0-19 (20)	3.25 (*1*)	30.7 (*21*)	4.82 (*4*)	16.75 (*14.5*)	7.7 (*4.5*)	15.05 (*11.5*)
20-39 (11)	2.95 (*2*)	33.45 (*30*)	4.77 (*4*)	15.63 (*10*)	5.55 (*4*)	12.09 (*10*)
40-59 (12)	4.00 (*2*)	32.33 (*35*)	4.83 (*5*)	17.58 (*15*)	8.91 (*5*)	7.66 (*6*)
60-79 (19)	4.37 (*3*)	33.37 (*25*)	5 (*5*)	16.95 (*20*)	8.32 (*4*)	13.74 (*10*)
80-99 (16)	4.81 (*1*)	28.31 (*25.5*)	4.19 (*3*)	12.87 (*10*)	6.25 (*3*)	9.81 (*6*)
All data	3.92 (*1.5*)	31.5 (*25*)	4.73 (*4*)	15.97 (*15*)	7.43 (*4.5*)	12.1 (*8.5*)
Pearson	0.096	-0.003	-0.047	-0.039	-0.013	-0.08
p-value	0.2	>0.5	>0.5	>0.5	>0.5	>0.5
Spearman	0.045	0.009	-0.007	-0.029	-0.013	-0.074
p-value	0.349	0.47	>0.5	>0.5	>0.5	>0.5

Table 5. Experiment 3, average and median WTP per quintile of anchor's distribution, and correlations.

[14] Wine could not be offered to students less than 21 years old, who are the majority of UCLA undergraduate students. Cordless trackballs were largely unknown to undergraduate students at the time of the experiment.

Item	Class Organizer	Cordless Keyboard	Milk Chocolates	Cordless Mouse	Design Book	Financial Calculator
z	0.148	-0.191	0.562	0.959	0.500	1.024
p-value	0.883	0.848	0.574	0.338	0.617	0.306

Table 6. Experiment 3, Wilcoxon test of homogeneity of the distributions of WTP for subjects in the first and fifth quintile of the range of random numbers.

We also preferred not to use the SSN-based anchor, since during a pilot some subjects showed privacy concerns. This might be another possible reason for the difference in our results. There is evidence that subjects might not view such numbers as truly arbitrary. For example, one-third of participants in one of the experiments of Chapman and Johnson (1999) stated that they though the SSN-based anchor number was informative. In another experiment, 57% said that they though that the experimenter wanted the number to influence their judgment. Perhaps a SSN-based anchor is not seen as random, in the sense that the value of the SSN has been determined long ago.

Our results do not confirm the very strong anchoring effects found in ALP. They are more in agreement with the results of Simonson and Drolet (2004) and of Tufano (2009), who also failed to confirm the robustness of the ALP (2003) experiments. He used the anchoring manipulation to increase the variance in subjects' WTA for a bad-tasting liquid, but the manipulation had no effect. Since this liquid offers a simple (negative) hedonic experience, like the "annoying sounds" of ALP, these results suggest limits on the robustness of anchoring effects.

Our data are consistent with the idea that people's initial choices are arbitrary, but subsequent behavior is constrained by previous choices. If consumers remember their previous choices, then they should have crystalized their WTP and WTA for normal market goods. This does not necessarily mean that their initial choices were informed by underlying fundamental preferences. What it does imply is that subjects should be expected to have monetary valuations for the goods that do not depend on irrelevant anchors. Their WTP may have still depended on irrelevant cues in the first time they made consumption choices for these goods.

APPENDIX : Instructions for the Lotteries Experiment (Experiment 2)

Welcome to CASSEL. Please turn off pagers and cellular phones now. It is important that you do not talk, or with any other way try to communicate during the experiment. Thank you for participating in this pricing exercise. We are interested in how much you value various lotteries. We will show you six different lotteries, and then ask you a pair of price-selling intent questions about each lottery:

- First, we will ask whether you would like to sell the lottery at a particular price. Your computer generated that price randomly.
- Next, we will ask you to state the minimum that you would be willing to accept for selling the lottery.

Three persons will be chosen at random from the class and the decisions of each about one of the lotteries will be carried out for real money. One lottery will be sold on the basis of answers to the first question, and two lotteries on the basis of answers to the second question. Each person is on the running for only one lottery.

<u>If you are chosen and the first question counts for you</u>, then we will look to see whether you decided to sell, or not to sell, the lottery at your random computer-generated price. If you stated that you wanted to sell the lottery at that price, then we will buy it from you at that price. If you stated that you did not want to sell the lottery at that price, you will keep it.

<u>If you are chosen and the second question counts</u>:

- The price of the lottery will be determined by drawing a number at random from the matrix in the overhead slide.
- If the value is smaller than this random price, then the lottery will be sold *at that randomly selected price*.

- If the value stated is higher than the price, then the lottery will not be sold.

Because your answer does not affect the amount you get paid, only whether you sell, it's to your advantage to state the minimum that you would be willing to be paid for selling the lottery.

Each of you will receive a 12-dollar participation fee. In addition, if you are one of the three persons chosen, all results from lotteries you have chosen to keep will be drawn right away. Any amount that you will earn from the sale of a lottery, or from the results of a lottery you keep, will be added to your participation fee. You will know your total earnings when you leave the experiment.

First, please copy your computer-generated random number in each of the 6 boxes

LOTTERY TICKET 1

50% chance of	50% chance of
$ **100**	$ **0**

Would you sell this lottery ticket for $ ☐ Circle YES or NO

The least I would be willing to receive to sell this lottery is $ ________

LOTTERY TICKET 2

25% chance of	75% chance of
$ **100**	$ **0**

Would you sell this lottery ticket for $ ☐? Circle YES or NO

The least I would be willing to receive to sell this lottery is $ ________

LOTTERY TICKET 3

10% chance of	90% chance of
$ **100**	$ **0**

Would you sell this lottery ticket for $ ☐. Circle YES or NO

The least I would be willing to receive to sell this lottery is $ ________

LOTTERY TICKET 4

75 % chance of	25% chance of
$ **60**	$ **20**

Would you sell this lottery ticket for $ ☐? Circle YES or NO

The least I would be willing to receive to sell this lottery is $ ________

LOTTERY TICKET 5

50% chance of	50% chance of
$ **60**	$ **20**

Would you sell this lottery ticket for $ ☐? Circle YES or NO

The least I would be willing to receive to sell this lottery is $ ________

LOTTERY TICKET 6

90% chance of	10% chance of
$ **60**	$ **20**

Would you sell this lottery ticket for $ ☐? Circle YES or NO

The least I would be willing to receive to sell this lottery is $ _______

References

Ariely, Dan, George Lowenstein, and Drazen Prelec. " 'Coherent Arbitrariness': Stable Demand Curves without Stable Preferences". The Quarterly Journal of Economics, Vol. 118, No. 1, 2003, pp. 73-104.

Becker, Gordon, Morris DeGroot, and Jacob Marschak. "Measuring Utility by a Single-Response Sequential Method". Behavioral Science 9, No. 3, 1964, pp. 226-232.

Bergman, Oscar, Tore Ellingsen, Magnus Johannesson and Cicek Svensson. "Anchoring and Cognitive Ability". Economic Letters, Vol. 107, No. 1, 2010, pp. 66-68.

Bettman, James, Mary Frances Luce and John Payne. "Constructive Consumer Choice Processes". The Journal of Consumer Research, Vol. 25, No. 3, 1998, pp. 187-217.

Butler, David and Graham Loomes. "Imprecision as an Account of the Preference Reversal Phenomenon". The American Economic Review, Vol. 97, No. 1, 2007, pp. 277-297.

Chapman, Gretchen and Eric Johnson. "Anchoring, Activation, and the Construction of Values". Organizational Behavior and Human Decision Processes, Vol. 79, No. 2, 1999, pp. 115-153.

Cox, James and David Grether. "The Preference Reversal Phenomenon: Response Mode, Markets and Incentives". Economic Theory, Vol. 7, 1996, pp. 381-405.

Grether, David and Charles Plott. "Economic Theory of Choice and the Preference Reversal Phenomenon". The American Economic Review, Vol. 69, No. 4, 1979, pp. 623-638.

Hershey, John and Paul Schoemaker. "Probability Versus Certainty Equivalence Methods in Utility Measurement: Are they Equivalent?". Management Science, Vol. 31, No. 10, 1985, pp. 1213-1231.

Johnson, Eric and David Schkade. "Bias in Utility Assessments: Further Evidence and Explanations". Management Science, Vol. 35, No. 4, 1989, pp. 406-424.

Loomes, Graham and Robert Sugden. "Regret Theory: an Alternative Theory of Rational Choice Under Uncertainty". The Economic Journal, Vol. 92, 1982, pp. 805-824.

MacCrimmon, Kenneth, and Maxwell Smith. "Imprecise Equivalences: Preference Reversals in Money and Probability." University of British Columbia Working Paper 1211, 1986.

Plott, Charless and Kathryn Zeiler. "The Willingness to Pay/ Willingness to Accept Gap, the 'Endowment Effect', Subject Misconceptions and Experimental Procedures for Eliciting Valuation". The American Economic Review, Vol. 95, No. 3, 2005, pp. 530-545.

Plott, Charles and Kathryn Zeiler. "Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?". The American Economic Review, Vol. 97, No. 4, 2007, pp. 1449-1466.

Pommerehne, Weiner, Friedrich Schneider and Peter Zweifel. "Economic Theory of Choice and the Preference Reversal Phenomenon: a Reexamination". The American Economic Review, Vol. 72, No. 3, 1982, pp. 569-574.

Slovic, Paul and Sarah Lichtenstein. "Preference Reversals: a Broader Perspective". The American Economic Review, Vol. 73, No. 4, 1983, pp. 596-605.

Slovic, Paul, Dale Griffin and Amos Tversky. "Compatibility Effects in Judgement and Choice". in R. H. Hogarth (Ed.) Insights in Decision Making: Theory and Applications, pp. 5-27. Chicago, University of Chicago Press 1991.

Tufano, Fabio. "Are 'True' Preferences Revealed in repeated Markets? An Experimental Demonstration of Context-Dependent Valuations". Experimental Economics, Vol. 13, No. 1, pp. 1-13.

Tversky, Amos and Richard Thaler. "Anomalies: Preference Reversals". Journal of Economic Perspectives, Vol. 4, No. 2, 1990, pp. 201-211.