BEE Paper

Konrad Kopp

2024-03-16

Abstract

The abstract is currently pending

Table of contents

Introduction

Intro

Literature review

face value paper (Raghubir and Srivastava 2002): - Arguably, the face value effect is due to the accessibility and perceptual salience of the face value of the foreign currency. Theoretically, therefore, the strength and persistence of the face value effect is likely to depend on the extent to which an individual has the opportunity or the time available to process exchange rate information and/or has experience in using a particular foreign currency. The last two studies show that the time available to process exchange information (study 5) and experience with a foreign currency, both measured (study 5) and manipulated (study 6), moderate the face value effect. - The asymmetric effect may be due to the differential difficulty in using whole numbers versus fractions or multiplying versus dividing - First, the reliance on face value may be a function of the ease with which the foreign money can be converted, implying that as r approaches a round number, a is likely to reduce. Second, a may vary with the absolute deviation of r from one, implying that as r deviates from one, the need to adjust increases implying a nonlinear relationship. From an effort-accuracy perspective, when r is close to one, it may not be worthwhile to adjust. Further, note that for currencies that are multiples of the home currency, r can range from one to , whereas for fractional currencies, r is bounded between zero and one. The lower bound may account for the lower likelihood of adjustment for a currency that is a fraction of the homecurrency. Future research should examine the asymmetric and nonlinear nature of this effect - Despite attempts to make the tasks as realistic as possible, the studies reported were laboratory experiments, and issues of generalizability when real money is on the line do arise. For instance, when people actually exchange foreign currency they may feel richer or poorer depending on the exchange rate, and the differences in perceptions of wealth may affect product valuation.

fang paper (Fang 2019): - did not implement bdm because infeasible

perceived value of money (Wertenbroch, Soman, and Chattopadhyay 2007): - In particular, we extend previous research in several important ways. We propose a conceptualization that not only considers the nominal prices faced by consumers but also takes salient reference values into account, such as the budget consumers have available to make a purchase or the price of competing purchase options. By including budget constraints and reference standards, our conceptualization links nominal valuations to standard economic as well as psychophysical theory. We propose and test two alternative computational mechanisms by which consumers might evaluate prices in different currencies relative to such ,reference standards. Specifically, do consumers use the difference between reference values and prices to derive their evaluations or the ratio of the two?

build on perceived value of money paper; where to extend: - incentive compatible mechanism - control for familiarity of currencies (easy on mind, easy on wallet) (Alter and Oppenheimer 2008) - examine the dispute between it and face value Paper - expand to virtual currencies (fang)

Model

ratio (traditional) model: - formulate valuation of good as ratio to budget in home currency - convert valuation of good into bid currency

assumptions: - is it rational if we always assume that the conversion will be accurate? what if this causes more mental effort than reward? - people can easily do those calculations in their head

if we agree with assumptions, then there should be (small) differences but not systematic ones (see methodology)

face value model: - formulate valuation of good in home currency - convert valuation of good into bid currency but weighted by nominal value of foreign currency

$$V = Vn$$

$$V = Vr$$

$$V = aVn + (1 - a)Vr$$

if this is the right model, then we should see group 1 overspend and group 2 underspend

value of money model: - difference in budget and item price in foreign currency - insufficiently adjust for exchange rate, leading to biased assesment of real value of transaction

if this is the right model, then we should see group 1 underspend and group 2 overspend

Method

There were 41 participants of the study, most of them current undergraduate students at the University of Oxford, but some participants were also recent graduates and postgraduate students. The study took place online, using Qualtrics, and students were incentivised to participate by taking part in a lottery, which will be explained in further detail below.

Study

The study itself was split into four parts: Consent Form, Introduction, Auction and Follow-Up Questions. The entire experiment instructions can be found in the Experiment Instructions section of the Appendix. The first section asked participants to read through and sign the consent form in order to ensure that they were informed about the study and what data would be collected. In the introductory section, the BDM mechanism was explained to participants and they were asked to complete four question that tested their understanding of the mechanism. The aim of this section was to ensure that participants properly understood how the mechanism worked in order to elicit their real WTP during the main part of the study.

The auction was the main part of the study, with the aim of eliciting the participants' WTP on a specific item and comparing these across different currencies used. In this section, participants were randomly split into two groups with almost-identical instructions. The participants were instructed to bid on the item shown below, a tabletop airhockey table, using the virtual currency "BlueCoin". The auction was explained to follow the BDM mechanism and it was also stated that three participants of the experiment would be randomly selected, their auction would be simulated and they would receive the outcome. The only difference between the groups was the exchange rate between BlueCoin and the British Pound (£) and the endowment that the participants received in BlueCoin. The first group received an endowment of 10 BlueCoin where one BlueCoin equals £5 and the second group recieved an endowment of 250 BlueCoin where one BlueCoin equals £0.2. To complete the auction, participants were asked to respond with their bid, in BlueCoin.

The final part of the experiment consisted of optional follow-up questions that were used to determine if there are any trends in bidding behaviour based on some characteristics of the participants. The questions asked were about whether they had previously lived in a different country than the UK, how often they tend to travel outside of the UK and what their typical expenditure is in a given month. The intention behind the first two questions is to use them as a proxy for experience dealing with foreign currencies, where presumably people that travel frequently or have lived in multiple countries are more experienced with currency conversion calculations. The final question was asked to be able

to determine whether there were any systematic differences in WTP based on a participants normal expenditure.

Auction Design

Since the aim of the study was to determine whether there were any systematic differences in WTP between consumers when the only differing factor is the currency used, we decided to not include a control group that bid on the item using Pounds, which was the design of some of the experiments run by RAS (Raghubir and Srivastava 2002). This allowed for a simpler design and more power given the same sample size. Further, unlike the studies run by RAS (Raghubir and Srivastava 2002) and WER (Wertenbroch, Soman, and Chattopadhyay 2007), we did not use a fiat currency or mix of fiat and fictional currencies, but used only one fictional currency for both groups. On the one hand, this simplification was due to a focus of wether the effects described in previous papers also held in similar ways for virtual currencies. On the other hand, it also allowed us to preclude any biases or previous experience that participants might have had towards or with certain currencies. In the course of this, we chose the relatively undescript name for a fictional, virtual currency: BlueCoin.

Another important decision in the design of the experiment was which item to select that a majority of participants would have a non-zero WTP. This is important since if the majority or even the entirety of participants were to bid zero on the item, then we would only learn that this was the participants' WTP for that good, but not whether there are any systematic differences in the WTP when using different currencies. Hence, we selected a tabletop airhockey table, something that few enough people have so that they won't want another one, but enough people want with a non-zero WTP, especially among a student population. In fact, only six participants of the experiment stated a WTP of 0.

In designing the main part of the experiment, we chose to conduct an auction, specifically using the BDM mechanism, in order to have an incentive-compatible method of eliciting the participants' WTP. The BDM's incentive compatibility in eliciting WTP is a well-established phenomenon that is superior to using participants' stated preferences (**cite?**). As mentioned above, the studies conducted by RAS (Raghubir and Srivastava 2002), WER (Wertenbroch, Soman, and Chattopadhyay 2007) (with the exception of one permutation of one of the studies) and Fang (Fang 2019) did not implement incentive-compatible methods of eliciting the participants' preferences, making this experiment and important addition to these previous studies in examining whether the effects describe still hold when a more accurate way of eliciting WTP is used. In order to incentivise the participants even further, we ran a lottery that paid out three randomly-chosen participants based on their actual choices made in the auction and by simulating the BDM mechanism using a random number generator. Due to budget constraints, we were only able to play this lottery for three people, but given that the sample size of the experiment was relatively small, the potential to receive one's outcome is likely to have helped in eliciting the participants' true WTP.

Finally, given our budget, we were able to give participants an endowment of £50 each, which was around double of what the item costs to buy. Assuming that the item is priced roughly around the average WTP (which turned out to be correct), this endowment was chosen to leave enough room for people with higher WTP to make their bid and for us to capture this in the data. This is important for the goals of the experiment, since if we had, for example, chosen an endowment of £20, the majority of participants might have bid their maximum amount, hindering our ability to meaningfully compare WTP across the two groups. We chose relatively simple exchange rates of 5 and 1/5 so that the

exchange rate calculations would be relatively easy to make, allowing us to be more confident that differences between groups were not due to the mental effort required to do calculations but due to something else.

Limitations

In setting up the experiment, we incorrectly used the BDM mechanism. Instead of explaining it as a sealed-bid second-price auction, the experiment instructions explained it as a sealed-bid first-price auction. Unlike in a second-price auction, the participants of a first price auction do not have a dominant strategy to bid their true valuations, in fact there generally is no dominant strategy in such an auction (**cite?**). (**todo?**)

Another limitation of the study is the low sample size of 41 participants, which entails lower statistical power and thus a lower likelihood of detecting a true positive when an effect actually exists. I will examine this in further detail in the next section, conducting a power analysis and estimating how big the sample size should have been to achieve an adequate level of power.

Results

There were 41 participants, 21 in the first group and 20 in the second. For consistency, I will refer to the first group as the fraction group, since 0.2 BlueCoin = £1, and the second as the multiple group, since 5 BlueCoin = £1, in line with the terminology used by RAS (Raghubir and Srivastava 2002). To conduct the data analysis, I normalised the bids across both groups by converting them into pounds and computed the following summary statistics:

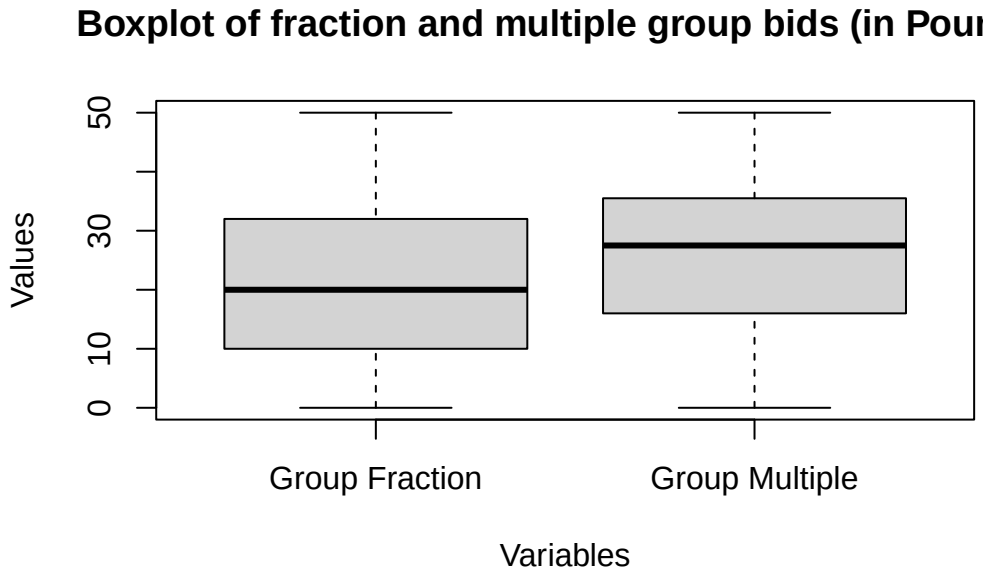
```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   0.00   10.00   20.00   19.74   32.00   50.00      20
```

```
[1] 15.49001
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
   0.00   16.50   27.50   25.57   35.25   50.00      21
```

```
[1] 14.8083
```

Further, I plotted the following boxplot, in order to visually see the difference between the groups:



As is obvious from the summary statistics and the boxplot, there is a difference between both groups, namely that the bids in the multiple group seem to be higher on average than those in the fraction group. To investigate this hypothesis, I conducted a t-test of a difference in means between the groups, specifically a Welch Two Sample t-test. The null hypothesis is that the difference in means is equal to 0 (or, alternatively, that the means are the same) and the alternative hypothesis is that this difference is not equal to 0 (or, alternatively, that the means are not the same):

$$h_0 : {}_1 = {}_2$$
$$h_1 : {}_1 \neq {}_2$$

Hence, I conducted the t-test:

```
    Welch Two Sample t-test

data:  experiment_data$group_fraction_bid_in_pounds and experiment_data$group_multiple_bid_in_pounds
t = -1.2323, df = 38.999, p-value = 0.2252
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -15.401901   3.740092
sample estimates:
mean of x mean of y
  19.7381   25.5690
```

The t-value of -1.2323 suggests that there is a difference in the means between the groups and, since it is negative, that the mean of the fraction group is lower. However, the p-value is 0.2252, which is quite high. If we take the typical significance level of $\alpha = 0.05$ we find that the p-value is higher than this value, and thus fail to reject the null hypothesis at the 5% significance level. Further, the 95% confidence interval is $[-15.401901, 3.740092]$ which includes 0, providing further evidence that the difference in means between the groups is not statistically significant. Thus, we must conclude that while there seems to be a difference in means based on the sample data, this difference is not statistically significant for this sample size.

To investigate this further, I have conducted a power analysis in order to determine the power of the t-test, which refers to the likelihood of the test detecting a true positive when an effect actually exists (**cite?**). High power would indicate that this likelihood is high and vice versa. To conduct the analysis, we first need to find the effect size, or Cohen's D and then implement the power analysis. Further, I will also conduct an analysis that investigates how big the sample size should have been to get a power of 80%.

```
     t test power calculation

             n1 = 21
             n2 = 20
              d = 0.3845806
      sig.level = 0.05
          power = 0.2246363
    alternative = two.sided


     Two-sample t test power calculation

              n = 107.1047
              d = 0.3845806
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

Based on the first test with the same significance level of 0.05, we can see that the power is 0.2246363 or 22%, which means that there is only a 22% likelihood of detecting a true positive if an effect actually exists. In other words, the likelihood of detecting a true difference in means between the two groups, when this difference exists, is only 22%, which is far below the usual value of 80%. The second test takes in this power of 80% and calculates the required sample size at the observed effect size. The result is $n = 107$ where the sample size is $2n$ or 214. Hence, using a sample size of 214 and observing the same effect, this t-test would yield a power of 80% at the 5% significance level.

Finally, I conducted a regression of the demographic values collected in the fourth part of the experiment on the bids of each group. This is done in order to determine whether there is any correlation between the demographic variables and the bid amount, which would suggest that randomisation did not occur correctly, which could have been exacerbated by the small sample size.

The regression yields the following outcomes:

```
Call:
lm(formula = group_fraction_bid_in_pounds ~ lived_abroad + travel_frequency +
    expenditure, data = experiment_data)

Residuals:
    Min      1Q  Median      3Q     Max
-20.744 -10.085   0.000   7.481  33.333

Coefficients: (2 not defined because of singularities)
                                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                                 16.667     11.144   1.496    0.169
lived_abroadNo                              16.883     34.290   0.492    0.634
lived_abroadYes                              4.651     28.453   0.163    0.874
travel_frequencyLess than once per 5 years  12.768     31.087   0.411    0.691
travel_frequencyOnce per 1-2 years          -4.952     27.679  -0.179    0.862
travel_frequencyOnce per 2-5 months          5.085     23.987   0.212    0.837
travel_frequencyOnce per 3-5 years         -18.005     42.079  -0.428    0.679
travel_frequencyOnce per 6-11 months        -6.488     24.395  -0.266    0.796
travel_frequencyOnce per month or more          NA         NA      NA       NA
expenditure£251-£400                        -6.317     17.686  -0.357    0.729
expenditure£401-£550                        -5.544     19.548  -0.284    0.783
expenditure£551-£700                        -8.884     20.685  -0.429    0.678
expenditureAbove £1000                      -13.549     37.738  -0.359    0.728
expenditureBelow £250                           NA         NA      NA       NA

Residual standard error: 19.3 on 9 degrees of freedom
  (20 observations deleted due to missingness)
Multiple R-squared:  0.3013,    Adjusted R-squared:  -0.5528
F-statistic: 0.3527 on 11 and 9 DF,  p-value: 0.9466


Call:
lm(formula = group_multiple_bid_in_pounds ~ lived_abroad + travel_frequency +
    expenditure, data = experiment_data)

Residuals:
    Min      1Q  Median      3Q     Max
-15.667  -4.550   0.000   5.383  17.200

Coefficients: (2 not defined because of singularities)
                                    Estimate Std. Error t value Pr(>|t|)
(Intercept)                           20.000     13.715   1.458   0.1788
lived_abroadNo                        -0.600     17.250  -0.035   0.9730
lived_abroadYes                       17.600     17.250   1.020   0.3342
```

```
travel_frequencyOnce per 1-2 years        4.600     21.843   0.211   0.8379
travel_frequencyOnce per 2-5 months      12.400     17.250   0.719   0.4905
travel_frequencyOnce per 6-11 months      7.267     19.717   0.369   0.7210
travel_frequencyOnce per month or more       NA         NA      NA       NA
expenditure£251-£400                     -5.000     18.025  -0.277   0.7877
expenditure£401-£550                    -17.200     15.801  -1.089   0.3046
expenditure£551-£700                    -43.487     21.461  -2.026   0.0734 .
expenditure£701-£850                     -8.400     17.250  -0.487   0.6379
expenditureAbove £1000                  -25.000     19.396  -1.289   0.2296
expenditureBelow £250                        NA         NA      NA       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 13.72 on 9 degrees of freedom
  (21 observations deleted due to missingness)
Multiple R-squared:  0.5937,    Adjusted R-squared:  0.1422
F-statistic: 1.315 on 10 and 9 DF,  p-value: 0.3456
```

Based on this regression, there are no statistically significant correlations at the 5% significance level or lower. However, there is one correlation that is significant at the 10% significance level, namely of `expenditure£551-£700` on the multiple group. The coefficient of -43.487 suggests that people whose usual expenditure is between £551 and £700 per month bid lower than the group overall. However, given that this is only significant at the 10% level and not at any lower ones, it is likely that this correlation is due to the low sample size rather than representing a true correlation between these variables.

Discussion

outline the results: - group one bid lower on average; statistically significant (but only power of 20%) - no statistically significant relationships between domographic vars and bids found

how to explain results: value of money paper suggests this, confirms this research, face value effect stipulates differernt result but seems to be incorrect here

explanations for effect: - model 3 is the correct one - anchoring: higher endowment anchors at higher endowment value (perceived value of money paper: When provided with a budget, consumers might anchor on the nominal value of that budget when evaluating transactions in foreign currencies. But this would not have predicted the biased evaluations in our follow-up study since the nominal budget was the same (i.e., quoted in US$) across all numerosity conditions) - people are worse at dividing than multiplying (or put in less effort) (or reverse) - people prefer using whole numbers over fractions - is this rational? mental ease? - sample size too small - power analysis

conclusion: agree with model 3 and get rid of some of its shortcomings: - incentive compatible mechanism - control for familiarity of currencies (easy on mind, easy on wallet) - examine the dispute between it and face value Paper - expand to virtual currencies (fang)

Summary

Limitations, Theoretical Implications and Future Research

References

Alter, Adam L., and Daniel M. Oppenheimer. 2008. 'Easy on the Mind, Easy on the Wallet: The Roles of Familiarity and Processing Fluency in Valuation Judgments'. *Psychonomic Bulletin & Review* 15 (5): 985–90. https://doi.org/10.3758/PBR.15.5.985.

Fang, Justin. 2019. 'In-Game Currency Design and Consumer Spending Behavior'. Bachelor's Thesis, Ross School of Business, University of Michigan. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/155343/Justin%20Fang_BA%20480%20Written%20Report.pdf.

Raghubir, Priya, and Joydeep Srivastava. 2002. 'Effect of Face Value on Product Valuation in Foreign Currencies'. *Journal of Consumer Research* 29 (3): 335–47. https://doi.org/10.1086/344430.

Wertenbroch, Klaus, Dilip Soman, and Amitava Chattopadhyay. 2007. 'On the Perceived Value of Money: The Reference Dependence of Currency Numerosity Effects'. *Journal of Consumer Research* 34 (1): 1–10. https://doi.org/10.1086/513041.

Appendix

Experiment Instructions

Source Code