Poverty and Economic Decision-Making

Do changes in one’s financial circumstances affect one’s decision-making process and cognitive capacity? In an experimental study, researchers randomly selected a group of US respondents to be surveyed before their payday and another group to be surveyed after their payday. Under this design, the respondents of the Before Payday group are more likely to be financially strained than those of the After Payday group. The researchers were interested in investigating whether or not changes in people’s financial circumstances affect their decision making and cognitive performance. Other researchers have found that scarcity induce an additional mental load that impedes cognitive capacity. This exercise is based on:

Carvalho, Leandro S., Meier, Stephen, and Wang, Stephanie W. (2016). “[Poverty and economic decision-making: Evidence from changes in financial resources at payday.](http://dx.doi.org/10.1257/aer.20140481)” *American Economic Review*, Vol. 106, No. 2, pp. 260-284.

In this study, the researchers administered a number of decision-making and cognitive performance tasks to the Before Payday and After Payday groups. We focus on the *numerical stroop task*, which measures cognitive control. In general, taking more time to complete this task indicates less cognitive control and reduced cognitive ability. They also measured the amount of cash the respondents have, the amount in their checking and saving accounts, and the amount of money spent. The data set is in the CSV file poverty.csv. The names and descriptions of variables are given below:

|  |  |
| --- | --- |
| Name | Description |
| treatment | Treatment conditions: Before Payday and After Payday |
| cash | Amount of cash respondent has on hand |
| accts\_amt | Amount in checking and saving accounts |
| stroop\_time | Log-transformed average response time for cognitive stroop test |
| income\_less20k | Binary variable: 1 if respondent earns less than 20k a year and 0 otherwise |

## Question 1

Load the poverty.csv data set. Look at a summary of the poverty data set to get a sense of what its variables looks like. Use histograms to examine the univariate distributions of the two financial resources measures: cash and accts\_amt. What can we tell about these variables’ distributions from looking at the histograms? Evaluate what the shape of these distributions could imply for the authors’ experimental design.

Now, take the *natural logarithm* of these two variables and plot the histograms of these tranformed variables. How does the distribution look now? What are the advantages and disadvantages of transforming the data in this way?

**NOTE:** Since the natural logarithm of 0 is undefined, researchers often add a small value (in this case, we will use $1 so that ) to the 0 values for the variables being transformed (in this case, cash and accts\_amt) in order to successfully apply the log() function to all values. Be sure to do this recoding only for the purposes of taking the logarithmic transformation – keep the original variables the same.

## Answer 1

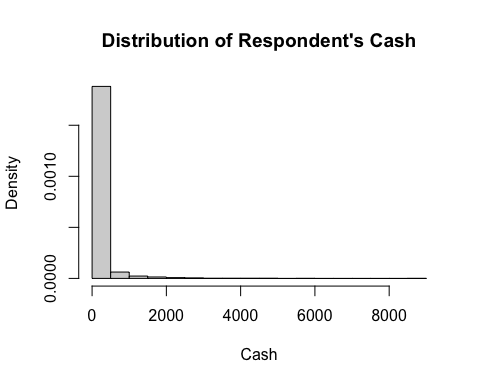
We can tell by looking at the univariate distribution that the finanical resources of the respondents is heavily skewed towards being poorer. However, there were a few respondents that had a lot of money >$50,000 in their checking and >$8,000 in cash. This means that the authors of the data might not have gotten a wholly normal or randomly distributed sample of respondents. However, if it is a representative sample depends on the distribution of the population they were trying to test.

In the natural log functions, the data is presented a little more cleanly. The data becomes more normalized (although not entirely normal, as we can see from the slight skew), which helps when we want to do analysis of the data. It makes for a cleaner looking graph, which assists in our understand of how the data lies. The disadvantage is it’s harder for us to know the exact numbers that the data is showing us by looking at the graph.

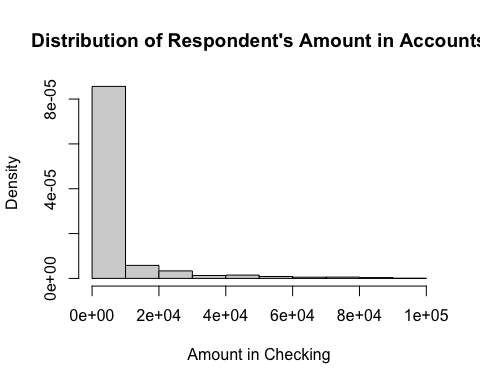
poverty <- read.csv("/Users/clayparham/Documents/Bush631-/Test 1/poverty-1.csv")  
summary(poverty)

## treatment cash accts\_amt stroop\_time   
## Length:2670 Min. : 0.0 Min. : 0 Min. :5.356   
## Class :character 1st Qu.: 15.0 1st Qu.: 176 1st Qu.:7.436   
## Mode :character Median : 49.5 Median : 1000 Median :7.564   
## Mean : 169.0 Mean : 6211 Mean :7.545   
## 3rd Qu.: 136.2 3rd Qu.: 5000 3rd Qu.:7.692   
## Max. :9000.0 Max. :95000 Max. :8.517   
## NA's :226 NA's :433   
## income\_less20k   
## Min. :0.0000   
## 1st Qu.:0.0000   
## Median :0.0000   
## Mean :0.4139   
## 3rd Qu.:1.0000   
## Max. :1.0000   
##

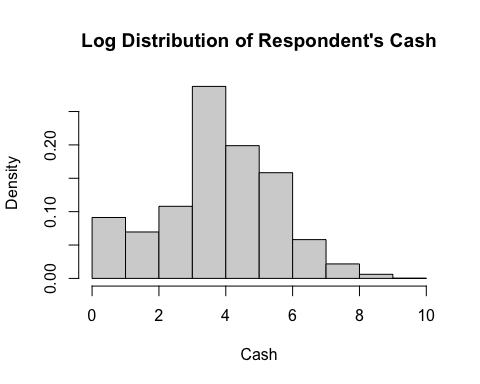
hist(poverty$cash, freq= FALSE, xlab = "Cash", main = "Distribution of Respondent's Cash")



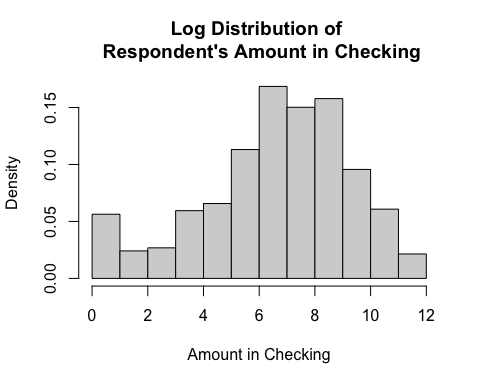
hist(poverty$accts\_amt, freq= FALSE, xlab = "Amount in Checking", main = "Distribution of Respondent's Amount in Accounts")



hist(log(poverty$cash + 1), freq= FALSE, xlab = "Cash", main = "Log Distribution of Respondent's Cash")



hist(log(poverty$accts\_amt + 1), freq= FALSE, xlab = "Amount in Checking", main = "Log Distribution of \n Respondent's Amount in Checking") #added one to log value to factor in those who had $0 in their bank accounts.



## Question 2

Now, let’s examine the primary outcome of interest for this study– the effect of a change in financial situation (in this case, getting paid on payday) on economic decision-making and cognitive performance. Begin by calculating the treatment effect for the stroop\_time variable (a log-transformed variable of the average response time for the stroop cognitive test), using first the mean and then the median. What does this tell you about differences in the outcome across the two experimental conditions?

Secondly, let’s look at the relationship between finanical circumstances and the cognitive test variable. Produce two scatter plots side by side (hint: use the par(mfrow)) before your plot commands to place graphs side-by-side), one for each of the two experimental conditions, showing the bivariate relationship between your *log-transformed* cash variable and the amount of time it took subjects to complete the stroop cognitive test administered in the survey (stroop\_time). Place the stroop\_time variable on the y-axis. Be sure to title your graphs to differentiate between the Before Payday and After Payday conditions. Now do the same, for the *log-transformed* accts\_amt variable.

Briefly comment on your results in light of the hypothesis that changes in economic circumstances will influence cognitive performance.

## Answer 2

We see that there is a small change in both the median and the mean when it comes to if there is a noticeable affect pre- or post- payday. The log-adjusted mean score increases by around .011 before and after payday and the median by about .014. This may signal that there is a correlation between the scores before and after payday. The graphs are fairly evenly distributed across the amount of money, but the scores are still mostly around 7, indicating that the amount of money may not play a large role in the scores.

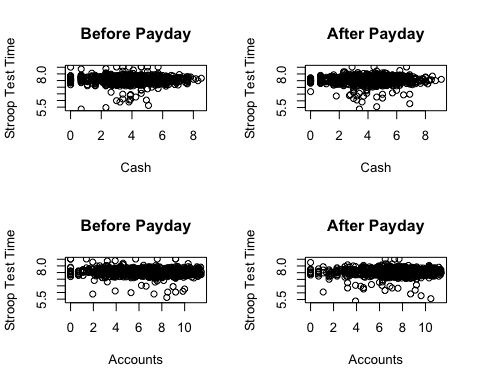
beforepayday <- subset(poverty, treatment == "Before Payday")  
afterpayday <- subset(poverty, treatment == "After Payday")  
mean(afterpayday$stroop\_time) - mean(beforepayday$stroop\_time)

## [1] 0.01142325

median(afterpayday$stroop\_time) - median(beforepayday$stroop\_time)

## [1] 0.01430528

par(mfrow = c(2,2)) #received error message when attempting side-by-side. It says "figure margins too large"  
plot(log(beforepayday$cash), beforepayday$stroop\_time, type = "p", main = "Before Payday", xlab = "Cash", ylab = "Stroop Test Time")  
plot(log(afterpayday$cash), afterpayday$stroop\_time, type = "p", main = "After Payday", xlab = "Cash", ylab = "Stroop Test Time")  
  
#par(mfrow = c(2,1)) #same with this one  
plot(log(beforepayday$accts\_amt), beforepayday$stroop\_time, type = "p", main = "Before Payday", xlab = "Accounts", ylab = "Stroop Test Time")  
plot(log(afterpayday$accts\_amt), afterpayday$stroop\_time, type = "p", main = "After Payday", xlab = "Accounts", ylab = "Stroop Test Time")



## Question 3

Now, let’s take a closer look at whether or not the Before Payday versus After Payday treatment created measurable differences in financial circumstances. What is the effect of payday on participants’ financial resources? To help with interpretability, use the original variables cash and accts\_amt to calculate this effect. Calculate both the mean and median effect. Does the measure of central tendency you use affect your perception of the effect?

## Answer 3

The amount of cash increased in both the mean and median. It increased in the mean by around $37 and in the median by around $10. However, for the amount in the checking account, the mean decreased after payday by almost $300, while the median increased by around $450. This is very surprising to me, as I would expect the mean to go up after payday. This alters greatly my interpretation of the data, since the outliers must greatly shift the data.

mean(beforepayday$cash, na.rm = TRUE) - mean(afterpayday$cash, na.rm = TRUE)

## [1] -36.77511

mean(beforepayday$accts\_amt, na.rm = TRUE) - mean(afterpayday$accts\_amt, na.rm = TRUE)

## [1] 251.9087

median(beforepayday$cash, na.rm = TRUE) - median(afterpayday$cash, na.rm = TRUE)

## [1] -10

median(beforepayday$accts\_amt, na.rm = TRUE) - median(afterpayday$accts\_amt, na.rm = TRUE)

## [1] -450

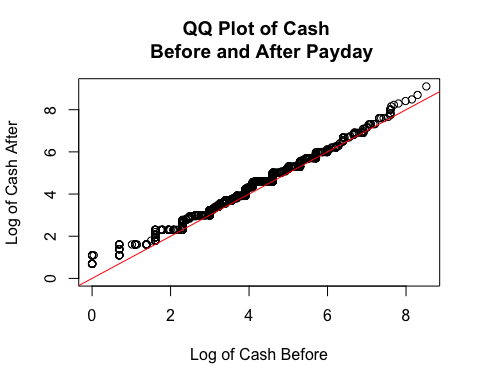
## Question 4

Compare the distributions of the Before Payday and After Payday groups for the *log-transformed* cash and accts\_amt variables. Use quantile-quantile plots to do this comparison, and add a 45-degree line in a color of your choice (not black). Briefly interpret your results and their implications for the authors’ argument that their study generated variation in financial resources before and after payday. When appropriate, state which ranges of the outcome variables you would focus on when comparing decision-making and cognitive capacity across these two treatment conditions.

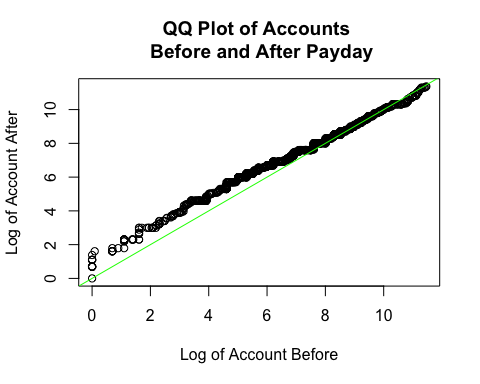
## Answer 4

The data seems to follow fairly closely the trend 45 degree line. This shows that the data is normally fairly distributed both before and after, which can help show that the authors can use the data to show that changes are possible before and after the information. This means that the distribution of people did not meaningfully shift between having before and after payday. This can help the researchers know that they can test on both groups before and after payday. I would focus on the outcome variables that are closest to the line. In this case, those are the outcomes that tend to be in the middle of the graphs.

qqplot(log(beforepayday$cash), log(afterpayday$cash), main = "QQ Plot of Cash \n Before and After Payday", xlab = "Log of Cash Before", ylab = "Log of Cash After")  
abline(a = 0, b = 1, col = "red")



qqplot(log(beforepayday$accts\_amt), log(afterpayday$accts\_amt), main = "QQ Plot of Accounts \n Before and After Payday", xlab = "Log of Account Before", ylab = "Log of Account After")  
abline(a = 0, b = 1, col = "Green")



## Question 5

In class, we covered the difference-in-difference design for comparing average treatment effects across treatment and control groups. This design can also be used to compare average treatment effects across different ranges of a *pre-treatment variable*- a variable that asks about people’s circumstances before the treatment and thus could not be affected by the treatment. This is known as *heterogeneous treatment effects* – the idea that the treatment may have differential effects for different subpopulations. Let’s look at the pre-treatment variable income\_less20k. Calculate the treatment effect of Payday on amount in checking and savings accounts separately for respondents earning more than 20,000 dollars a year and those earning less than 20,000 dollars. Use the original accts\_amt variable for this calculation. Then take the difference between the effects you calculate. What does this comparison tell you about how payday affects the amount that people have in their accounts? Are you convinced by the authors’ main finding from Question 2 in light of your investigation of their success in manipulating cash and account balances before and after payday?

## Answer 5

It seems that for people making less than 20k, they have more money in their accounts before payday than they do after payday. This explains some of the discord in the numbers from above. It also makes me question the validity of the random selection. It does make me question their main finding. If the data for people making more than 20k decreases from before to after payday, that might signal that any changes, or lack of changes, in the score from before and after may be signalling issues in how the data is distributed.

lessthan20k <- subset(poverty, income\_less20k == "1")  
morethan20k <- subset(poverty, income\_less20k == "0")  
  
mean(lessthan20k$accts\_amt[lessthan20k$treatment == "After Payday"], na.rm = TRUE) - mean(lessthan20k$accts\_amt[lessthan20k$treatment == "Before Payday"], na.rm = TRUE)

## [1] 81.0914

mean(morethan20k$accts\_amt[morethan20k$treatment == "After Payday"], na.rm = TRUE) - mean(morethan20k$accts\_amt[morethan20k$treatment == "Before Payday"], na.rm = TRUE)

## [1] -649.9476

(mean(morethan20k$accts\_amt[morethan20k$treatment == "After Payday"], na.rm = TRUE) - mean(morethan20k$accts\_amt[morethan20k$treatment == "Before Payday"], na.rm = TRUE)) - (mean(lessthan20k$accts\_amt[lessthan20k$treatment == "After Payday"], na.rm = TRUE) - mean(lessthan20k$accts\_amt[lessthan20k$treatment == "Before Payday"], na.rm = TRUE))

## [1] -731.039