Final Project

E ß)wJ i wekeq s

Hi gi q f i v 645<

R Markdown

Introduction

Abstract

This analysis seeks to determine if people who are exposed to social norms or empirical norms will punish more frequently than people who are not shown these norms. For example, does the phrase "most participants in a previous study believe that you should punish people who lie" cause more punishments than the the phrase "most participants in a previous study punished people who lie"? In this study there were three treatment groups (social norms, empirical and control) and the frequency of punishment for lying behavior is being analyzed. Overall the study found that the data trended that social punished more than empirical which punished more than control, but the results were not statistically significant.

Motivation and Experimental Design

Our student-run study intends to examine how empirical versus normative information impacts the willingness to punish deviant behavior in a dice roll game. In more detail, the first stage of our experiment had participants roll a dice, and the number on the die equalled the number of tickets the participant recieved in a raffle for a $20 giftcard. Therefore, some of the participants lied about the number on the die to increase their chance of winning the raffle. Next, the experiment presented other participants with 2-4 deceitful decisions by previous participants. These new participants were given the option to punish (or take away tickets) from the previous participants. A third of the participants were presented with social normative information, a third were presented with empirical information and the control group was not given any statements. Empirical information means that our behavior depends on how we think others behave in a given situation. Social normative information means an individual's behavior is impacted by how they believe others expect them to behave in a given situation (Bicchierri, 2016). Essentially, this experiment is trying to test if people punish cheaters more when they believe that "other people punished cheaters" or if they believe that "other people expect you to punish cheaters". Based on previous literature we believe that normative information will increase punishments.

Hypothesis: $\Pi_{social} > \Pi_{empirical} > \Pi_{control}$.

Null : $\Pi_{social} = \Pi_{empirical} = \Pi_{control}$.

The extension to analyze punisher behavior is novel and interesting for both academics and policy-makers. There is literature on how normative information impacts cheating behavior, but there is limited academic information about how punishers are impacted by normative information (Bicchierri et al., 2018). Therefore, this experiment and the statistical analysis will be interesting for academics. In terms of policy applications, it is important to encourage people to report to management when there is deviant behavior occurring in an organization.Therefore, regulators may want to understand if presenting employees with normative information can encourage them to come forward and punish bad behavior.

Note: Original project was conducted in BDS 501 with Cole Hammond and Marco Bauer. This may explain similiarites between data, analysis and motivation. Please let us know if you have any questions or concerns.

Analytical Strategy

Data

The data was all gathered at the University of Pennsylvania and all participants were students. Data was collected using a qualtrics survey and there was a comprehension check to ensure that participants understood the experiment.

After the data was cleaned to remove participants who did not understand the experiment, there were 80 participants in total. The breakdown by treatment was:

Treatment	# Participants
Control	24
Empirical Information	30
Social Norms	26

The average age was 21.1 and the median age was 20. The treatments were fairly evenly split on gender, which will be discussed further later in this analysis.

Treatment	Males	Females	Other or N/A
Control Treatment	58.3%	41.7%	0%
Empirical Information Treatment	50.0%	46.7%	3.3%
Social Norm Treatment	46.15%	50%	3.9%
Total	51.25%	46.25%	2.5%

When analyzing frequency of punishment data, all data points are a percentage. Participants were shown several lies and the frequency measure was determined by the number of lies they chose to punish out of the total possible lies they could have punished. In the following analyses the dependent variable is the frequency of punishment and the independent variable is the normative information the participant was shown. The dependent variable is measured on an ratio scale (percentages have intervals and a meaningful zero) and the independent variable is categorical.

Methods

Hypothesis Testing

In order to test the hypothesis that participants who were given social normative information will punish more than those who were shown empirical information who will punish more than the control. Again our hypothesis is that:

$\Pi_{qmax\_j} > \Pi_{ck\ ngoga\_j} > \Pi_{amh\ rpmji}$.

First, I will graph the data to understand the absolute differences between the treatments.

```
barplot(c(mean(newcontrol$X..lies.punished)*100, mean(Empirical$X..lies.punished)*100,me
an(Social$X..lies.punished )*100), main = "Percent of Lies Punished ", xlab = "Treatmen
t", ylab = "Percent", col = "#95001a", names.arg = c("Control", "Empirical", "Social"),
 ylim = c(0,100))
```

Percent of Lies Punished
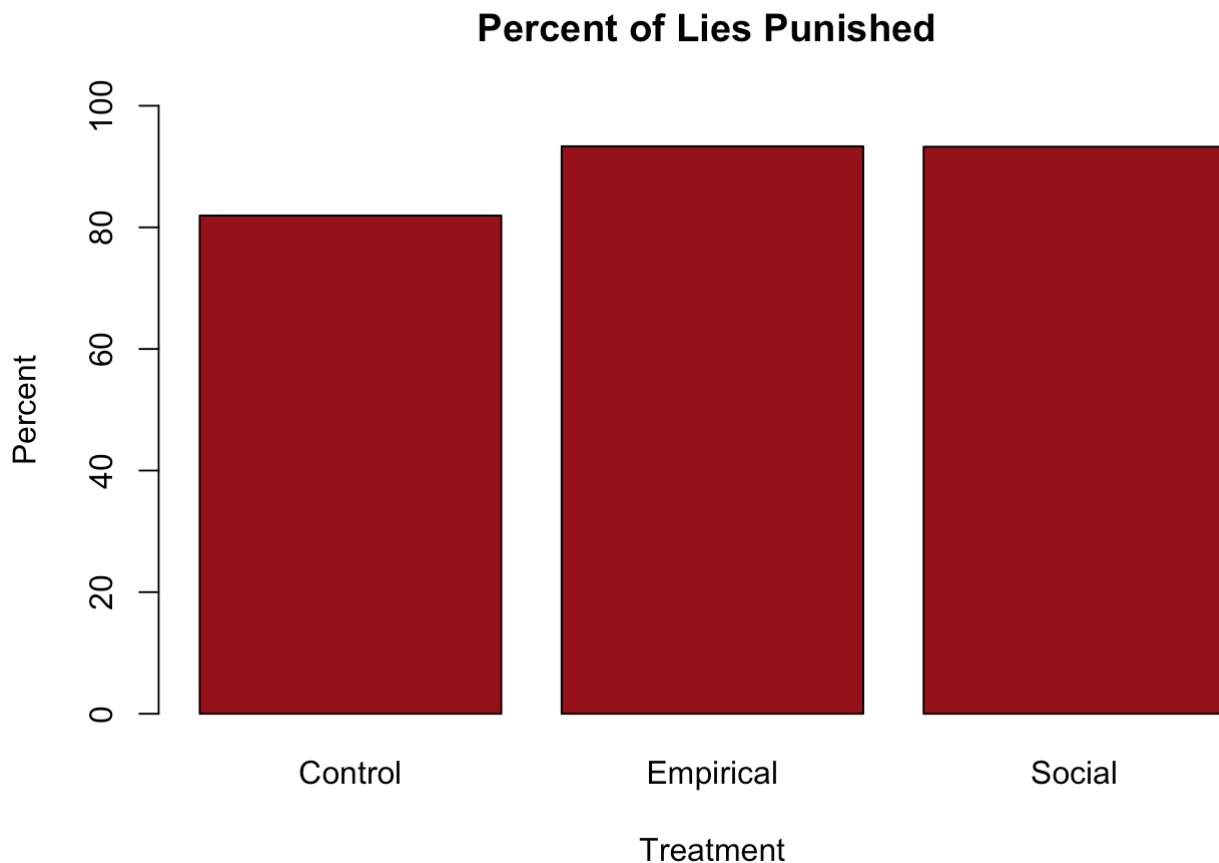
From this graph, it is clear there is not a large difference between any of the treatments, but it is necessary to test for significance. First, I will use a Kruskal Wallis test between all three groups and then will cross-validate this by using three pair-wise Mann Whitney tests to test the difference between control/ social, control/ empirical, social/empirical. I chose to primarily use the Mann Whitney U tests, because I am looking to find a difference between two distributions in terms of their location (mean) and the dataset cannot be assumed to be parametric. The data works for a Mann Whitney U test because the dependent variable is an interval scale and the independent is categorical with at least two categories.

Based on the analysis below, it can be concluded that there is no significant difference between any of the normative treatments and the frequency of punishment. In order for these results to be significant the p-value would have to be below 0.05, which means that there is less than a 5% chance of getting these results if the null is true.

```
KruskalTest
```

```
##
##  Kruskal-Wallis rank sum test
##
## data:  punishers$X..lies.punished by punishers$Treatment
## Kruskal-Wallis chi-squared = 2.4867, df = 3, p-value = 0.4777
```

Control_vs_Social

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  newcontrol$X..lies.punished and Social$X..lies.punished
## W = 279.5, p-value = 0.3297
## alternative hypothesis: true location shift is not equal to 0
```

Control_vs_Empirical

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  newcontrol$X..lies.punished and Empirical$X..lies.punished
## W = 310, p-value = 0.1391
## alternative hypothesis: true location shift is not equal to 0
```

Empirical_vs_Social

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  Empirical$X..lies.punished and Social$X..lies.punished
## W = 407, p-value = 0.5836
## alternative hypothesis: true location shift is not equal to 0
```

Gender Analysis

Previous literature has found that men may be more willing to punish than women, therefore it was important to run a gender analysis on the data to check for potential gender-based effects which would confound potential results (Eckel et al., 1996). First, I determine the gender makeup of the overall participant set and each treatment.

Treatment	Males	Females	Other or N/A
Control Treatment	58.3%	41.7%	0%
Empirical Information Treatment	50.0%	46.7%	3.3%
Social Norm Treatment	46.15%	50%	3.9%
Total	51.25%	46.25%	2.5%

Overall there is some variation between men and women across treatments, but it is fairly evenly split. However, the gender difference only matters if men and women behave differently when exposed to norms. Therefore, next I ran a Mann-Whitney U test to test the difference in distribution location of punishment behavior for men and women across all treatments. Again, I chose this test because I am testing a difference of non-parametric data that has categorical independent variable (male/ female) and interval scale dependent variables (frequency of punishment). Overall, there was not a difference between men and women in punishment behavior across all treatment.

Gender Difference in Punishment Behavior

Punishment Type	Difference Between Men and Women	P-Value
Frequency	6.65% (Men > Women)	0.34
Magnitude	14.32% (Women > Men)	0.51

Please see below for more detail.

```
Men_vs_WomenFrequency
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  females$X..lies.punished and males$X..lies.punished
## W = 1372, p-value = 0.3427
## alternative hypothesis: true location shift is not equal to 0
```

This analysis tested the diference in distributions between how frequently men punish and how frequently women punish across all treatments. This test found that there is no difference between the genders.

```
Men_vs_WomenMagnitude
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  females$X..Magnitude.of.Lies.Punished and males$X..Magnitude.of.Lies.Punished
## W = 1203, p-value = 0.5128
## alternative hypothesis: true location shift is not equal to 0
```

This analysis tested the difference in means between how much men punish (in relation to the lie they saw) and how much women punish across all treatments. This test found that there is no difference between genders.

Finally, to ensure that there is no gender effect between treatments, a similar set of tests was run between men and women in each treatment. Since this involved six Mann-Whitney U tests, output of the tests is not provided but the results are displayed in the table below.

Frequency of Punishment by Gender per Treatment

Treatment	Difference Between Men and Women	P-Value
Control Treatment	.033 (Men > Women)	1

Treatment	Difference Between Men and Women	P-Value
Social Normative Treatment	0.064 (Men > Women)	0.95
Empirical Information Treatment	0.133 (Men > Women)	0.17

Magnitude of Punishment by Gender per Treatment

Treatment	Difference Between Men and Women	P-Value
Control Treatment	.152 (Men > Women)	1
Social Normative Treatment	0.487 (Men > Women)	0.139
Empirical Information Treatment	0.079 (Men > Women)	0.609

Discussion

First, there are no significant results, thus we cannot reject the null that the means across all treament groups are equal. The lack of results is unlikely to be due to a gender effect. There was a fairly even distribution of men and women across treatment groups, and there was no differnce between how men and women punished.

The lack of results may be caused by the experimental design. In the experiment, there was no "cost" associated with punishing other participants who lied. This caused most participants to punish across all treatments, so there is not enough room for variation. If I were to re-run this experment, I would have a cost of punishing each participant, which may make the norms more effective than no normative information. This is also more likely to be similar to real world situations (increase external validity), since punishing peers is likely to come with either financial or social costs. Additionally, the study is not highly powered (80 participants), so a larger sample size may have led to better results.

In terms of implications, there are no significant results to suggest any policy on encouraging whistleblowers. However, it is clear that in a costless environment people were fairly willing to punish, and there is no need to worry about men or women punishing less. So in environments with costless punishment, this research implies that policy makers do not have to worry about people punishing unfair behavior. There are implications for future research in that punishment and norms is still a very underexplored area. Therefore, it could be interesting to examine how norms impact:

1. a rewards based system for non-deviants with the punishments for deviants
2. a non-anonymous environment in which others could see punishment decisions
3. a financially costly environment for punishers

References

Bicchieri, C., Dimant, E., & Xiao, E. (2017). Deviant or Wrong? The Effects of Norm Information on the Efficacy of Punishment (No. 2017-14).

Bicchieri, Cristina. Norms in the wild: How to diagnose, measure, and change social norms. Oxford University Press, 2016.

Eckel, C. C., & Grossman, P. J. (1996). The relative price of fairness: Gender differences in a punishment game.