

Jocelyn Jingjing Hu^a, Richelle Ju^a, Tiphany Chow^a, Katherine Lim¹

^a*Department of Economics, Pierce Hall, Northampton, MA, 01063*

Abstract

An important principle of economics is that people respond to incentives. For that reason, government and organizations implement incentives to encourage prosocial behavior. This paper looks at how monetary and nonmonetary incentives change people's decisions to act prosocially, and how these effects differ across gender. It will start by discussing the paper about the effect of monetary incentives on charity donations, and will then look at the literature on the drivers of prosocial behavior and the effectiveness of incentives. Lastly, we will tackle unanswered questions by suggesting a more comprehensive experimental design that addresses the gaps in the literature by considering a new population.

JEL Classification: C91, C93, D90

Keywords: Individual Behavior, Laboratory Experiments, Field Experiment, Prosocial behavior

1. Introduction

2. Main paper

3. Reproduction of [Author Names and date of paper]

The following section will discuss the reproduction of all the relevant figures, p-values and regression results to gamble choices and group formation experiments presented in the paper. The discussed summary statistics, figures, and regression tables are within the text while the relevant regression for the mentioned p-values of the logit regression model can be found in the appendix.

The bar chart and all the summary tables in this paper are reproduced successfully and exactly the same. All the regression tables are reproduced successfully but there are some minor differences. For table 5, we are able to reproduce the regression model with the same coefficients and very similar marginal effects. However, R and STATA calculate marginal effects differently and R's margins package is not able to generate standard errors when a factor variable is included in the model. Therefore, the marginal effects we present are very close but not identical. In addition, the linear regression model with clustering we reproduce are a little bit different from the one in the paper since

Email addresses: jhu70@smith.edu (Jocelyn Jingjing Hu), tju@smith.edu (Richelle Ju), tchow@smith.edu (Tiphany Chow), kql15@hampshire.edu (Katherine Lim)

R treats NA differently and drops one more observation than STATA from the data set we are given. Yet the difference is very minimal as most numbers in the regressions are identical and only a few have different second or third digits after decimal point.

Figure 1: Assortative match with respect to risk attitudes by different types o

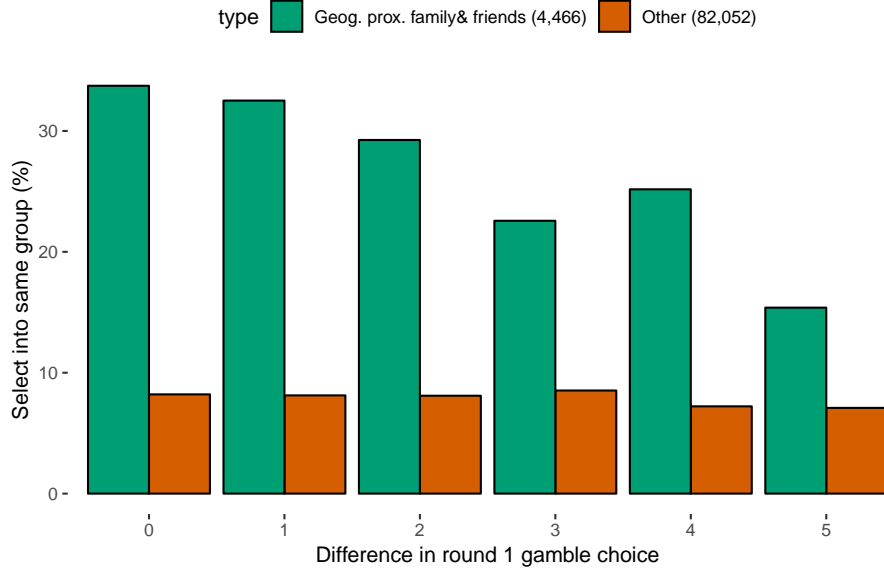


Figure 1 below could be regarded as an informal analysis by visualizing the relationship between grouping choices, risk attitude, and inter-relationship. It shows how the proportion of dyads choosing to group together varies with the difference in their first round gamble choices and depending on whether they are close family and friends or not.

Descriptive statistics for the experimental participants are presented in table 1 and reflect the fact that only households in the poorest of six income categories defined by the Colombian government are eligible for the financial assistance program, such as FeA that makes cash transfers to households conditional on a pledge from them that all of their children will complete primary school and that the senior woman will attend some nutrition workshops(@). The first and second columns contain the proportions, means, and corresponding standard errors for as many of the 2, 512 subjects as we can match to the survey data in the case of each variable, and the third and fourth columns present the same statistics but for the sample upon which the dyadic regression analysis was ultimately performed. Eighty seven percent were female, 77 percent were married, 29 percent were heads of households. This means that the current experimental results could not be generalized to all sexualities and might contain gender bias. Their average age was 42 years and, on average, they had 3.7 years of education. Thirty-four percent lived in municipal centers where the experimental sessions

Table 1: Experimental subjects						
	Full Sample			Sample Analysed		
	Obs.	Mean/Prop	s.d.	Obs.	Mean/Prop	s.d.
Female	2420	87.52%	33.06	2321	87.2%	33.41
Age (years)	2396	41.78	11.39	2321	41.72	11.37
Education (years)	2397	3.7	3.12	2321	3.7	3.13
Household head	2423	28.64%		2321	28.78%	
Married	2420	77.11%		2321	77.73%	
Lives in municipal centre	2478	33.86%	47.33	2321	34.38%	47.51
Household consumption ('000 Pesos/month) 2	2478	433.64	254.9	2321	427.28	249.91
Log household consumption per month	2478	12.82		2321	12.81	
Household size	2452	7.34	3.19	2321	7.27	3.13
No. of kin recognized in session	2506	0.32		2321	0.32	
No. of friends recognized in session	2506	2.39	2.57	2321	2.42	2.57

were conducted, while the remaining 66 percent lived in the surrounding rural clusters. The average monthly household consumption for this sample at the time of the experiments was 430, 000 Colombian Pesos (approximately US\$190).

Table 1: Gamble choices						
Gamble Choice	Low payoff(yellow)	High payoff(blue)	Expected value	Standard Deviation	Risk aversion class	CRRA range
Gamble1	3000	3000	3000	0	Extreme	infinity to 7.49
Gamble2	2700	5700	4200	2121	Severe	7.49 to 1.73
Gamble3	2400	7200	4800	3394	Intermediate	1.73 to 0.81
Gamble4	1800	9000	5400	5091	Moderate	0.81 to 0.46
Gamble5	1000	11000	6000	7071	Slight-neutral	0.47 to 0.00
Gamble6	0	12000	6000	8485	Neutral-negative	0 to -ve infinity

Table 2 presents the expected returns on each gamble and the associated risk appetites. The expected returns vary between 3, 000 and 6, 000 Colombian Pesos.

Table 3 presents the summary statistics of the experimental data. The first and second columns contain the proportions, means, and corresponding standard errors for all 2, 512 participants and the third and fourth columns present the same statistics but for the sample upon which the dyadic regression analysis was ultimately performed. Overall, there is a shift towards more risk-taking in the second round in which 10 percent more people chose Gamble 5 or 6. In the second round, 85% of the subjects chose to join a group of people to share the risk and the average number of subjects in each group is 5. In addition, eight percent of the people defected most of which left after winning their gambles instead of losing their gambles.

The summary statistics of the dyadic variables in the logit regression are shown in table 4. We see that 1) nine percent of all the possible within-municipality dyads grouped together. 2) The average difference in gamble choices, as a proxy of risk attitudes difference, was two, meaning that one

Table3: Experimental data				
Variables	Full Sample		Sample Analysed	
	Mean/Prop	s.d.	Mean/Prop	s.d.
Gamble choice 1st round				
Gamble 1(safe)	8.739%		8.746%	
Gamble 2	17.757%		17.665%	
Gamble 3	18.196%		18.311%	
Gamble 4	29.29%		29.168%	
Gamble 5	11.253%		11.116%	
Gamble 6(riskiest)	14.765%		14.994%	
Won gamble in 1st round	54.7%		54.5%	
Winings 1st round ('000 Pesos)	5.842	3.832	5.835	3.838
Joined a group	86.2%		86.9%	
Number of co-group members	4.195	5.776	3.687	3.899
Gamble choice 2nd round				
Gamble 1(safe)	6.028%		5.991%	
Gamble 2	12.854%		12.759%	
Gamble 3	17.685%		17.759%	
Gamble 4	28.942%		28.75%	
Gamble 5	17.206%		17.328%	
Gamble 6(riskiest)	17.285%		17.414%	
Won gamble in 2nd round	57.7%		57.7%	
Reneged having won gamble	6.3%		6.4%	
Reneged having lost gamble	1.8%		1.8%	
Winnings 2nd round ('000 Pesos)	6.134	4.046	6.133	4.052
Observations	2506		2321	

member of the dyad choosing the modal Gamble 4 and the other choosing either Gamble 2 or Gamble 6. 3) In ten percent of the dyads one or both of the members recognized that they shared a tie of kinship or friendship. Among them, kinships are more rare than friendships by looking at the proportions, but the overall number of recognitions was still large due to the huge sample size. 4) Table 4 also presents the average differences in and sums of individual characteristics for the dyads.

Table 5 presents the marginal effects of five logit regressions. In the paper, two models are mainly analyze the test the hypothesis. The first model is:

$$m_{ij} = \beta_0 + \beta_1 f_{ij} + \beta_2 |l_i^1 - l_j^1| + \beta_3 (f_{ij} * |l_i^1 - l_j^1|) + s_{ij} + u_{ij} \quad (2)$$

The dependent variable is a binary variable indicating whether people join the same group or not. β_1 gives the evidence on whether close family and friends are more likely to group together. β_2 is the difference in gamble choices, which is a proxy of risk attitudes. β_3 , one of the most important predictors, tells us whether close family and friends assort more strongly with respect to risk attitudes than who are unfamiliar. If its marginal effect is significantly negative, the answer would be yes. This model corresponds to the first column in table 5. The result confirms two hypotheses. First, grouping is more likely among

Table4: Experimental data				
	All dyads		Close family and friends	Other dyads
	Mean/Prop	s.d.	Mean/Prop	s.d.
Joined same group in round 2(%)	9.21%		29.467%	8.107%
Difference in gamble choice(round1)	1.639	1.265	1.682	1.637
Sum of gamble choices (round 1)	7.179	2.123	7.056	7.186
Friends and family: One or both recognized friendship or kinship(%)	10.49%		100%	5.618%
Both recognized friendship(%)	2.427%		29.019%	0.98%
Both recognized kinship(%)	0.176%		2.015%	0.076%
One recognized friendship, other kinship(%)	0.453%		5.464%	0.18%
One recognized friendship(%)	6.9%		57.77%	4.132%
One recognized kinship(%)	0.534%		5.732%	0.251%
Stranger(%)	89.51%		0%	94.382%
Geographically proximate friends and family(%)	5.162%		100%	0%
One lives in the municipal centre, one not(%)	30.951%		0%	32.635%
Different genders(%)	20.537%		16.57%	20.753%
Difference in age(years)	12.404	9.682	11.481	12.454
Difference in education (years)	3.235	2.77	2.633	3.267
Difference in marital status(%)	34.684%		30.721%	34.9%
Difference in household consumption ('000s Pesos/month)	232.84	227.244	226.585	233.181
Difference in log household consumption per month	0.589	0.489	0.583	0.59
Difference in household size	3.111	2.907	2.803	3.128
Difference in round 1 winnings ('000 Pesos)	4.182	3.213	4.079	4.188
Number who live in the municipal centre	0.715	0.78	1.161	0.691
Number of females	1.75	0.482	1.765	1.749
Sum of ages (years)	83.673	16.012	84.456	83.631
Sum of education (years)	7.352	4.512	6.71	7.387
Number married	1.55	0.592	1.595	1.548
Sum of household consumption ('000s Pesos/month)	850.188	359.942	850.359	850.179
Sum of log household consumption per month	25.621	0.845	25.601	25.622
Sum of household sizes	14.568	4.529	14.114	14.592
Sum of round 1 winnings ('000s Pesos)	11.708	5.484	11.524	11.718
observations	86518		4466	82052

close family and friends than the unfamiliar. Second, the significant negative marginal effect of the interaction term indicates that among close family and friends grouping is strongly assortative with respect to risk attitudes and among unfamiliar individuals grouping may or may not be assortative with respect to risk attitudes.

A possible explanation is that being close friends and family is associated with a higher likelihood of grouping together. Because family and friends are considered more trustworthy than strangers, and people with available connections but group with unfamiliar are suspected of being untrustworthy.

The second model includes the three variables from model 1 and is added more controls regarding the nature of relationships of friends and kinship. It corresponds to second column of table 5.

$$\begin{aligned}
m_{ij} = & \beta_0 + \beta_{11}f_{1ij} + \beta_{21}f_{2ij} + \dots + \beta_{k1}f_{kij} + \beta_2|l_i^1 - l_j^1| + \beta_3(f_{ij} * |l_i^1 - l_j^1|) \\
& + \beta_4|z_i - z_j| + \beta_5(z_i + z_j) + \beta_6(l_i^1 + l_j^1) + s_{ij} + u_{ij}
\end{aligned}$$

The results of the second model support the robustness of the findings of the first model. The additional controls lead to a significant reduction in the

Table 5: Dyadic analysis of assortative matching on risk attitudes

	All dyads		Close friends and family	Other dyads	
	1	2	3	4	5
Difference in gamble choice (round 1)	-0.001	0.001	-0.018	0	-0.001
	-0.004	-0.001	-0.018	0	-0.001
Close friends and family	0.153	0.043			
	-0.049	-0.019			
Diff. in gamble choice 1 x Close friends and family	-0.014	-0.012			
	-0.001	-0.003			
Both recognised friendship		0.169	0.229	0.159	0.16
		-0.034	-0.065	-0.042	-0.003
Both recognised kinship		0.136	0.213	0.105	0.106
		-0.047	-0.095	-0.055	-0.042
One recognised friendship, other kinship		0.134	0.153	0.13	0.131
		-0.066	-0.112	-0.095	-0.056
One recognised friendship		0.069	0.017	0.066	0.067
		-0.015	-0.059	-0.015	-0.095
One recognised kinship		0.068		0.045	0.045
		-0.034		-0.038	-0.015
Max.no of close friends and family options					-0.007
					-0.003
Diff. in gamble choice x Max no. close friends and family options					-0.007
					0.00042
Other control variables included##	no	yes	yes	yes	yes
Municipality dummy variables included	yes	yes	yes	yes	yes
Number of observations	86518	86518	4440 #	82052	82052
Pseudo R-squared	0.131	0.16	0.183	0.139	0.14

^a Notes: Marginal effects reported. Corresponding standard errors (in parentheses) adjusted to account for non-independence within municipalities by clustering; *** - sig. at 1% level; ** - sig. at 5% level; * - sig. at 10% level. # - 4 additional municipalities were dropped from this regression because, in one, all of the dyads joined the same group and, in three, none of the dyads joined the same group; ## - Controls included: One lives in municipal centre, one not, Different genders, Difference in age, Difference in years of schooling, Difference in marital status, Difference in log household consumption, Difference in household size, Difference in round 1 winnings, Sum of gamble choices, Number who live in municipal centre, Number of females, Sum of ages, Sum of years of schooling, Number who are married (not to each other), Sum log household consumption, Sum of household sizes, Sum of round 1 winnings.

size of the marginal effect of ‘Close friends and family’. While still positive and significant, the marginal effect is less than one fifth of its former size.

In the last column, the third hypothesis is tested and according to the statistical significance of ***Max no. of close friends and family options, we could conclude that the more family and friends are available, unfamiliar dyads are less likely to group together.

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University.
E-mail: hlavac at fas.harvard.edu % Date and time: Sun, May 05, 2019 - 2:50:07 PM

Other regressions the researchers conduct are simple linear regressions with clustering. They are used to test the assumptions on people’s decision to defect (i.e., leave the group) that the predictions of the paper are based on. According

Table 2: Table A1: Group-level analysis of defections: Dependent variable = proportion of members that default

	Proportion of members that default		
	<i>felm</i>	<i>OLS</i>	
	Groups of 2 or 3	All groups	
	(1)	(2)	(3)
1:Density of close friends and family network within group	−0.067*** (0.021)	−0.034 (0.029)	−0.155** (0.066)
2:Number of group members	0.030 (0.021)	0.001 (0.004)	−0.004 (0.005)
1x2			0.039** (0.019)
Average gamble choice	0.012 (0.011)	0.012 (0.009)	0.011 (0.009)
Proportion of females	0.050 (0.048)	0.032 (0.037)	0.031 (0.037)
Average age	−0.000 (0.002)	−0.000 (0.001)	−0.000 (0.001)
Proportion living in municipal centre	0.082** (0.036)	0.031 (0.031)	0.021 (0.031)
Average years of education	−0.013*** (0.005)	−0.010** (0.005)	−0.010** (0.005)
Proportion married	−0.047 (0.048)	−0.001 (0.035)	−0.004 (0.035)
Average log household consumption	0.007 (0.027)	−0.019 (0.025)	−0.018 (0.025)
Average household size	0.004 (0.006)	0.004 (0.004)	0.004 (0.004)
Constant	−0.124 (0.340)	0.212 (0.312)	0.218 (0.311)
Municipality dummies	no	no	yes
Observations	251	526	526

Notes:

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Linear regression coefficients reported.

Standard errors (in parentheses)

adjusted to account of non-independence within municipalities by clustering

to the results from the first (for groups of 2 or 3) and third column (all groups regardless of sizes), the higher the density of within-group friend and family network, the lower the defaults(leaving the group). Moreover, the effect declines with group size.

Although the paper and our reproduction give the same conclusion on the main variables, there is a difference on the statistical significances of the gender and the residence variables. Th paper’s results show that proportion of females are important predictor of default rate while ours shows that gender is not significant but the proportion living in the municipal centers is important. To figure out the reason that this difference exists, we use several functions from different packages to conduct the regression, including the `lm.cluster()` function from the `miceadds` package, standard `lm`, and `lm` with clustering option from the `lfe` package. They all give the same results (column 1 in the reproduced column 1 of table A1) different from that of the original paper. This means that the original paper might be wrong or STATA an R have different assumptions about the linear regression model with clustering effects. If our model is correct, the more percenatge of subjects that live in non-rural areas like towns and villages, the higher the rate of default. Since we know from the literature review that only people from poor areas would usually need group formation to deal with financial challenges, it is understandable that people from urban areas would have a higher default rate since they lack the incentive to stay and could be comfortable with taking risk by themselves after a period of playing. It is also possible that people from urban areas are more risk averse and lose averse, meaning that urban subjects might be willing to form a group with trustworthy and risk loving person, but want to leave the group when loosing to decrease the risk of sharing his/her own gains.

4. Related literature

5. Experiment proposal

Appendix

```
library(stargazer)
stargazer(c1,c2,c3,c4,c5,type = "html", column.labels = c("1","2","3","4","5"), column.separ
```

References