TBD*

TBD

Bongju Yoo and Najma Osman

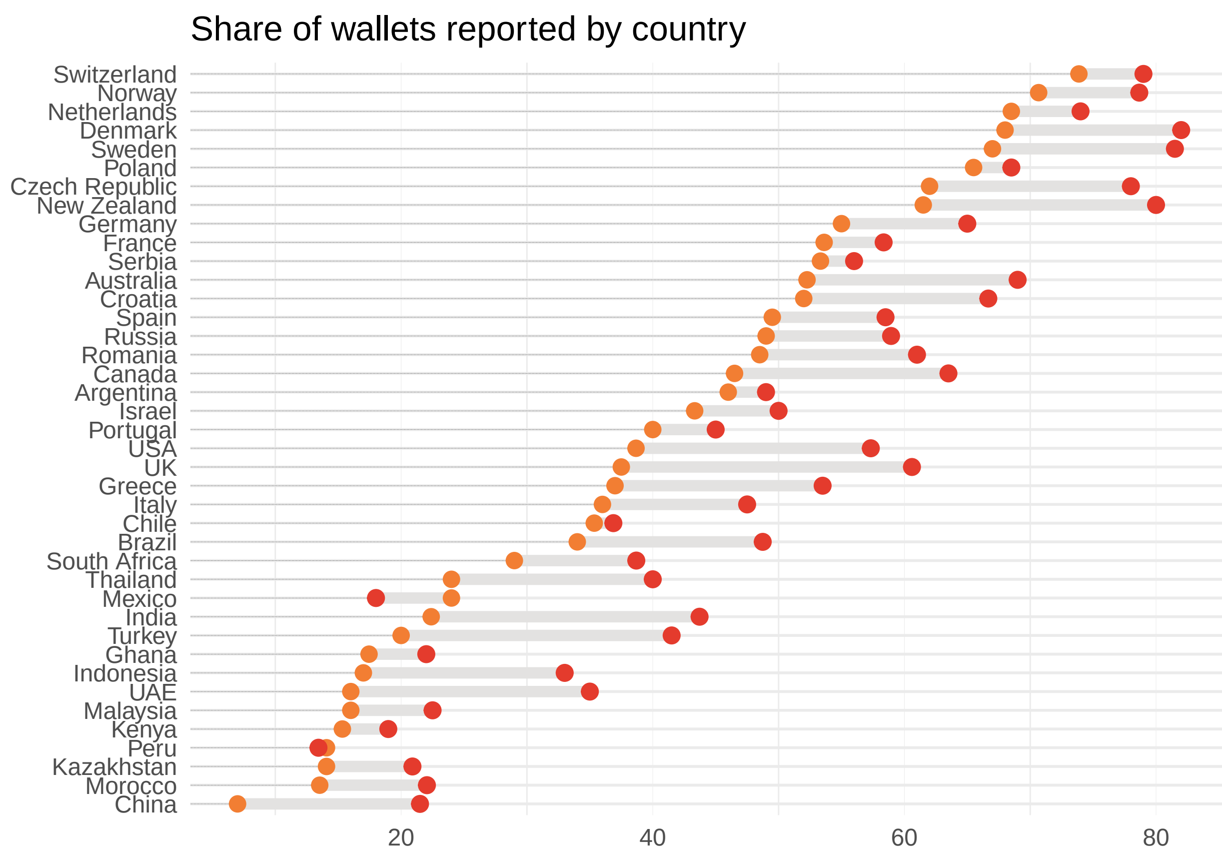
23 March 2021

Abstract

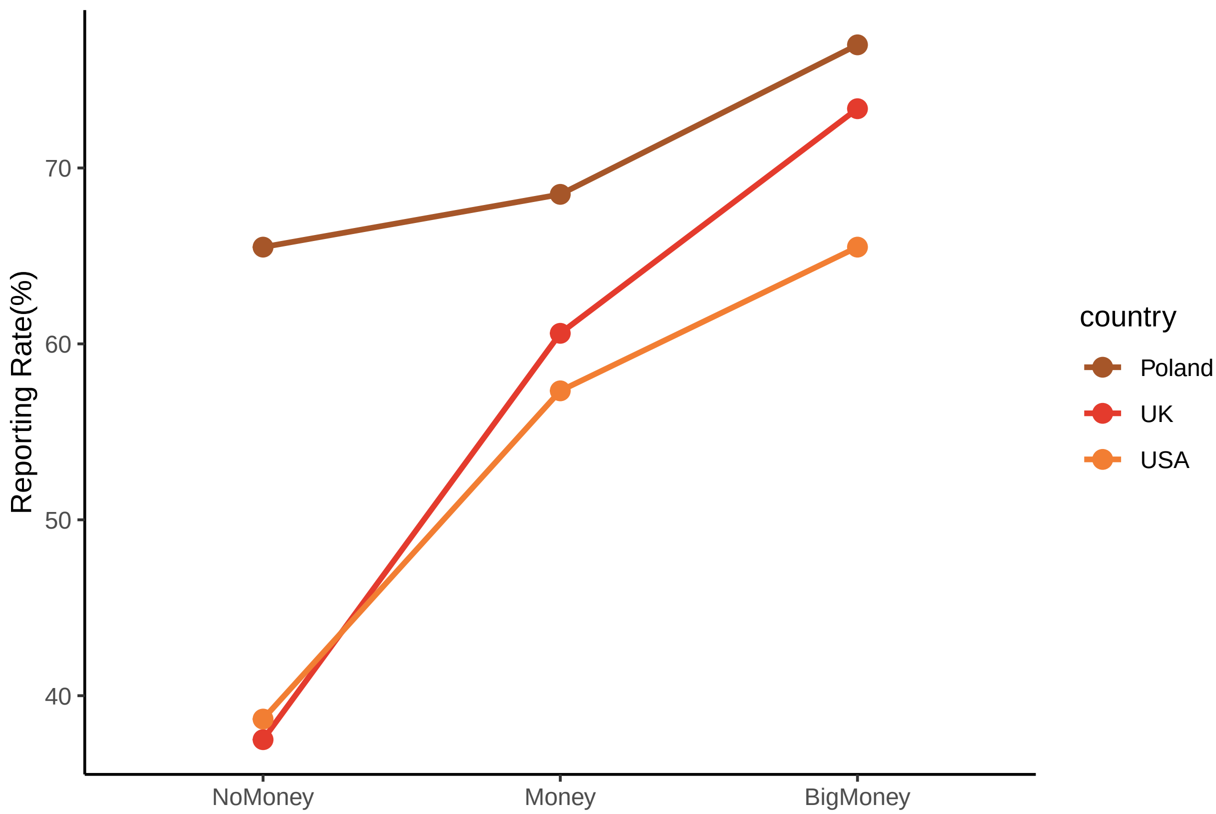
First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

2 Data



*Code and data are available at: https://github.com/bonjwow/lost-wallet



2.1 Description of Study

2.2 Methodology and Data Collection

2.3 Power Analysis

```
##
##    Cell Contents
## |-------------------------|
## |                   Count |
## | Chi-square contribution |
## |             Row Percent |
## |          Column Percent |
## |           Total Percent |
## |-------------------------|
##
## Total Observations in Table:  16099
##
##              | dfPwr$response
##   dfPwr$cond |         0  |       100  | Row Total |
## -------------|-----------|-----------|-----------|
##            0 |      4740  |      3148  |      7888  |
##              |    46.843  |    55.899  |           |
##              |    60.091% |    39.909% |    48.997% |
##              |    54.116% |    42.888% |           |
```

```
##              |   29.443% |   19.554% |           |
## -------------|-----------|-----------|-----------|
##            1 |     4019  |     4192  |     8211  |
##              |   45.000  |   53.700  |           |
##              |   48.947% |   51.053% |   51.003% |
##              |   45.884% |   57.112% |           |
##              |   24.964% |   26.039% |           |
## -------------|-----------|-----------|-----------|
## Column Total |     8759  |     7340  |    16099  |
##              |   54.407% |   45.593% |           |
## -------------|-----------|-----------|-----------|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## ------------------------------------------------------------
## Chi^2 =  201.4426     d.f. =  1     p =  1.011638e-45
##
## Pearson's Chi-squared test with Yates' continuity correction
## ------------------------------------------------------------
## Chi^2 =  200.9936     d.f. =  1     p =  1.267682e-45
##
##
##        Minimum expected frequency: 3596.367
```

```
##
##      Difference of proportion power calculation for binomial distribution (arcsine transformation)
##
##               h = 0.2242923
##               n = 312.0382
##       sig.level = 0.05
##           power = 0.8
##     alternative = two.sided
##
## NOTE: same sample sizes
```

3 Model

We used a linear regression model, specifically ordinary least squares (OLS), which are used to examine relationships between variables, specifically looking to determine which independent variables hold the most and/or significant influence over dependent variables - in this case, whether the wallet is reported or not. In addition to this, we can see how changes in the independent variables are related to changes in the dependent variable through use of dummy variables. It allows us to understand the mean change in a dependent variable, given a 1 unit change in each independent variable.Three regressions were run. Table 1's second column and Table 2 consists of multiple linear regressions, Table 1's first column and Table 3 are simple linear regressions.

3.1 Formulae

Table 1 (column 1) + Table 3:

$$Y \sim \beta_0 + \beta_1 X_1 + \epsilon$$

Table 1 (column 2):

$$Y \sim \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

Table 2:

$$Y \sim \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

3.2 Model Features/Aspects

Like stated prior, 3 OLS regressions are run within this paper to replicate tables S8, S9, and S10 in the original paper, respectively Tables 1, 2, and 3 within this paper.

- Treatment Condition (cond, as found in the cleaned model dataset):
  - Originally, the treatment condition is a categorical variable from 0-3, where 0 = NoMoney, 1 = Money, 2 = Big Money, 3 = Money-NoKey.
  - To replicate S8, the treatment condition had to be re-coded into a dichotomous variable where 0 is the NoMoney condition, and 1 comprises the Money, BigMoney, and Money-NoKey conditions.
  - For S9, dummy variables had to be created in order to focus on the BigMoney and Money conditions,
  - And for S9, the treatment condition was again re-coded so 1 referred to the Money-NoKey condition, and 0 comprised all other conditions.

In doing so, we could isolate for the correct conditions when running the model.

The remaining variables within the model, such as whether the respondent is male, above 40, coworkers, a computer, or other bystanders are present, are all binary variables. Using dummy variables, and/or avoiding categorical ones was necessary to ensure that those demarcated with a value of 0 will cause that particular variable to have no role in influencing the dependent variable, which in this instance is response/reporting of the wallet.

3.3 Justification

For the purposes of this study, a regression analysis makes sense due to the use of a categorical/binary dependent variable, whether the person reports the lost wallet to the owner, and continuous or binary independent variables. Linear regression, as opposed to nonlinear regression models, are a better fit for this study due to their simplicity, assuming adequate fit in the residual plots.

Assumptions underpinning OLS models include:

- No heteroscedasticity: the error term has a constant variance,
- Linearity in Parameters: the regression model is linear in coefficients and error term,
- Zero conditional mean: error term has a population mean of zero
- Variation in X: there is variation in the explanatory variable,
- Normality - No IV is a perfect linear function of other explanatory variables,
- And Random Sampling: the observed data represents a random sample from the population

3.3.1 Alternatives

According to Cohn et al's (2019) supplementary materials, OLS was chosen over non-linear models (e.g., logistic regression) as they return virtually identical results. In addition to this, OLS was used for simplicity and clarity as "coefficients can be directly interpreted as percent point changes." OLS is easier to perform, and interpret. Nonlinear regressions use iterative algorithms as opposed to the linear, solving with matrix equations. This introduces worry about algorithm choice, starting values, and convergence possibilities.

3.4 Model Convergence/Checks

RMSE is the square root of the variance of residuals, and indicates the absolute fit of the model to the data. Specifically, it tells us how close the observed data points are to the model's predicted values, and as such, can be interpreted as the standard deviation of the unexplained variance. Lower values of RMSE indicate better fit, with 0 meaning a perfect fit. Once again, the "response" in this particular model is whether the recipient of the wallet does or does not report it stolen - 0, or 100. The RMSE for each regression are as follows:

- Table 1
  - Col 1: RMSE = 49.76789
  - Col 2: RMSE = 49.01223
- Table 2
  - Col 1: RMSE = 48.53371
  - Col 2: RMSE = 48.30933
  - Col 3: RMSE = 46.21193
  - Col 4: RMSE = 49.22719
- Table 3
  - Col 1: RMSE = 49.13414
  - Col 2: RMSE = 49.43042
  - Col 3: RMSE = 46.35512
  - Col 4: RMSE = 49.9143

The high RMSE here may partially be as a result of failure to account for fixed effects in the models, such as institution and city. That withstanding, since the response data ranges from 0 to 100, the RMSE's result uses the same unit, the numbers here are incredibly large and do not suggest a great fit. However, considering doesn't truly range from 0 to 100, and is rather a binary 0 or 100 for non-response or response, this may have affected the result.

3.5 Software Used

To run the OLS model, we used the R Core Team (2020) inbuilt function, lm(), which is used to fit linear models. It has two key arguments: formula, and data. Formula takes on the form y ~ x1 + x2, where y is the dependent variable and x1, x2, and onward are the independent variable, while data is the data frame containing the columns specified in the formula. The stargazer package by Hlavac (2018) was used to create the LaTeX regression tables.

```r
### S10/Table 3 (main result)

s10_main <- dfBehavModel %>%
  filter(country == 39 | country == 27 | country == 40)
```

```r
s10_main$cond <- ifelse(s10_main$cond == 3, 1, 0)


s10_main_result <- lm(response ~ cond, data = s9_main)
  # UK
s10_uk <- s9_main %>%
  filter(country == 39)
s10_uk_result <- lm(response ~ cond, data = s10_uk)

  # Poland
s10_poland <- s10_main %>%
  filter(country == 27)

s10_poland_result <- lm(response ~ cond, data = s10_poland)

  # USA
s10_usa <- s10_main %>%
  filter(country == 40)
s10_usa_result <- lm(response ~ cond, data = s10_usa)

# RMSE
s10_rmse <- c(sqrt(mean(s10_main_result$residuals^2)),
              sqrt(mean(s10_uk_result$residuals^2)),
              sqrt(mean(s10_poland_result$residuals^2)),
              sqrt(mean(s10_usa_result$residuals^2)))
```

4 Results

Table 1's results are aggregated across the 40 countries visited to conduct the experiment. Results show the following: * Lost wallet reporting rates increase by 12 percentage points in the Money relative to the NoMoney condition * Men are less likely to report a wallet than woman the presence of a computer increased likelihood of reporting the lost wallet, * but the presence of other bystanders decreased reporting rates, * However, unlike in the original paper, age groups (above and under 40), and the presence of coworkers had no statistical significance regarding reporting or failing to report lost wallets.

Table 2 examines reporting rates for the Money and BigMoney conditions in the UK, Poland, and the USA. * Column 1 shows that across the three countries, they were more likely to report a lost wallet containing greater amounts of money. * Columns 2-4, where the countries are isolated, shows a trend that wallets with larger amounts of money are more likely to be reported. * With smaller amounts of money, only the UK and US showed statistically significant likelihood of reporting the lost wallets.

Table 3 looks at reporting rates for the Money-NoKey conditions in the UK, Poland, and the USA. * Column 1 shows that fewer wallets were reported when they did not contain a key in the 3 countries. * Columns 2-4 show that this pattern holds for the UK, Poland, and the USA, but the difference was only statistically significant for the UK and Poland.

Here, we see that it is less likely for Money-NoKey wallets to be reported compared to Money condition wallets in Table 2. This may suggest that having a key present indicates greater loss than the small (USD$13.45) sum of money, as losing the key may cause the owner greater inconvenience.

Table 1: Reporting rates in the Money and No Money Condition

	Dependent variable:	
	Response	
	(1)	(2)
Money	12.157***	10.527***
	(1.540)	(1.527)
Male		−6.910***
		(1.041)
Above 40		−1.449
		(1.035)
Computer		17.210***
		(1.275)
Coworkers		1.011
		(1.075)
Other Bystanders		−6.121***
		(1.085)
Constant	51.053***	45.185***
	(0.549)	(1.544)
Observations	9,407	9,407
R^2	0.007	0.037
Adjusted R^2	0.006	0.036
Residual Std. Error	49.773 (df = 9405)	49.030 (df = 9400)
F Statistic	62.281*** (df = 1; 9405)	59.378*** (df = 6; 9400)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Reporting rates in NoMoney, Money, and Big Money condition

	Dependent variable:			
	Response			
	UK, Poland, and US	United Kingdom	Poland	United States
	(1)	(2)	(3)	(4)
Money	11.317^{***}	17.600^{***}	5.779	12.333^{***}
	(2.025)	(3.200)	(4.020)	(3.600)
Big Money	22.225^{***}	30.367^{***}	14.600^{***}	20.500^{***}
	(2.403)	(4.196)	(4.033)	(4.125)
Constant	49.923^{***}	43.000^{***}	63.065^{***}	45.000^{***}
	(1.348)	(2.419)	(2.321)	(2.205)
Observations	2,926	1,132	794	1,000
R^2	0.030	0.050	0.016	0.028
Adjusted R^2	0.029	0.048	0.014	0.026
Residual Std. Error	48.559 (df = 2923)	48.373 (df = 1129)	46.299 (df = 791)	49.301 (df = 997)
F Statistic	45.429^{***} (df = 2; 2923)	29.585^{***} (df = 2; 1129)	6.582^{***} (df = 2; 791)	14.162^{***} (df = 2; 997)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3: Reporting rates in Money-No Key condition

	Dependent variable:			
	Response			
	UK, Poland, and US	United Kingdom	Poland	United States
	(1)	(2)	(3)	(4)
Money-NoKey	3.605^{***}	3.691^{**}	-10.875^{***}	2.125
	(0.858)	(1.521)	(3.794)	(3.950)
Constant	53.484^{***}	51.633^{***}	70.875^{***}	52.375^{***}
	(1.489)	(2.528)	(1.904)	(1.767)
Observations	2,926	1,132	794	1,000
R^2	0.006	0.005	0.010	0.0003
Adjusted R^2	0.006	0.004	0.009	-0.001
Residual Std. Error	49.151 (df = 2924)	49.474 (df = 1130)	46.414 (df = 792)	49.964 (df = 998)
F Statistic	17.655^{***} (df = 1; 2924)	5.891^{**} (df = 1; 1130)	8.215^{***} (df = 1; 792)	0.289 (df = 1; 998)
Note:				*p<0.1; **p<0.05; ***p<0.01

5 Discussion

This paper seeks to replicate the results found in the study on civic honesty around the globe. Using their datasets, and limiting to results pertaining to reporting rates across the 4 treatment conditions (NoMoney, Money, BigMoney, and Money-NoKey), we attempt to validate their findings on civic honesty. Like Cohn et al. (2019), from the results, we found that citizens appear to display greater civic honesty when wallets contain money, the likelihood of contacting the wallet owner increasing when the wallet has larger amounts of money (Table 1, Table 2. This result holds when controlling for recipient and situational characteristics, such as gender, presence of coworkers, a computer, or bystanders, and age, as seen in Table 1. However, the effects of a coworker/coworkers being present, as well as being ages 40+ did not produce statistically significant results as found in the original paper. In addition, when the wallet contains money but no key, citizens are less likely to contact the owner compared to instances where there was money and a key (Table 3).

5.1 Limitations

There are several limitations associated with the study that must be addressed. The first relates to the survey data that was gathered. Cross-country surveys come with a number of limitations, e.g., comparisons of survey data may be biased due to cultural differences in interpretations of questions, how participants make use of response scales, and to what degree responses are influenced by social desirability concerns (Tannenbaum et al., (2020). It is challenging to draw clean comparisons through survey data alone with these limitations, not including the question of whether survey responses actually translate to meaningful behaviours despite cognitive biases and social desirability effects (Bertrand and Mullainathan, (2001). The limitations listed here raise the likelihood that surveys regarding social capital have little similarity with objective measures of social capital (Tannenbaum et al., (2020). In addition to survey limitations, it is possible that failure to return the wallet or contact the owner is not a dishonest act, but caused by the policies in place at the institutions the wallets were lost in, e.g., use of a lost and found as mentioned by Sulitzeanu-Kenan et al., (2020), or even fatigue.

5.2 Next Steps

Based on the above limitations, we have some suggestions for how the experiment can be extended and/or improved upon. To start, the researchers need to ensure that all institutions used in the study do not have a policy against contacting owners of lost items would allow for a firmer conclusion in whether non-returns are a result of dishonesty or altruistic behaviour. The current model leaves room for non-responses being codified as dishonest where they may not be. Second, in the initial study, additional representative surveys were given in the UK, Poland, and the US, where respondents were given detailed descriptions of wallet drop-off procedure, and asked how likely it was they'd contact the owner. Across the 3 countries, the average was 100%, substantially higher than the wallet return rates seen in the regression results and our replication. Social desirability in the face of them knowing their answers will be seen by someone differs greatly from receiving and/or finding a wallet and the task being left up to you with no indication of who the owner may be.

5.3 Ethics/Bias

40 countries, and 355 cities were used to represent the globe, with 400 observations per country. However, is it fair to say that the 40 countries chosen, or even those 5-8 largest cities per country that were selected, are representative of the world? If we look to the selection criteria for the countries/cities, as seen in the methodology section, countries without 5-8 cities with a minimum population of 100,000, countries that have strict customs, immigration, and banking regulations, and those deemed difficult to visit and/or unsane are omitted. As such, entire regions of the world are omitted, such as north and central Africa (not including

Middle Eastern nations), and the Pacific Islands, while nearly half of all nations included are within Europe, introducing a heavy Western culture bias while attributing results to the world. There is also failure to mention the average socioeconomic status of the areas in which the wallets are dropped off, which can be a confounding variable.

In addition to this, drop-off procedure for the wallets involved research assistants waking into a building, and approaching an employee at the counter to give them the counter and immediately leave the building without leaving contact details or asking for a receipt. Two factors come into mind: is it reasonable to treat the employee as the average person, and did their telling the employee that the wallet was found outside the building and around the corner, as opposed to inside the building, have an effect on the responses? In the event they were told that the wallet was found on the premises, it is possible that reporting rates could have increased due to added likelihood that someone would come searching for the wallet. Furthermore, all research assistants were pooled from two German speaking universities. Assuming they were German or Swiss nationals, it is incredibly likely that their English is accented, and that they are white, which may have had an effect on response rates compared with seeking out assistants within the target countries.

Appendix

References

Bertrand, Marianne, and Sendhil Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data." *American Economic Review* 91 (2): 5. https://doi.org/10.1257/aer.91.2.67.

Hlavac, Marek. 2018. *Stargazer: Well-Formatted Regression and Summary Statistics Tables*. Bratislava, Slovakia: Central European Labour Studies Institute (CELSI). https://CRAN.R-project.org/package=stargazer.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sulitzeanu-Kenan, Raanan, Omer Yair, and Markus Tepe. 2020. "Are Public Sector Employees More Honest? Field Evidence from Forty Countries." *SSRN*, 37.

Tannenbaum, David, Alain Cohn, Christian Lukas Zund, and Michel Andre Marechal. 2020. "What Do Cross-Country Surveys Tell Us About Social Capital?" https://doi.org/10.5167/uzh-188351.