Vote Brokers Replication Paper

Cristopher Patvakanian

Abstract

Frye, Reuter and Szakony (2019) examine the voter behavior Russia and Venezuela and find different types of brokers, appeals, and targets have different effects on voter turnout. I successfully replicated all of their results. As a robustness test, I impute missing values in the dataset and find results in line with that of the original study, but of a smaller magnitude. These results confirm the authors' original findings and suggest that the missing values in their sampled population do not bias the results.

Contents

1 Introduction

I am replicating "Vote Brokers, Clientelist Appeals, and Voter Turnout: Evidence from Russia and Venezuela" by Timothy Frye, Ora John Reuter and David Szakonyi.[1] The authors use survey-based framework experiments conducted in Russia and Venezuela to analyze voter patterns and turnout in both countries respsectively. Their main finding is that there is higher voter turnout after a voter has been induced by an employer as opposed to a party activist. The results are consistent in both Venezuela and Russia, and are also consistent with existing literature on voter turnout, such as Mares and Young (2016), who also find that employer intimidation significantly influences voter turnout. The data are rich and from face to face interviews in both Russia and Venezuela, with the Venezuelan data coming from a stratified sample whereas the Russian from 20 regions. Thus, it appears that the Venezulean data is a more representative than the Russian. However I believe due to the random sampling and oversampling,the Russian data is still an accurate representation of Russian voters. The models used in the paper include using difference in means between brokers and between inducements, predicted probabilities from individual ordered logit regressions on respondents' likelihood of voting in the survey experiment, and fixed effect linear regressions using interaction variables on willingness to vote.

I was able to sucessfully replicate the paper using the softward of RStudio Team (2015). The source of the data and code used for this replication was generously provided online at the authors' publicly accesible Dataverse.[2] The repo with my replication code and data can be found on Github at the link provided in the footnote.[3]

For my extension of this paper, I decided to look at the missing values in this dataset. Given that the treatment assignment was random for all participants, it seemed unlikely that the misssing values the data would likely lead to any to any bias. However, as a robustness check, it is still an important contribution to ensure that the results we see are not due to potential sample bias in either country and that the data truly are missing at random. After imputing data using multivariate imputation by chained equations (MICE), I was able to create mutiple imputations or replacement values, for the gaps in the data. After running similar regressions provided in Table 3, I find results consistent with paper's originally findings, but lower in magitude.

[1] Frye, Reuter, and Szakonyi 2019
[2] Frye, Reuter, and Szakonyi 2019 Dataverse
[3] Author's Github Replication Repository

2 Literature Review

For this paper I looked at a few papers on voting structure in Russia, as I was able to find less on Venezuela. In "How Capitalism was Built", by Anders Aslund, the literature suggests that in many post soviet countries, voting patterns were heavily influenced by the transition to democracy in instiutions built. In Russia, the case was that there was not enough a big push to transform after communism, and thus the country had to face more difficulties in long term in ensuring fair and free elections. Additionally, in Olga Popova's "Corruption, Voting and Employment Status: Evidence from Russian Parliamentary Elections", Popova finds that controlling for different employment statuses and corruption, people are stil likely to vote differently, and more corruption generally induces people to vote more, which I think is to expected. Additionally, the analysis I found closest to this paper was in Mares and Young (2016), who find that employer intimidation significantly influences voter turnout.

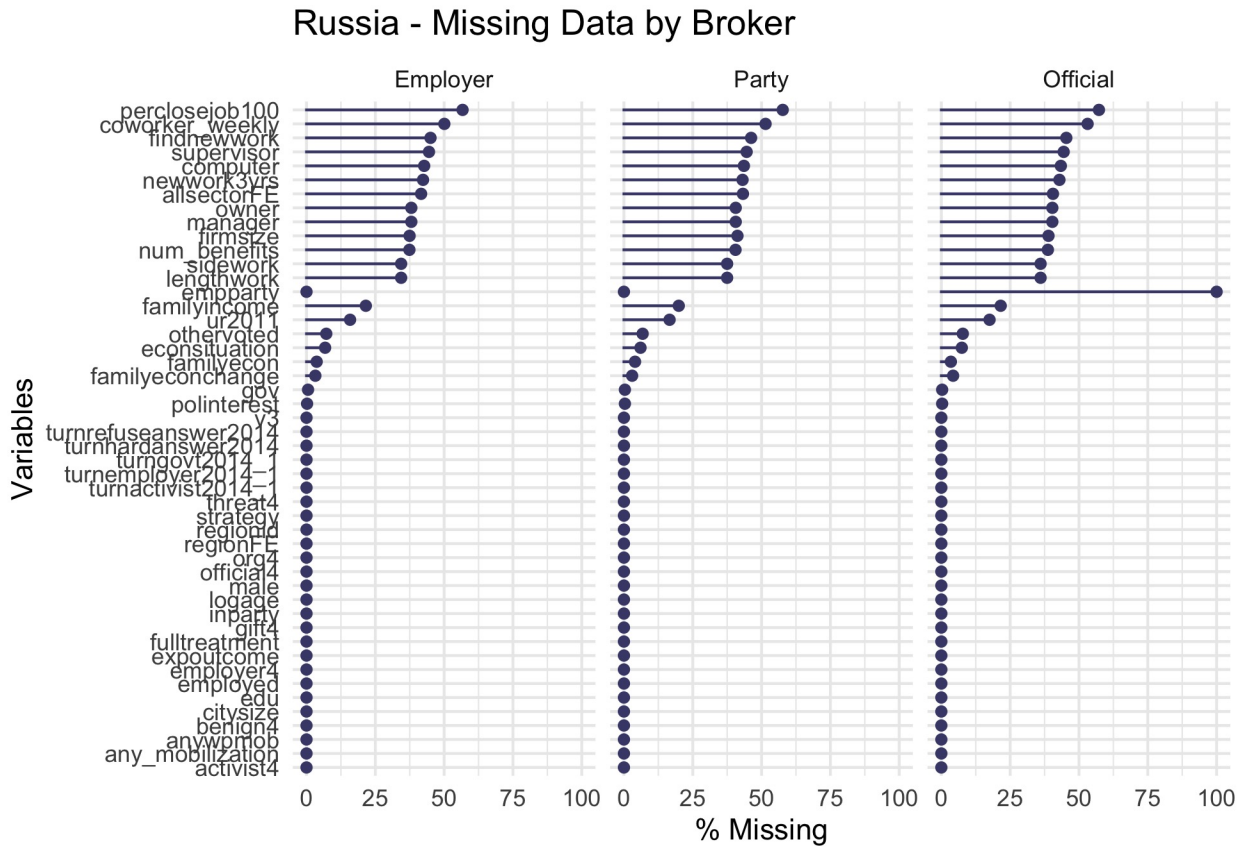
3 Replication

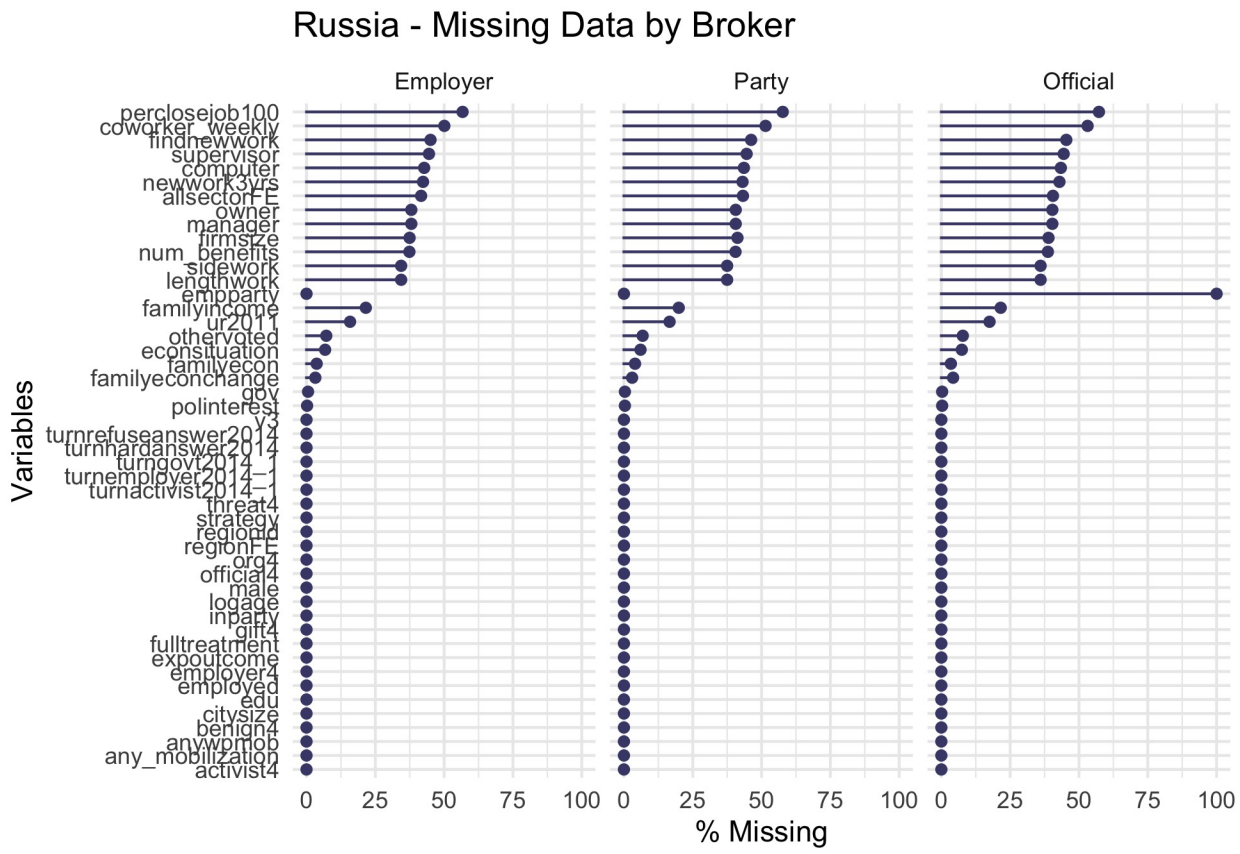
I was succesfully able to replicate all of the figures and the regression table provided in the paper based on the authors' provided data and code.

4 Extension

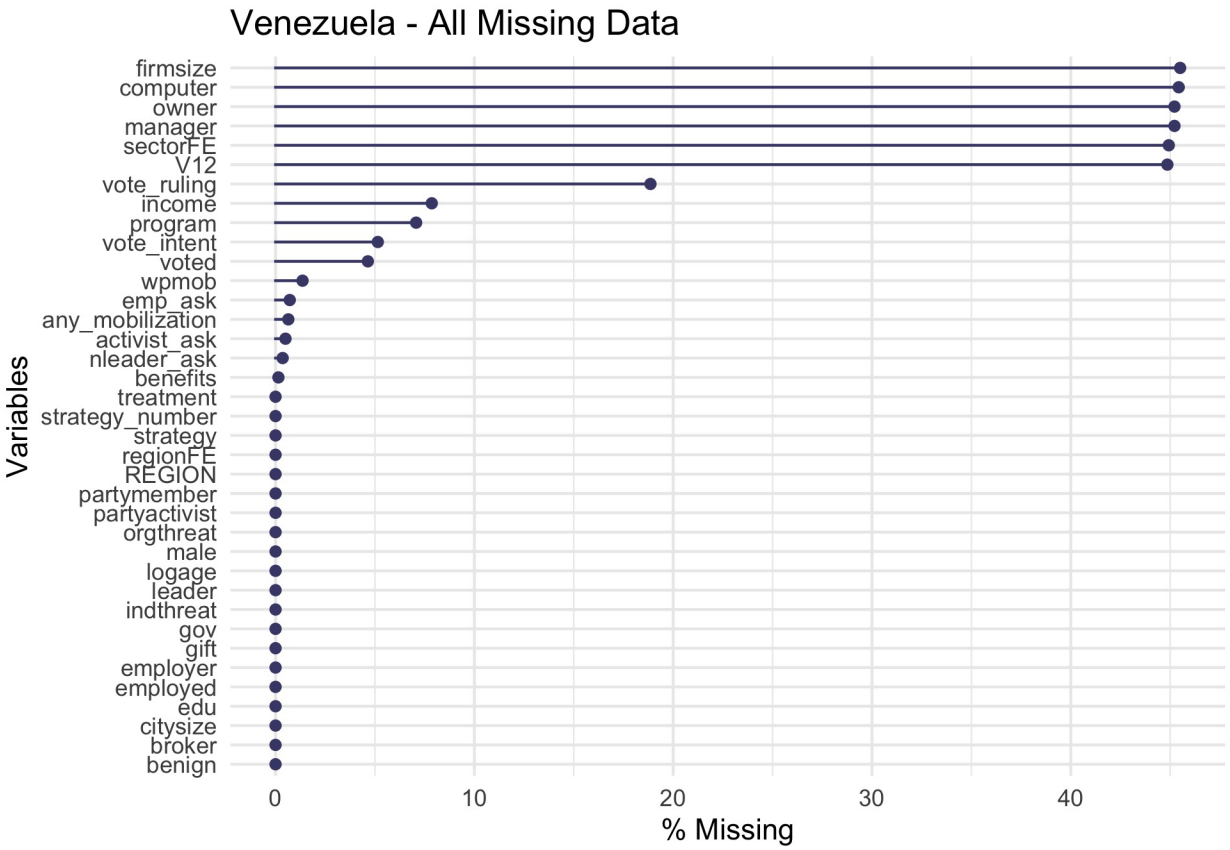
For my extension, I decided to look at missing data in the data sets and impute data for the regressions used in Table 3.

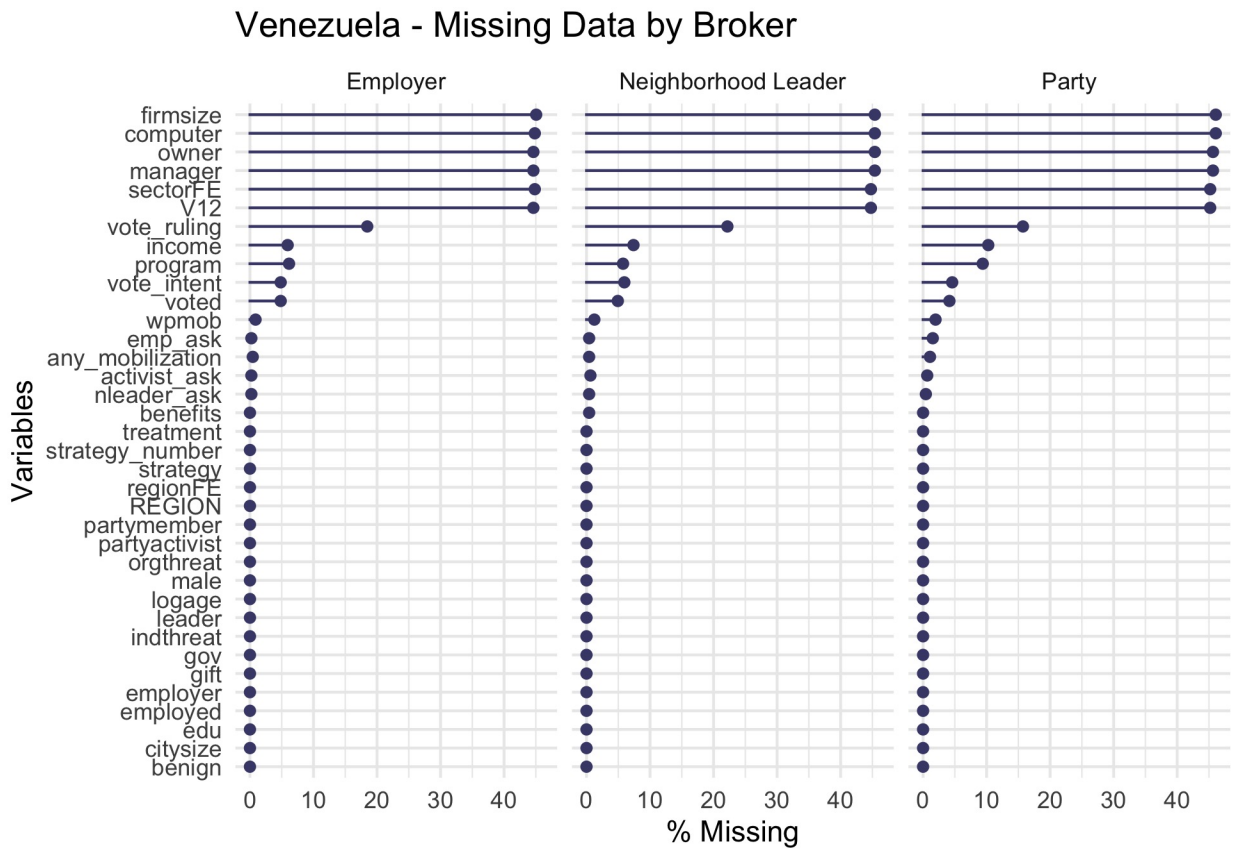
\begin{center}


Russia - Missing Data by Broker
Employer
Party
Official
Variables
perclosejob100
coworker_weekly
findnewwork
supervisor
computer
newwork3yrs
allsectorFE
owner
manager
firmsize
num_benefits
sidework
lengthwork
empparty
familyincome
ur2011
othervoted
econsituation
familyecon
familyeconchange
gov
polinterest
v3
turnrefuseanswer2014
turnhardanswer2014
turngovt2014_1
turnemployer2014_1
turnactivist2014_1
threat4
strategy
regionid
regionFE
org4
official4
male
logage
inparty
gift4
fulltreatment
expoutcome
employer4
employed
edu
citysize
benign4
anyworrijob
any_mobilization
activist4
0
25
50
75
100
% Missing



We see that for Russia, there is a lot of data missing for percentclose job, cowoerker_weekly, and findnewwork variables. When we further example data missing by broker, there generally seems to be the same variables missing with the exception being empparty for official, which is missing for 100% of the data.


Venezuela - All Missing Data
firmsize
computer
owner
manager
sectorFE
V12
vote_ruling
income
program
vote_intent
voted
wpmob
emp_ask
any_mobilization
activist_ask
nleader_ask
benefits
treatment
strategy_number
strategy
regionFE
REGION
partymember
partyactivist
orgthreat
male
logage
leader
indthreat
gov
gift
employer
employed
edu
citysize
broker
benign
Variables
0
10
20
30
40
% Missing



\end{center}

We see that for Venezuela, there is a lot of data missing for firmsize, computer, manager, and sector variables. When we further example data missing by broker, there generally seems to be the same variables missing for all of the brokers.

Given that the regression for Table 3 uses only data from Russia, I decided to impute the missing values using the mice(), function and re-run the regressions, which can be found in Table 2 of the appendix. When I compare my results with the imputed data and compare it to the original study, I find the same significance and sign of the coefficients, however the magnitude of the coefficients appears to be smaller in general. I think that this just shows the original study is valid and robust, and by being able to not only replicate the data but get very similar results to the original after imputing data, I feel even more confident in the authors' findings.

Table 1

	Dependent variable:						
	Outcome Respondent Would Vote						
	Leverage				Monitoring		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
empparty:gov				0.155**			
				(0.065)			
empparty:perclosejob100	0.252**						
	(0.126)						
empparty:findnewwork		0.087***					
		(0.031)					
empparty:num_benefits			0.057*				
			(0.033)				
empparty:supervisor					0.105**		
					(0.051)		
empparty:lengthwork						0.012**	
						(0.006)	
empparty:coworker_weekly							0.042
							(0.059)
Observations	10,324	10,647	10,839	10,921	10,682	10,921	10,504

Note: *p<0.1; **p<0.05; ***p<0.01

The outcome variable is the willingness to turnout outcome (five-point scale) from the survey experiment. The sample includes only respondents who received the employer or political party broker treatment. The employer treatment collapses the data along the inducement treatment arm of the factorial design used in the experiment. The sample is limited to only those who are employed. Chance of job loss measures the probability a respondent believes he or she will lose his or her job in the next twelve months. Hard to find a new job uses a five-point scale to capture the likelihood that if he or she were to lose his or her job, a respondent could find a similar one; higher values indicate more difficulty. Receives benefits captures the number of in-kind benefits (health care, education, transportation sub- sidies, etc.) respondents received from their employer. Higher values on the three-point scale used in knows supervisor well indicate better familiarity with one's boss. Number of years employed measures the length of time at one's work. Socializes with coworkers captures whether respondents spend time with colleagues outside work. All models include the constituent terms and basic demographic characteris- tics (gender, age, education, size of settlement, and an indicator for government employment). Models are estimated via ols and cluster errors at the region level.

5 Conclusion

In this paper, I was succesffuly able to replicate all of the results of Frye, Reuter and Szakony (2019). The the use of a robustness test, I was able impute missing values in the dataset and find results in line with that of the original study, but of a smaller magnitude. These results confirm the authors' original findings and suggest that the missing values in their sampled population do not bias the results. The implications of this replication and extension are that future studies in this area can build on these existing results and have confidence that the findings of Frye, Reuter and Szakony (2019) are in fact robust.

6 References

I make use of RStudio Team (2015), Aslund (2012), Popova (2010), Mares and Young (2016) and Frye, Reuter, and Szakonyi (2019).

Aslund, Anders. 2012. *How Capitalism Was Built: The Transformation of Central and Eastern Europe, Russia, the Caucasus, and Central Asia.* 2nd ed. Cambridge University Press. https://doi.org/10.1017/CBO9781139207850.

Frye, Timothy, Ora John Reuter, and David Szakonyi. 2019. "Vote Brokers, Clientelist Appeals, and Voter Turnout: Evidence from Russia and Venezuela." *World Politics* 71 (4): 710–46. https://doi.org/10.1017/S0043887119000078.

Mares, Isabela, and Lauren Young. 2016. "Buying, Expropriating, and Stealing Votes." *Annual Review of Political Science* 19 (1): 267–88. https://doi.org/10.1146/annurev-polisci-060514-120923.

Popova, Olga. 2010. "Corruption, Voting and Employment Status: Evidence from Russian Parliamentary Elections." *SSRN Electronic Journal.*

RStudio Team. 2015. *RStudio: Integrated Development Environment for R.* Boston, MA: RStudio, Inc. http://www.rstudio.com/.

7 Appendix

7.1 Figure 1

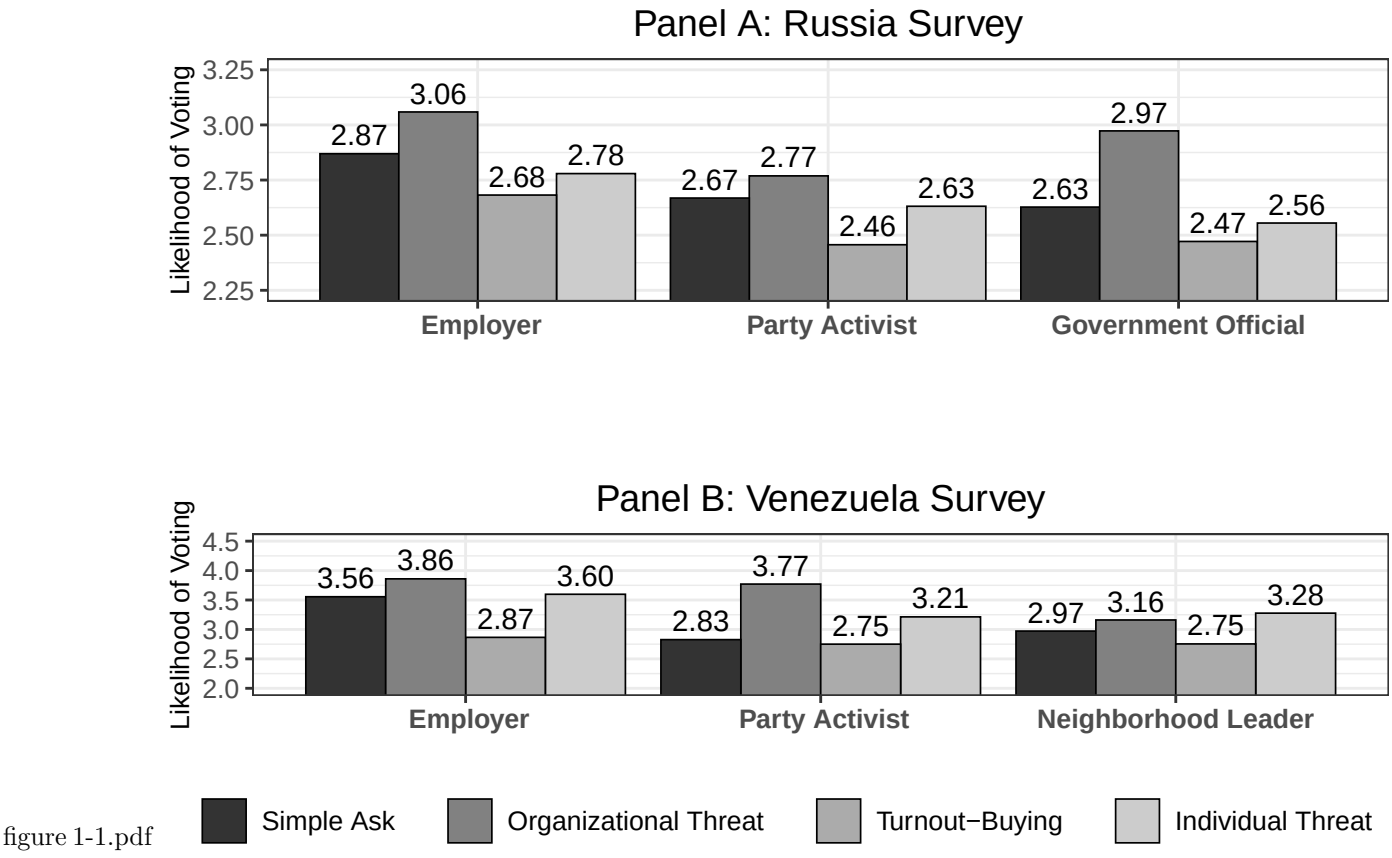


figure 1-1.pdf

7.2 Figure 2

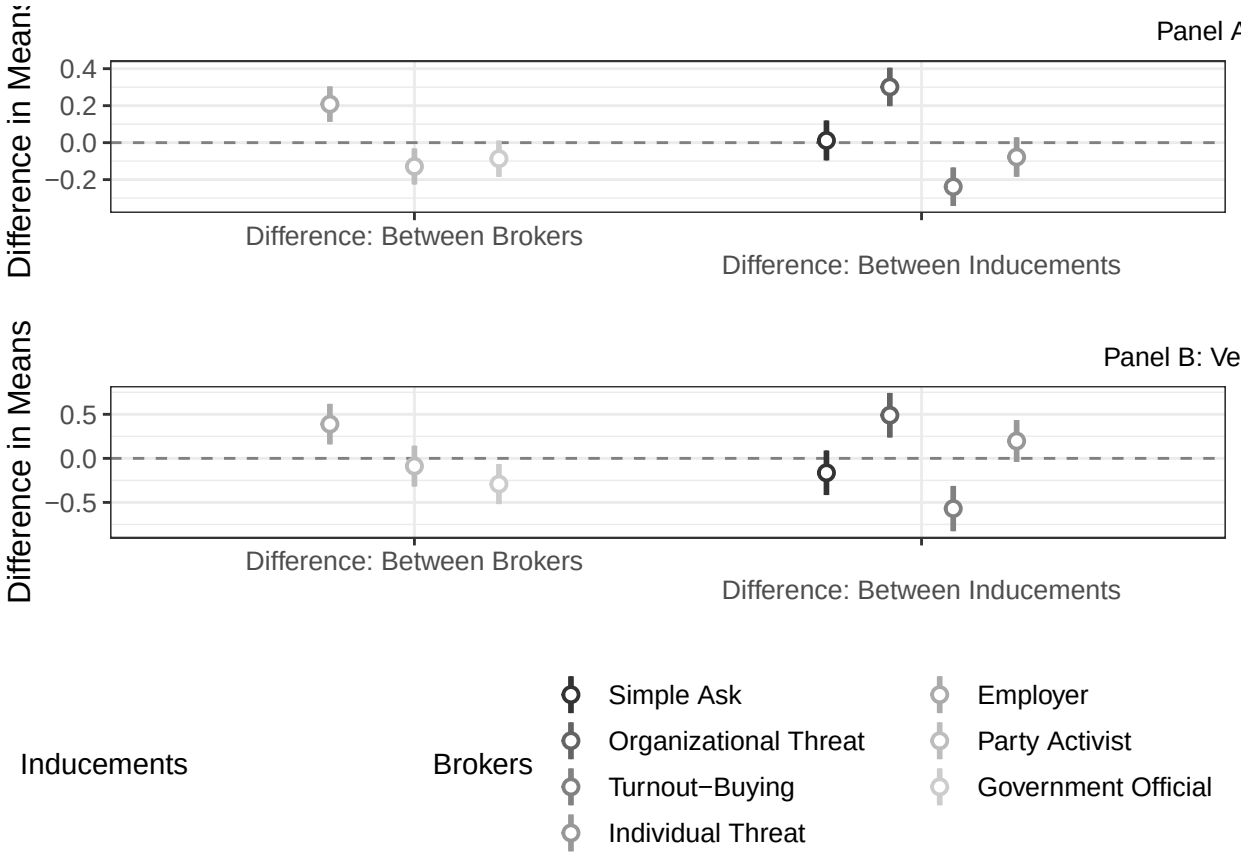


figure 2-1.pdf

7.3 Figure 3

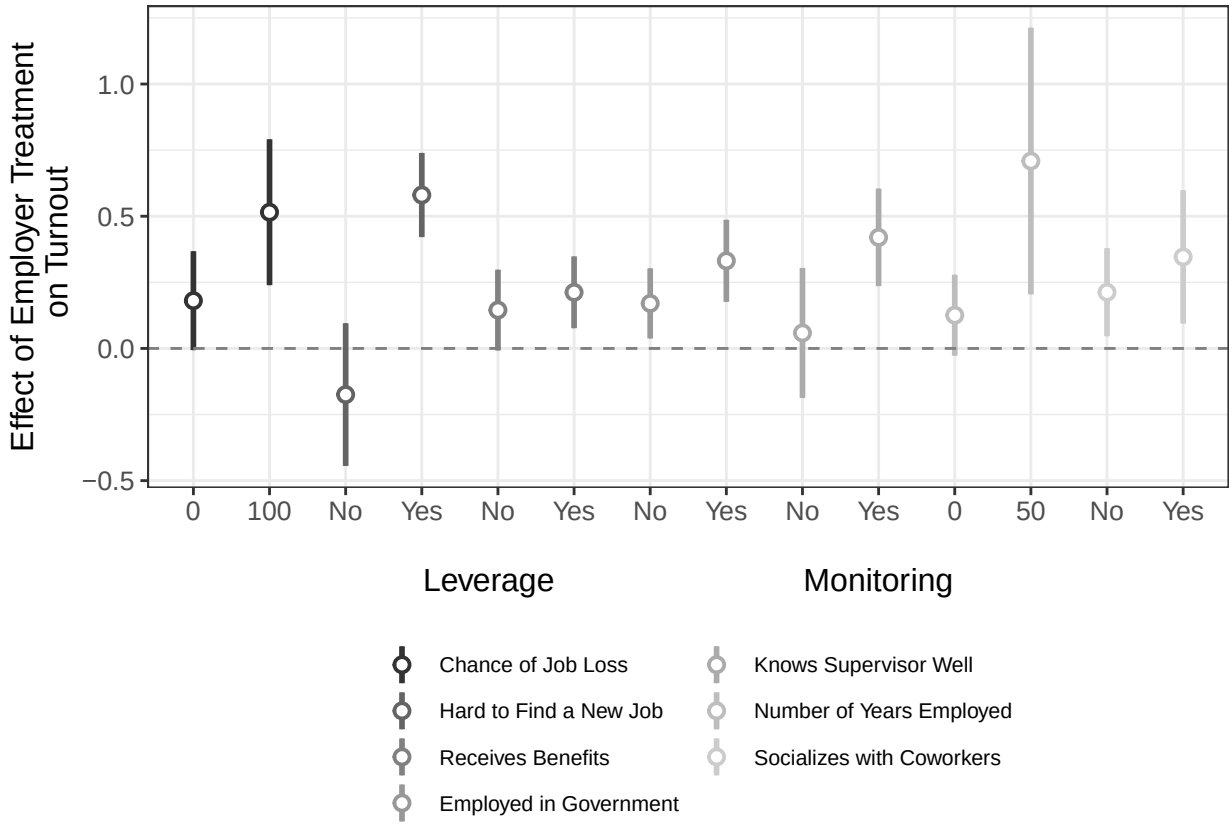


figure 3-1.pdf

7.4 Table 3

Table 2

	Dependent variable:						
	Outcome Respondent Would Vote						
	Leverage				Monitoring		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
empparty:gov				0.166**			
				(0.072)			
empparty:perclosejob100	0.360**						
	(0.167)						
empparty:findnewwork		0.156***					
		(0.038)					
empparty:num_benefits			0.070**				
			(0.036)				
empparty:supervisor					0.119**		
					(0.059)		
empparty:lengthwork						0.012**	
						(0.006)	
empparty:coworker_weekly							0.063
							(0.076)
Observations	1,209	1,532	1,724	1,806	1,567	1,806	1,389

Note: *p<0.1; **p<0.05; ***p<0.01

The outcome variable is the willingness to turnout outcome (five-point scale) from the survey experiment. The sample includes only respondents who received the employer or political party broker treatment. The employer treatment collapses the data along the inducement treatment arm of the factorial design used in the experiment. The sample is limited to only those who are employed. Chance of job loss measures the probability a respondent believes he or she will lose his or her job in the next twelve months. Hard to find a new job uses a five-point scale to capture the likelihood that if he or she were to lose his or her job, a respondent could find a similar one; higher values indicate more difficulty. Receives benefits captures the number of in-kind benefits (health care, education, transportation sub- sidies, etc.) respondents received from their employer. Higher values on the three-point scale used in knows supervisor well indicate better familiarity with one's boss. Number of years employed measures the length of time at one's work. Socializes with coworkers captures whether respondents spend time with colleagues outside work. All models include the constituent terms and basic demographic characteris- tics (gender, age, education, size of settlement, and an indicator for government employment). Models are estimated via ols and cluster errors at the region level.

TABLE 3
EXAMINING MECHANISMS: RUSSIA SURVEY EXPERIMENT [a]

	Outcome: Respondent Would Vote						
	Leverage				*Monitoring*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Employer treatment *	0.360**						
Chance of job loss	(0.167)						
Employer treatment *		0.156***					
Hard to find a new job		(0.038)					
Employer treatment *			0.070**				
Receives benefits			(0.036)				
Employer treatment *				0.166**			
Employed in government				(0.072)			
Employer treatment *					0.119**		
Knows supervisor well					(0.059)		
Employer treatment *						0.012**	
Number of years employed						(0.006)	
Employer treatment *							0.063
Socializes with coworkers							(0.076)
Constituent terms	yes	yes	yes	yes	yes	yes	yes
Demographics	yes	yes	yes	yes	yes	yes	yes
Observations	1209	1532	1724	1806	1567	1806	1389

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$

[a]The outcome variable is the willingness to turnout outcome (five-point scale) from the survey experiment. The sample includes only respondents who received the employer or political party broker treatment. The employer treatment collapses the data along the inducement treatment arm of the factorial design used in the experiment. The sample is limited to only those who are employed. *Chance of job loss* measures the probability a respondent believes he or she will lose his or her job in the next twelve months. *Hard to find a new job* uses a five-point scale to capture the likelihood that if he or she were to lose his or her job, a respondent could find a similar one; higher values indicate more difficulty. *Receives benefits* captures the number of in-kind benefits (health care, education, transportation subsidies, etc.) respondents received from their employer. Higher values on the three-point scale used in *knows supervisor well* indicate better familiarity with one's boss. *Number of years employed* measures the length of time at one's work. *Socializes with coworkers* captures whether respondents spend time with colleagues outside work. All models include the constituent terms and basic demographic characteristics (gender, age, education, size of settlement, and an indicator for government employment). Models are estimated via OLS and cluster errors at the region level.

Figure 1: Original Table 3 in Paper I am replicating