Breaking bad:
Malfunctioning institutions erode good behavior

Agne Kajackaite

joint paper with Rustam Hakimov

September 6, 2022
WZB

Introduction

- In social and economic interactions, individuals often exploit informational asymmetries and behave unethically to pursue private ends

- Given the high costs of dishonesty for the society, understanding the motives underlying such behavior is important

Introduction

- In social and economic interactions, individuals often exploit informational asymmetries and behave unethically to pursue private ends

- Given the high costs of dishonesty for the society, understanding the motives underlying such behavior is important

- Extensive literature on lying behavior:
  - e.g., Gneezy, 2005; Mazar, Amir and Ariely, 2008; Shalvi, Dana, Handgraaf and De Dreu, 2011; Fischbacher and Föllmi-Heusi, 2013; Abeler, Becker and Falk, 2014; Cohn, Fehr and Marechal, 2014; Dai, Galeotti and Villeval, 2018; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo, and Raymond, 2019

Introduction

- In social and economic interactions, individuals often exploit informational asymmetries and behave unethically to pursue private ends
- Given the high costs of dishonesty for the society, understanding the motives underlying such behavior is important
- Extensive literature on lying behavior:
  - e.g., Gneezy, 2005; Mazar, Amir and Ariely, 2008; Shalvi, Dana, Handgraaf and De Dreu, 2011; Fischbacher and Föllmi-Heusi, 2013; Abeler, Becker and Falk, 2014; Cohn, Fehr and Marechal, 2014; Dai, Galeotti and Villeval, 2018; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo, and Raymond, 2019
- Focus on: psychological cost of lying, incentives, demographics, etc.

Introduction

- In social and economic interactions, individuals often exploit informational asymmetries and behave unethically to pursue private ends
- Given the high costs of dishonesty for the society, understanding the motives underlying such behavior is important
- Extensive literature on lying behavior:
  - e.g., Gneezy, 2005; Mazar, Amir and Ariely, 2008; Shalvi, Dana, Handgraaf and De Dreu, 2011; Fischbacher and Föllmi-Heusi, 2013; Abeler, Becker and Falk, 2014; Cohn, Fehr and Marechal, 2014; Dai, Galeotti and Villeval, 2018; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo, and Raymond, 2019
- Focus on: psychological cost of lying, incentives, demographics, etc.
- This paper: How do control institutions affect lying?

Introduction

- In social and economic interactions, individuals often exploit informational asymmetries and behave unethically to pursue private ends
- Given the high costs of dishonesty for the society, understanding the motives underlying such behavior is important
- Extensive literature on lying behavior:
  - e.g., Gneezy, 2005; Mazar, Amir and Ariely, 2008; Shalvi, Dana, Handgraaf and De Dreu, 2011; Fischbacher and Föllmi-Heusi, 2013; Abeler, Becker and Falk, 2014; Cohn, Fehr and Marechal, 2014; Dai, Galeotti and Villeval, 2018; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo, and Raymond, 2019
- Focus on: psychological cost of lying, incentives, demographics, etc.
- This paper: How do control institutions affect lying?
- In particular: How do uncertain control institutions affect lying?

Uncertain control institutions

- In many situations, strong institutions are announced but not implemented
- Despite the presence de jure of robust and modern institutions in many cases, they malfunction de facto
- Examples:
  - Quarantine rules
  - Checks of tax reports
  - Drug use in Indonesia
- In all these examples, even though institutions are used as a threat, there is an uncertainty whether the institution will be enforced
- Such uncertain institutions are the main focus of this paper

Uncertain control institutions

- What all these institutions have in common is that, when introduced, uncertainty exists regarding their implementation
- Over time, with experience, people develop beliefs about the probability that the law will be enforced
- The longer they observe unpunished violations, the greater the likelihood that the institution is malfunctioning
- We introduce this uncertainty in our main treatment

Two streams of literature on control/punishment

- Two streams of literature on control/punishment institutions:
  - Punishment institutions have been found to be useful for maintaining cooperation (e.g., Fehr and Gächter, 2000; Gürerk et al., 2006; Nikiforakis and Normann, 2008).
  - Punishment institutions crowd out intrinsic motivation (e.g., Frey and Oberholzer-Gee, 1997; Gneezy and Rustichini, 2000; Frey and Jegen, 2001; Fehr and Rockenbach, 2003; Dickinson and Villeval, 2008).

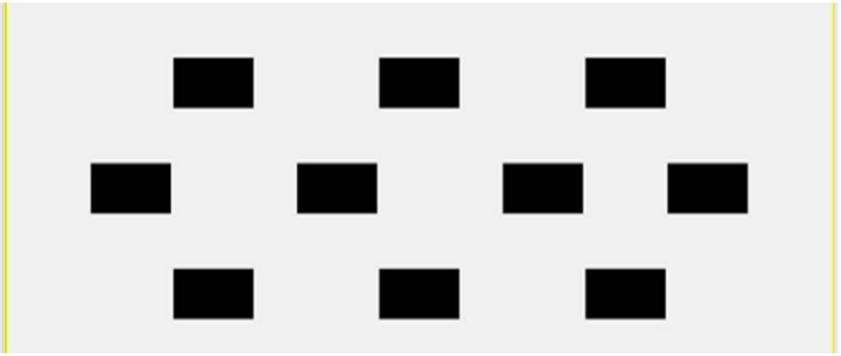
Two streams of literature on control/punishment

- Two streams of literature on control/punishment institutions:
  - Punishment institutions have been found to be useful for maintaining cooperation (e.g., Fehr and Gächter, 2000; Gürerk et al., 2006; Nikiforakis and Normann, 2008).
  - Punishment institutions crowd out intrinsic motivation (e.g., Frey and Oberholzer-Gee, 1997; Gneezy and Rustichini, 2000; Frey and Jegen, 2001; Fehr and Rockenbach, 2003; Dickinson and Villeval, 2008).

- Our paper: how do control institutions affect lying?
- Especially: how do *uncertain* control institutions affect lying?
- Focus on the second stream of the literature

The game

- Repeated version of the observed game (Gneezy et al., 2018)
- 20 rounds
- Three treatments
- Between-subjects design



Treatments

- **No-institution treatment (NoControl):**
  - Payoff equal to the number reported in euros
  - Subjects were told they would not be controlled

- **Strong-institution treatment (Control30).**
  - In each round, control with a 30% probability
  - Independently of their truthfulness, after each round, subjects learned whether they had been controlled in that round
  - If controlled and lied – no payoff; otherwise – payoff equal to the number reported in euros

50-50 treatment

- **50% chance of being in either No-institution or in Strong-institution treatment. (50-50):**
  - Subjects received instructions for both NoControl and Control30
  - Told that with a 50% chance, they were in one of the treatments
  - Independently of their truthfulness, after each round, subjects learned whether they had been controlled in that round

50-50 treatment

- We distinguish between two sub-treatments:
  1. **50-50Control30:**
     - Participants in Control30
     - They had learned it deterministically after the first control

  2. **Malfunctioning-institution sub-treatment (50-50NoControl):**
     - Participants in NoControl.
     - They had never deterministically learned they were in NoControl
     - However, after 10 rounds of absent control, p(NoControl)=97.25%; and after 19 rounds of absent control, p(NoControl)=99.89%

Hypotheses

- Truthfulness in the last 10 rounds:

$$Control30 = 50\text{-}50Control30 > NoControl > 50\text{-}50NoControl$$

Hypotheses

- Truthfulness in the last 10 rounds:

$$Control30 = 50\text{-}50Control30 > NoControl > 50\text{-}50NoControl$$

- Control30 and 50-50Control30 crowd out intrinsic motivation for honesty but at the same time have monitoring function

- 50-50NoControl crowds out intrinsic motivation but has no monitoring function

Procedure

- Laboratory of the University of Valencia (LINEEX); general population

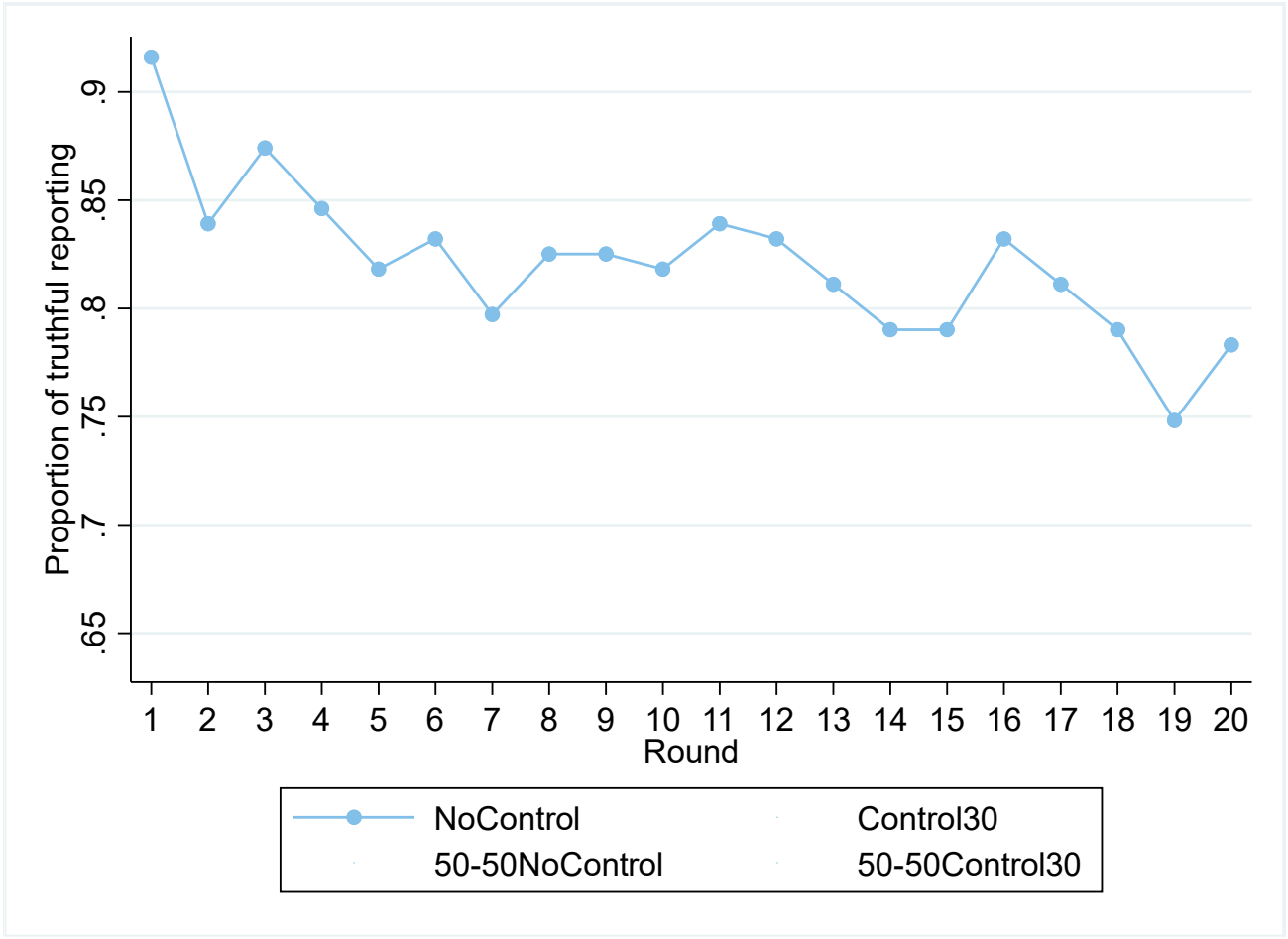
- 615 participants in the main treatments:
  - N(NoControl)=143
  - N(Control30)=145
  - N(50-50NoControl)=168
  - N(50-50Control30)=159

- Average duration of 1 hour; average pay of 11.46 EUR

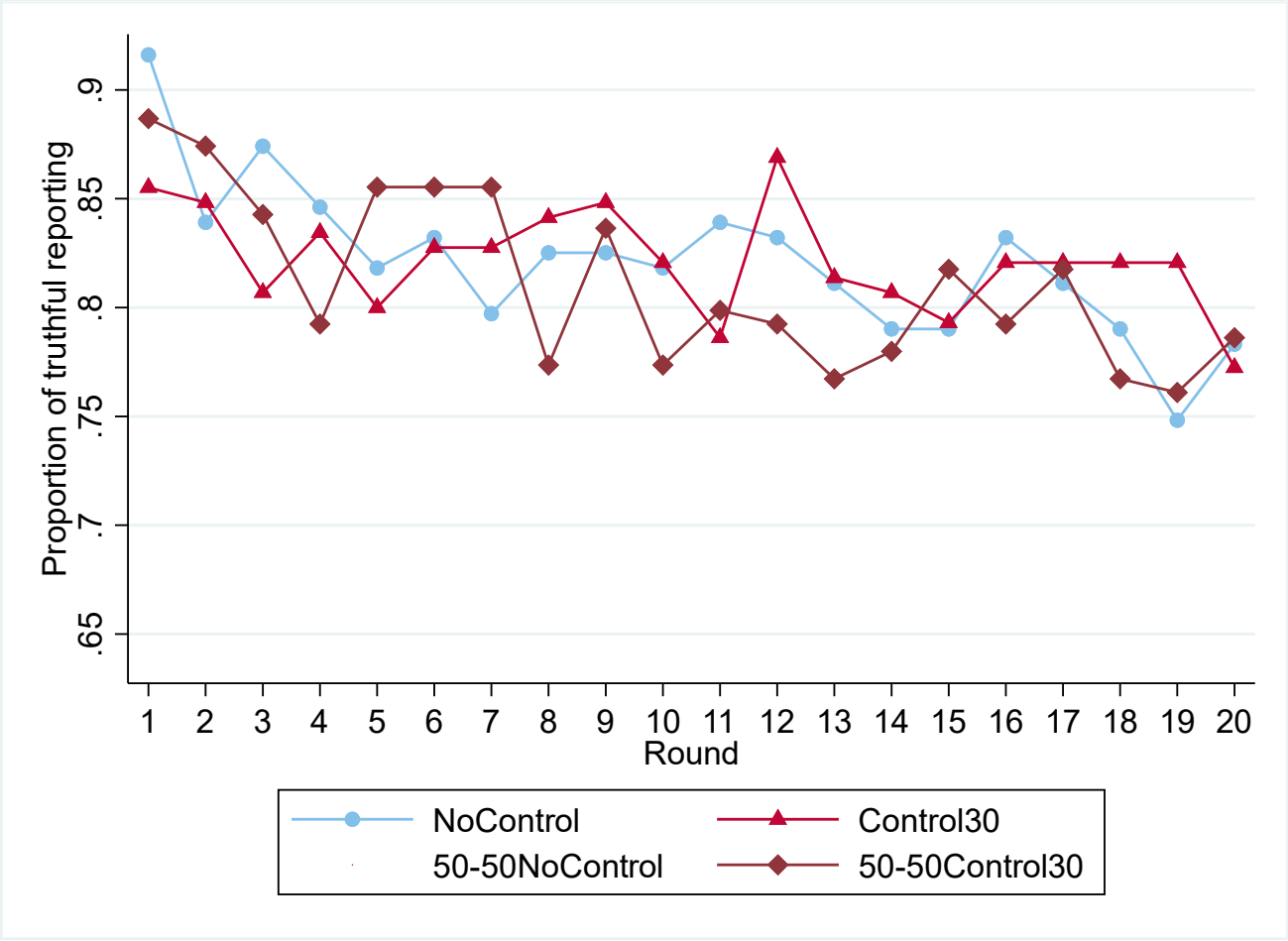
Results:

Overall truthfulness

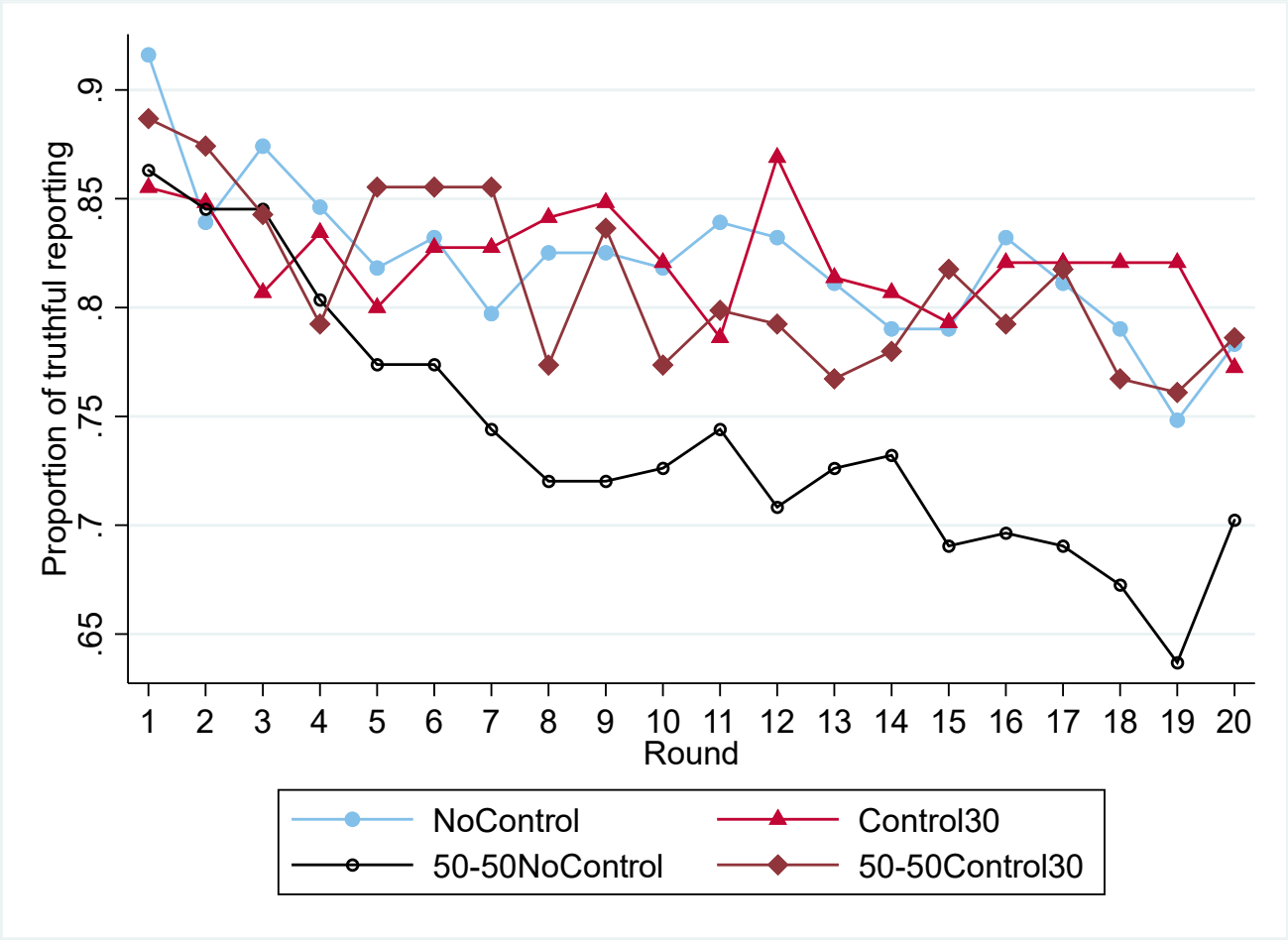
Overall truthfulness



Overall truthfulness



Overall truthfulness



Overall truthfulness

	Truthful (1)	Truthful (2)
Control30	0.010	0.016
	(0.037)	(0.035)
50-50NoControl	-0.097***	-0.111***
	(0.036)	(0.035)
50-50Control30	-0.015	-0.008
	(0.035)	(0.033)
Female		0.073***
		(0.023)
Age		0.010***
		(0.002)
Draw		0.052***
		(0.002)
Round		-0.005***
		(0.001)
Observations	6150	6150
Number of clusters	615	615
Sample	Last 10 rounds	Last 10 rounds
Pseudo *R*-squared	0.039	0.173

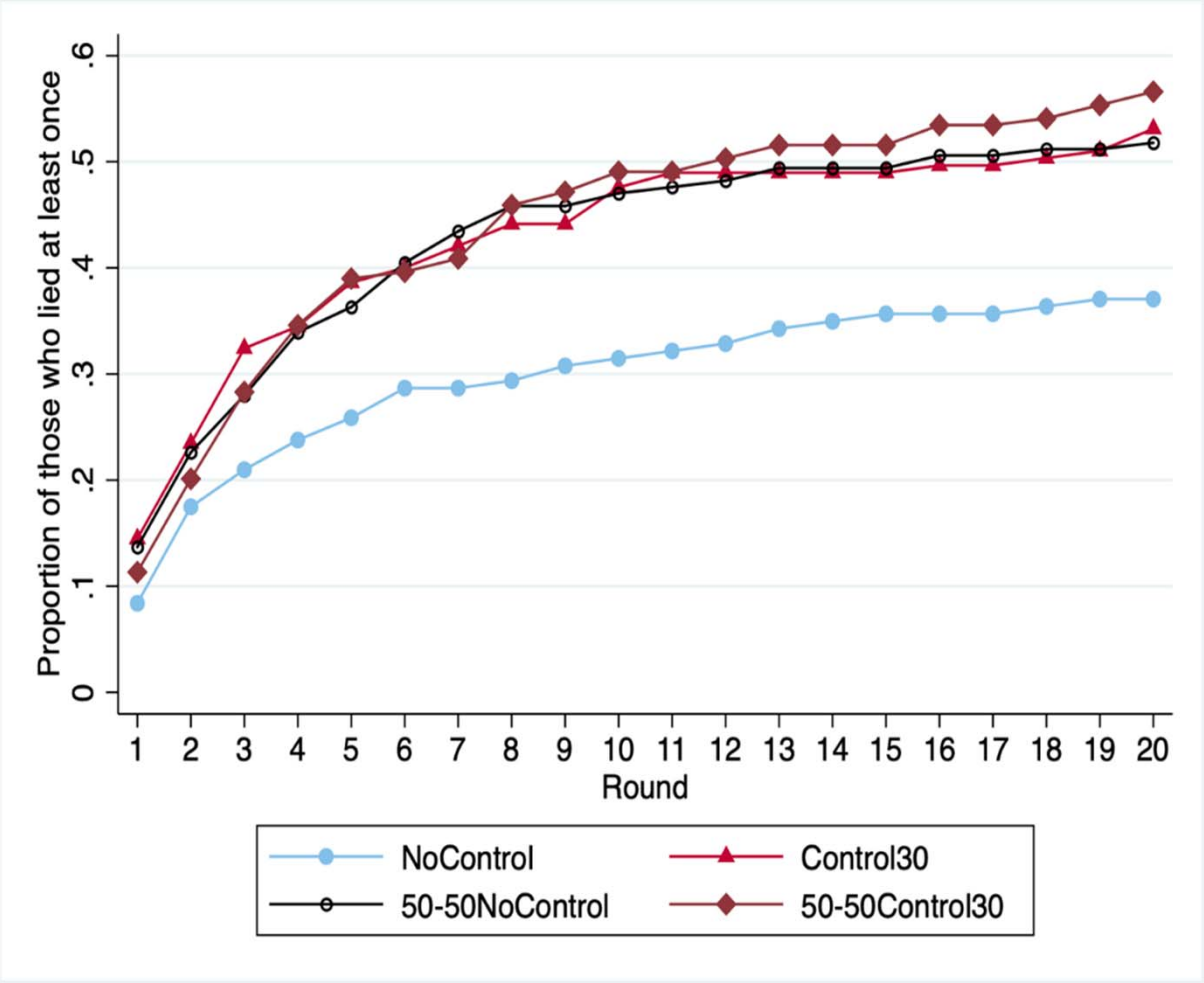
Overall truthfulness

- Participants break bad and cheat the most when the institution is malfunctioning
- This breaking bad starts early – the effect is significant within the first seven rounds
- Contrary to our hypotheses, a strong control institution does not lead to more overall truthfulness than NoControl
- The results on the overall truthful reporting might be masking some underlying shifts in behavior on the extensive and intensive margins

Results:

Extensive margin

Extensive margin: Lying at least once



Extensive margin: Lying at least once

	Lied at least once (1)	Lied at least once (2)
Control30	0.160***	0.154***
	(0.057)	(0.055)
50-50NoControl	0.147***	0.169***
	(0.056)	(0.054)
50-50Control30	0.195***	0.197***
	(0.056)	(0.053)
Female		-0.176***
		(0.039)
Age		-0.019***
		(0.003)
Average draw		-0.044
		(0.028)
Observations	615	615
Sample	All rounds	All rounds
Pseudo *R*-squared	0.016	0.080

Extensive margin

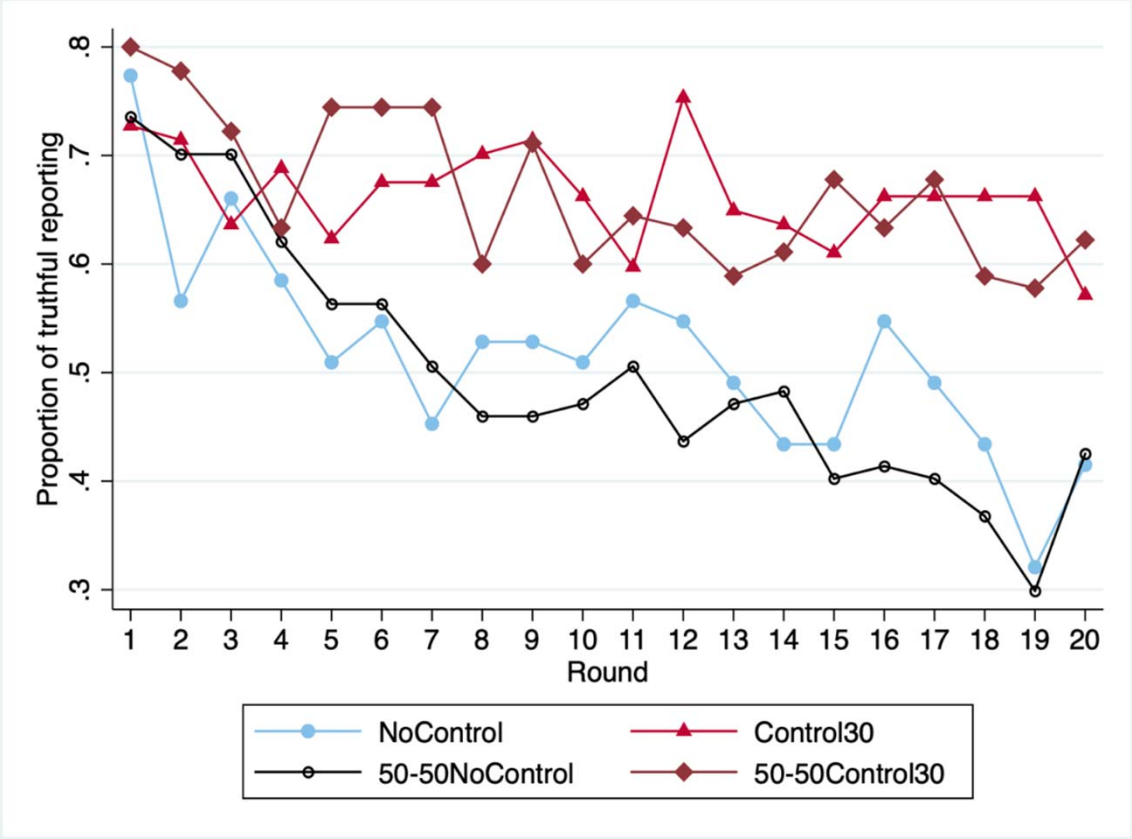
- Participants are more likely to lie at least once when they are in either the strong institution or the 50-50 treatment than when no institution is in place

- The crowding out starts early – the effect is significant within the first four rounds

Results:

Intensive margin

Intensive margin: Proportion of truthful reporting for those who lied at least once



Intensive margin: Proportion of truthful reporting for those who lied at least once

	Truthful (1)	Truthful (2)
Control30	0.175***	0.174***
	(0.049)	(0.048)
50-50NoControl	-0.046	-0.051
	(0.049)	(0.049)
50-50Control30	0.154***	0.156***
	(0.046)	(0.045)
Female		0.015
		(0.029)
Age		0.003
		(0.004)
Draw		0.088***
		(0.002)
Round		-0.008***
		(0.002)
Observations	3070	3070
Number of clusters	307	307
Sample	Lied at least once in the experiment; last 10 rounds	Lied at least once in the experiment; last 10 rounds
Pseudo *R*-squared	0.029	0.341

Intensive margin

- In the last 10 rounds, participants lie less often in Control30 and 50-50Control30 than in other conditions, conditional on lying at least once during the whole experiment

- The analyses of dynamics reveal that 50-50NoControl leads to less often lies, conditionally on lying at least once, than NoControl in the first rounds, but the effect vanishes over time (not significant within the first eight rounds)

Malfunctioning institution

- The 50-50NoControl treatment combines the worst of the two worlds:
  - it makes one more likely to lie at least once than in NoControl (extensive margin)
  - and it makes one lie as frequently, conditional on lying at least once, as in the NoControl (intensive margin)

Malfunctioning institution vs. weak institution

- One might argue the detrimental effect of the 50-50NoControl treatment is driven by, in expectation, a low probability of control, and subjects perceiving it as a weak institution ("pay-enough-or-don't-pay-at-all" type of behavior)

- We argue, however, that the detrimental effect of the malfunctioning institution goes beyond the effect of a small probability of control

- We argue the ***threat*** of the ***strong*** control institution and no monitoring possibility are what drive the detrimental result

Weak institution

- **Weak-institution treatment (Control0.03):**
  - In each round, control with a 0,03% probability
  - Independently of their truthfulness, after each round, subjects learned whether they had been controlled in that round
  - If controlled and lied – no payoff; otherwise – payoff equal to the number reported in euros

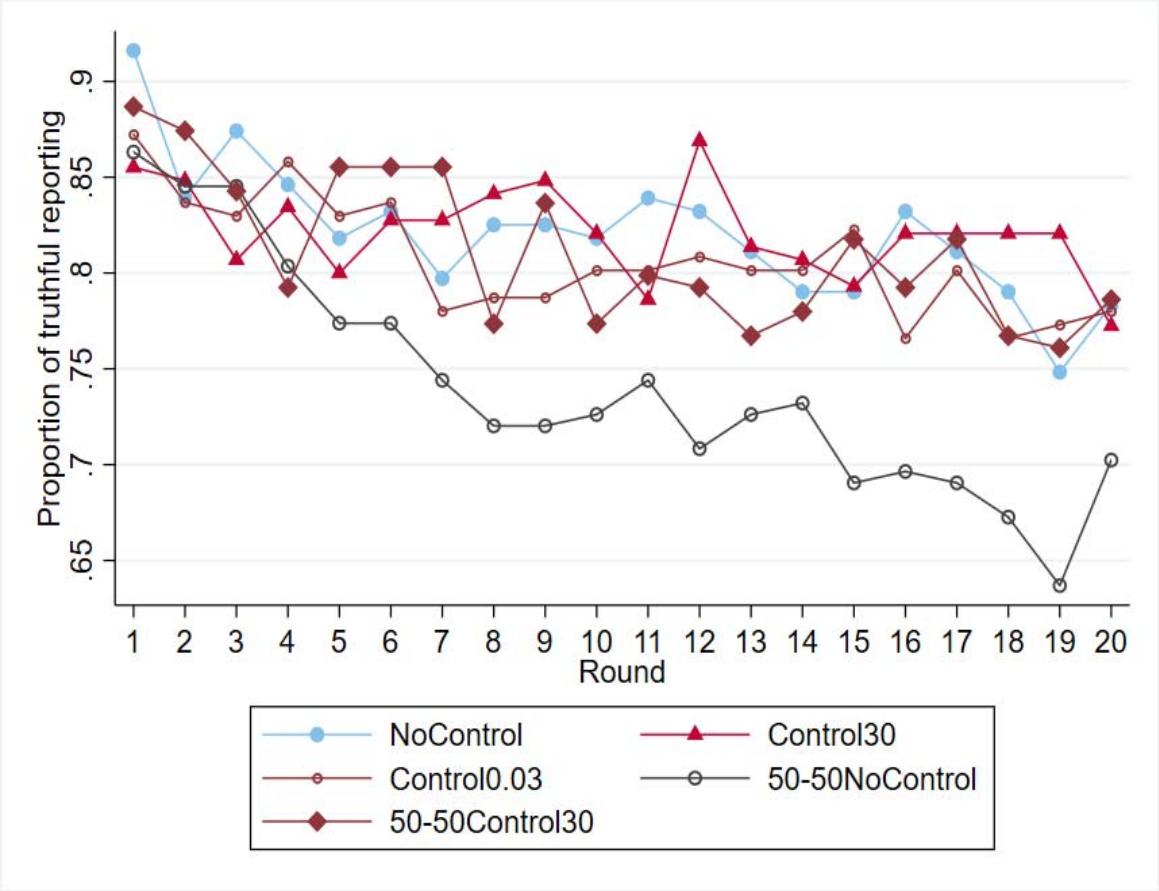
Hypotheses

- Truthfulness in the last 10 rounds:

NoControl > Control0.03 > 50-50NoControl

- Control0.03 crowds out intrinsic motivation and has no real monitoring function
- But the crowding out is not as high as when threatening with a strong institution

Weak institution



Weak institution

- We find no overall effect of weak institutions on truthfulness
- A weak institution does not crowd out people's honesty, but a malfunctioning institution does
- Overall truthfulness in the last 10 rounds:

Control30 = 50-50Control30 = NoControl = Control0.03 > 50-50NoControl

Conclusion

- The detrimental effect of malfunctioning institutions originates from the threat of a strong control institution crowding out intrinsic motivation for some individuals to tell the truth

- Learning almost with certainty that no strong institution is in place does not crowd in intrinsic motivation

Conclusion

- The detrimental effect of malfunctioning institutions originates from the threat of a strong control institution crowding out intrinsic motivation for some individuals to tell the truth

- Learning almost with certainty that no strong institution is in place does not crowd in intrinsic motivation

- The main contribution of this paper is the discovery of a new behavioral regularity: we show that a threat of punishment that is not implemented leads to severe detrimental effects

Conclusion

- The detrimental effect of malfunctioning institutions originates from the threat of a strong control institution crowding out intrinsic motivation for some individuals to tell the truth

- Learning almost with certainty that no strong institution is in place does not crowd in intrinsic motivation

- The main contribution of this paper is the discovery of a new behavioral regularity: we show that a threat of punishment that is not implemented leads to severe detrimental effects

- A clear policy implication from our study is that having trust institutions in place is better than threats of strong institutions that will not be implemented

Thank you!

Other papers on control and cheating

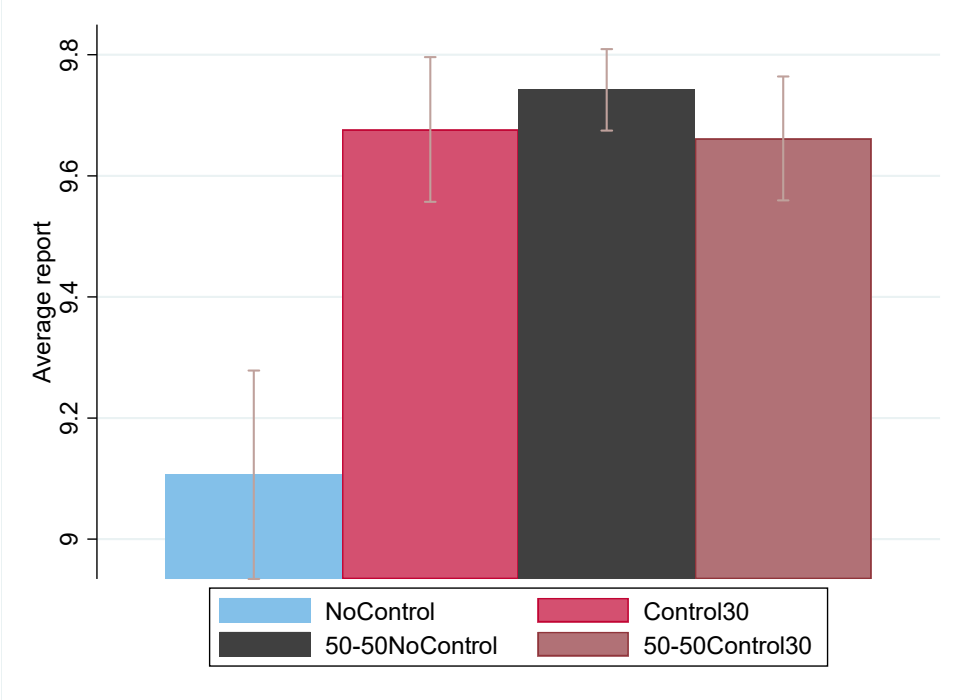
- Laske et al. (2018):
  - How does the size of the fine and the probability of being caught affect cheating?
  - Repeated deception game
  - The higher and more likely the fine, the less participants cheated

- Galeotti et al. (2021):
  - Spillover effects of control
  - A quasi experiment with two stages
  - Stage 1: being checked on a public transport
  - Stage 2: actor picks up a banknote of 5 euros and asks the participant whether it is theirs
  - More cheating about the banknote after seeing checks on the public transport

Intensive margin: Dynamics

	Truthful (1)	Truthful (2)
Control30	0.194***	0.173***
	(0.053)	(0.050)
50-50NoControl	0.141***	-0.073
	(0.052)	(0.051)
50-50Control30	0.221***	0.165***
	(0.050)	(0.046)
Female	0.011	0.008
	(0.034)	(0.031)
Age	0.000	0.001
	(0.004)	(0.004)
Draw	0.089***	0.088***
	(0.002)	(0.002)
Observations	784	1491
Number of clusters	216	307
Sample	Those who lied at least once in rounds 1 to 5; first 5 rounds	Those who lied at least once in rounds 1 to 20; last 5 rounds
Pseudo *R*-squared	0.400	0.349

Intensive margin: Average report, given a lie



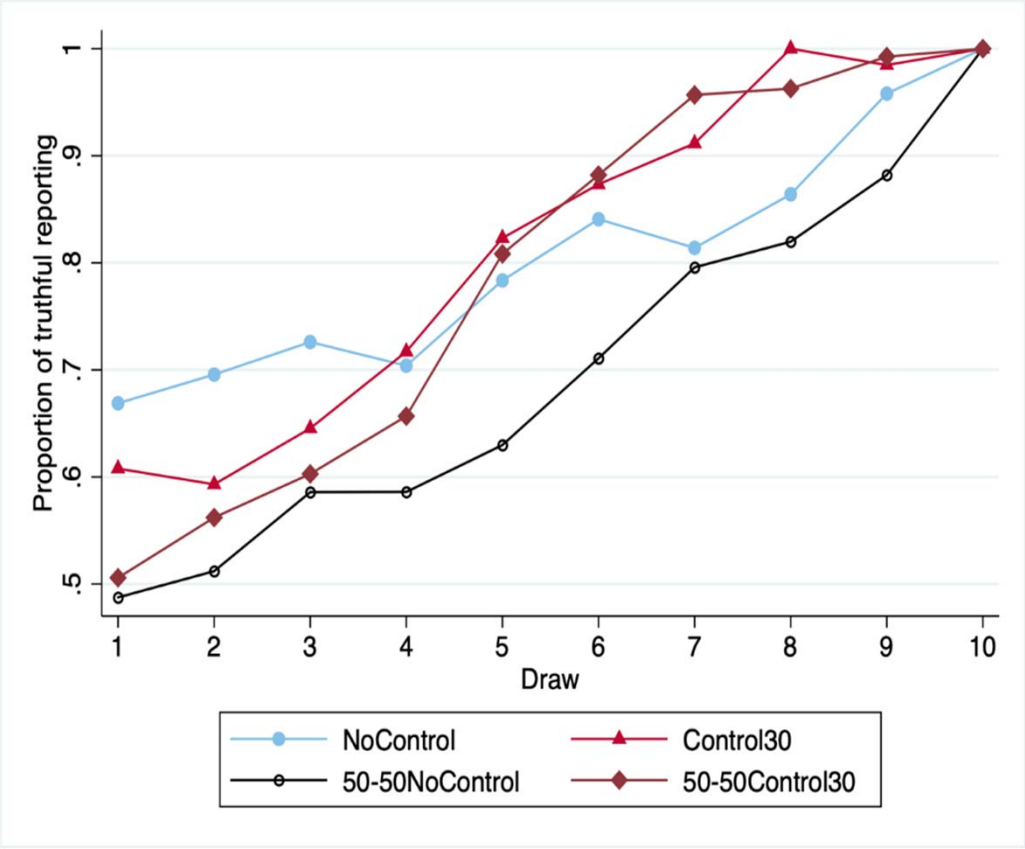
Intensive margin: Average report, given a lie

	Report (1)	Report (2)
Control30	0.570***	0.630***
	(0.187)	(0.184)
50-50NoControl	0.636***	0.648***
	(0.175)	(0.169)
50-50Control30	0.555***	0.631***
	(0.184)	(0.179)
Age		-0.002
		(0.010)
Female		-0.085
		(0.096)
Draw		0.075***
		(0.015)
Round		0.014
		(0.009)
Constant	9.106***	8.673***
	(0.163)	(0.311)
Observations	1395	1395
Number of clusters	289	289
Sample	Dishonest; last 10 rounds	Dishonest; last 10 rounds
R-squared	0.052	0.076

Does observing control affect lying differently in Control30 and 50-50?

	Truthful (1)	Truthful (2)	Truthful (3)	Truthful (4)
No control so far	-0.007	0.000	-0.048**	-0.053**
	(0.026)	(0.026)	(0.022)	(0.023)
Female	0.090**	0.091**	0.072***	0.067**
	(0.036)	(0.036)	(0.027)	(0.028)
Age	0.004	0.005	0.009***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)
Draw	0.049***	0.050***	0.054***	0.056***
	(0.005)	(0.005)	(0.003)	(0.003)
Observations	1450	2175	3270	4905
Number of clusters	145	145	327	327
Sample	First 10 Control30	First 15 Control30	First 10 50-50	First 15 50-50
Pseudo R-squared	0.20	0.20	0.21	0.20

Proportion of truthful reporting by draws and treatments in the last 10 rounds



Proportion of truthful reporting by draws and treatments in the last 10 rounds for those who lied at least once during the experiment