Spring 2022

Essays in Corruption and Capture

Trevor Incerti
Yale University Graduate School of Arts and Sciences, tnincerti@gmail.com

Abstract

Essays in Corruption and Capture

Trevor Incerti

2022

Corruption is commonly defined as the abuse of public office for private gain. However, private gain can be extracted from political processes or connections in ways that often do not constitute "abuse" or meet legal definitions of corruption. This volume explores some of the myriad ways private and public actors use political decisions to extract private gains—legal or extralegal—as well as explores potential mechanisms to reduce these actions.

The first two chapters examine the degree to which individual citizens can reduce extraction of private gain from public office through participation in politics. The first chapter examines the average electoral cost of corruption to politicians across countries. A systematic meta-analysis of all randomized experiments examining voting for candidates revealed to be corrupt to date shows that corrupt candidates are punished by zero percentage points on average across field experiments, but 32 points in survey experiments. This suggests that the "true" or average effect of voter punishment of revealed corruption is likely to be small in magnitude in actual elections, and that the high levels of electoral punishment of corruption found in survey environments may stem from social desirability bias and the lower and hypothetical nature of costs.

Next, I examine how to mobilize individuals to organize against interest group capture of local political institutions and regulations. Specifically, I examine the case of zoning regulations in large US cities. In these cities, homeowners are more likely than renters to participate in local politics across numerous dimensions, and zoning regulations reflect the preferences of this group. I conducted eight field experiments to investigate how to motivate renters (n=19,951 households) to comment at city council meetings in opposition to regulations that harm them. Opening a message highlighting high costs of abstention

caused a large increase in both total public comments and pro-housing comments across all treated meetings. These results suggest that increasing the perception that abstention is costly is an effective motivator of collective action, and that outreach can make civic bodies greater reflect the broader public where increases in accessibility alone do not.

The third chapter investigates the revolving door—a phenomenon long hypothesized to be a conduit to corruption and/or regulatory capture. However, data tracing flows of civil servants from the bureaucracy to the private sector remains rare. I create a dataset containing individual-level data of all Japanese bureaucrats who retired into positions outside of the bureaucracy over the past decade. I document how the data was constructed through a combination of digitization of thousands of PDF documents, web-scraping, and linkages with financial databases. Next, I describe what the data illuminates about the revolving door in Japan, such as a bifurcated job market in which high ranking officials are more likely to join the private sector while lower ranked bureaucrats commonly join non-profit organizations. I conclude by discussing how the data can be used to investigate empirical questions in subjects such as corruption, regulatory capture, procurement, pork, government waste, bureaucratic representation, and international political economy.

The final chapter examines rent extraction by rational actors of the global financial system in a legal capacity, but with large consequences for capital access in developing economies. Extant theories posit that political instability discourages investment. I challenge this consensus by showing that instability does not systematically depress access to financial capital, but rather capital access can increase when instability is expected to lead to increased financial returns in the future. Using an event study approach, I examine daily returns of national financial indices in every country that experienced an irregular regime change subject to data availability. Returns following resignations are positive, while those following assassinations and coups are negative and smaller in magnitude. However, a pro-capitalist coup results in large positive returns (+10%), and authoritarian or anti-capitalist regime changes are more likely to lead to capital flight than democratic or pro-business changes.

The immediate impact of political instability on capital access is therefore dependent on the type of regime change and its expected impact on future growth.

In sum, I document four channels through which individuals or interest groups extract rents from political processes—political corruption, political capture of regulatory bodies, the revolving door, and the global financial system. I first examine the degree to which citizen action can be utilized to reduce corruption and capture. Next, I create the most detailed dataset of bureaucratic revolving door hires constructed to date and illuminate previously under-appreciated aspects of revolving door networks. Finally, I detail and measure how global investors extract financial returns from political instability.

Essays in Corruption and Capture

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Trevor Incerti

Dissertation Director: Frances McCall Rosenbluth

May 2022

Acknowledgments

This dissertation is dedicated to Frances McCall Rosenbluth. The level of support—both academic and personal—she provided was one that few PhD students get to experience, and for which I will always be grateful.

Frances demonstrated what academia could look like with more kindness, thoughtfulness, and fewer boundaries. She was an intellectual tour-de-force, with research ranging from Japanese politics to the political economy of gender, partisan competition in America, war, diplomacy, and more. She pursued this amazing breadth of research because she was intellectually curious and it brought her fulfillment. Frances wanted this sense of fulfillment for her students, and knew that it could never come from work alone. She encouraged her students to live balanced lives and pursue activities that brought us joy. She taught us that academia—and work in general—should be a part of that balance, and that this meant pursuing your own curiosities and goals. Frances was never performative, she was genuine.

In my first conversation with Frances as my advisor, she asked if I needed help finding an apartment. She regularly asked about my life at home, my partner, and my family. After my father suffered a serious injury, she checked in regularly and encouraged me to take time away from work. She regularly invited her advisees to dinner and took us on walks with her dogs. She broke the bureaucratic and funding barriers that cost students time and energy. She signed her emails with XXOO. Reading this list of kindnesses big and small from an advisor would shock most PhD students. If academia is in my future, I will do all I can to treat my students with even a modicum of the kindness and sincerity Frances showed me.

I am also grateful to the other members of my committee who helped me through the program. It was an honor to learn from a giant of the field like Susan Rose-Ackerman, whose accomplishments are awe-inspiring. She cemented my passion for the study of corruption, always gave thoughtful advice, and offered new perspectives and directions on my research whenever we spoke. P.M. Aronow went above and beyond by workshopping papers in ways that were often critical to their success, and I never left a conversation with P without new ideas for how to improve my work. Last but not least, P was an indispensable source of methodological training throughout my PhD. Josh Kalla's feedback was essential, as he pushed my work from idea to the finish line with concrete advice that always steered me in the right direction. Ken Scheve also deserves special mention for stepping in as a senior advisor without hesitation during an incredibly trying time, and offering the exact support I needed.

Other faculty members beyond my committee also provided support. A special mention goes to Charles Crabtree for regularly exceeding all reasonable expectations of a what a great mentor, colleague, and friend can be, and for being another exemplar of what a kinder academy could look like. Alex Coppock, Greg Huber, Dan Mattingly, Fredrik Sävje, and Ian Shapiro always had their door open to me and provided helpful advice. Alex Debs was a welcoming and incredibly helpful DGS, especially during hard times.

I would also like to thank the organizations that collaborated with me or provided generous funding and support: Abundant Housing LA, the Japan Foundation Center for Global Partnership, the University of Tokyo Institute of Social Science, the Waseda University Institute of Political Economy, the Yale Center for the Study of American Politics, the Yale Council for East Asian Studies, the Yale Institution for Social and Policy Studies, the Yale South Asian Studies Council, and YIMBY Law.

The entire Japanese politics community also deserves special mention for their general spirit of kindness of support. I would be remiss not to give special thanks to Professors Phillip Lipscy, Kenneth McElwain, Daniel M. Smith, and Steve Vogel in particular for their immense support throughout various stages of my career.

A large debt of thanks goes to my graduate student colleagues at Yale and elsewhere who read and commented on this work throughout its development: Amy Basu, Angèle Delevoye, Matthew Graham, Annabelle Hutchinson, Hilary Matfess, Colin Moreshead, Cleo O'Brien Udry, Lilla Orr, Mina Pollmann, Tomoya Sasaki, Collin Schumock, and Hikaru Yamagishi.

I feel fortunate that I will look back on graduate school fondly due to those I shared the experience with, and regret that COVID-19 cut much of our time short. Hikaru Yamagishi was the first friend I made at Yale, and our journey from her initial kind welcome to close friend and serial collaborator was a true highlight of my graduate school experience. Colin Moreshead always managed to remind me that academia—both the work and the journey—can and should be fun. Nights spent around the fire with Matt Graham were always a welcome respite from the stresses of teaching and research. Trips to Ordinary with Annabelle Hutchinson and Lilla Orr were always the highlight of my week. Abeer Obaid and Rachel Ooi were wonderful roommates and partners in comedy and commiserations. Amy Basu was a constant source of support and is the most stylish political scientist I know. Angèle Delevoye schooled us all both in the classroom and on whatever court she stepped on. Hanging out with Collin Schumock after work was always the perfect way to unwind. Mike Goldfien was a huge support through some of the most academically challenging points of the PhD. I will never forget the sounds of laughter that were a clear signal that Cleo O'Brien Udry and Hilary Mattfess were near. Nate Grubman and Louis Wasser deserve a mention for organizing our softball team, the Standard Errors. The wonderful grad students at MIT who I become close with after moving to Cambridge in the pandemic years of my PhD also deserve mention, particularly Zachary Burdette, Apekshya Prasai, and Tomoya Sasaki.

My family was my constant source of support and motivation. Nobody did more to support me throughout this journey than my partner Mina. I use the term journey literally, as we bounced from New England to California to the Pacific Northwest to Tokyo and back. She shared all of my trials and tribulations, and drove me to also invest in life beyond the computer screen. It would not be an exaggeration to state that this document would never have been completed without her support. My brother Devin inspired my academic pursuits, gave me the passion to study political economy, and continued to support my research and even partner in it throughout graduate school. I cannot imagine a better role model. Finally, I would not be where I am today without the relentless support of my parents, and the value they taught me existed in education in my early years.

Dedicated to Frances McCall Rosenbluth.

A towering intellect, caring mentor, and dearest friend.

Contents

List of Figures **viii**

List of Tables **xi**

1 Corruption information and vote share: A meta-analysis and lessons for experimental design **1**

Introduction 2
Corruption information and electoral accountability 3
Research Design and Methods 5
Results 10
Exploring the discrepancy 12
Discussion 26
Conclusion 28
Appendix 30

2 Combatting capture in local politics: Evidence from eight field experiments **57**

Motivation 60
Theory and hypotheses 62
Research design 68
Results 75
Conclusion 82
Appendix 84

3 Amakudata: a new dataset of bureaucratic revolving door hires **98**

Introduction 99
A brief review of revolving door literature 100
Dataset creation 101
Reexamining bureaucratic connections 104
Empirical applications and future research 109

Conclusion . . . 112
Appendix . . . 113

4 Are regime changes always bad economics? Evidence from daily financial data 120
Reexamining capital response to political instability . . . 122
Data . . . 124
Estimation . . . 129
Impact of Political Instability on Stock Returns . . . 132
Exploring possible mechanisms . . . 139
Robustness . . . 143
Conclusion . . . 147
Appendix . . . 149

5 References 156

List of Figures

1.1 Field experiments: Average treatment effect of corruption information on incumbent vote share and 95% confidence intervals 11
1.2 Survey experiments: Average treatment effect of corruption information on incumbent vote share and 95% confidence intervals 11
1.3 Breitenstein (2019) conjoint: average marginal component effects 24
1.4 Breitenstein (2019) conjoint: can the right candidate overcome corruption? . 24
1.5 Breitenstein (2019) conjoint decision tree: predicted probabilities of voting for candidate 26
1.6 Lab experiments: Average treatment effect of corruption information on vote share 30
1.7 Field experiments: Average treatment effect of corruption information on incumbent vote share (excluding Banerjee, Green, Green and Pande (2010) and Banerjee, Kumar, Pande and Su (2011)) 35
1.8 Survey experiments: Average treatment effect of corruption information on incumbent vote share (including De Figueiredo, Hidalgo and Kasahara (2011)) 35
1.9 P-curve: all experiments 39
1.10 P-curve: survey experiments 40
1.11 P-curve: field experiments 41
1.12 Funnel plot: all experiments 42
1.13 Funnel plot including trim and fill “missing” studies: all experiments 43
1.14 Funnel plot: all experiments with field experiment moderator 44
1.15 Funnel plot: field experiments 45
1.16 Funnel plot: survey experiments 46
1.17 All experiments by publication status: Average treatment effect of corruption information on vote share and 95% confidence intervals 47
1.18 Survey experiments by information quality: Average treatment effect of corruption information on vote share and 95% confidence intervals 48
1.19 Breitenstein (2019) conjoint: can the right candidate overcome corruption (clean challenger)? 49

1.20 Breitenstein (2019) conjoint: probability of choosing candidate (by clean or corrupt) . . . 49
1.21 Breitenstein (2019) conjoint decision tree: predicted probabilities of voting for corrupt politician with clean challenger . . . 50
1.22 Franchino and Zucchini (2015) conjoint: AMCEs . . . 51
1.23 Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (conservative respondents)? . . . 52
1.24 Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (liberal respondents)? . . . 52
1.25 Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (conservative respondents and clean challenger)? . . . 53
1.26 Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (liberal respondents and clean challenger)? . . . 53
1.27 Mares and Visconti (2019) conjoint: AMCEs . . . 54
1.28 Mares and Visconti (2019) conjoint: can programmatic offerings and experience overcome corruption? . . . 55
1.29 Mares and Visconti (2019) conjoint: can programmatic offerings and experience overcome corruption (conditional on other illicit activities)? . . . 55
1.30 Chauchard, Klašnja and Harish (2019) conjoint: can performance, partisanship, and coethnicity overcome corruption? . . . 56
1.31 Chauchard, Klašnja and Harish (2019) conjoint decision tree: predicted probabilities of voting for corrupt politician . . . 56

2.1 Treatment groups . . . 67
2.2 Intent-to-treat effect and complier average causal effect, all cities . . . 75
2.3 Meta-analysis of complier average causal effects, by council meeting . . . 76
2.4 Effects by treatment group, all cities . . . 77
2.5 Complier average causal effects by turnout . . . 79
2.6 CACE by type of comment . . . 79
2.7 Change in housing net worth by age and income percentile . . . 84
2.8 Map of cities in Los Angeles county by experiment status . . . 87
2.9 Average treatment effect on email opening, all cities . . . 89
2.10 Average treatment effect on email opening, by city . . . 89
2.11 Intent-to-treat effect and complier average causal effect, all cities (without covariate adjustment) . . . 94
2.12 Effects by treatment group, all cities (without covariate adjustment) . . . 94
2.13 Meta-analysis of complier average causal effects by city, excluding pilot studies 95
2.14 Distibution of outcomes by treatment group (compliers only) . . . 96
2.15 Bayesian multilevel model: coefficient estimates and posterior distributions (includes city fixed effects) . . . 96
2.16 Posterior distributions of costly abstention treatment, instructions only treatment, and difference . . . 97

3.1 Top 10 *amakudari* destinations vs. overall economy . . . 104
3.2 *Amakudari* destinations by firm type . . . 105

3.3 *Amakudari* ministry of origin 106
3.4 Flows of bureaucrats from ministries to top 10 public interest corporations (by number of hires) 107
3.5 Example original PDF document 113
3.6 *Amakudari* hires by industry and ministry (all years) 114
3.7 *Amakudari* appointments by ministry over time 115
3.8 *Amakudari* ministry of origin adjusted for ministry size (all years) 115
3.9 Cumulative abnormal returns from assistant vice-minister and vice- minister appointments 116
3.10 *Amakudari* vs. non-*amakudari* firm financials 116
3.11 Age of exit from ministry, by ministry 117
3.12 Cumulative abnormal returns from assistant vice-minister and vice- minister appointments 118
3.13 University hires of MEXT officials 119

4.1 Absolute value of daily returns 129
4.2 Mean of volatility estimates from GARCH(1,1) models 133
4.3 Abnormal returns surrounding the 2002 Venezuelan coup attempt 142
4.4 Mean cumulative abnormal returns by type of regime change 145
4.5 Cumulative abnormal returns during the Egyptian revolution 150
4.6 Time-shifted placebo sensitivity analysis of mean event day abnormal return by type of regime change 153
4.7 Abnormal returns surrounding the 2016 Turkish coup attempt 154

List of Tables

1.1 Field experiments . . . 7
1.2 Survey experiments . . . 8
1.3 Lab experiments . . . 30
1.4 Excluded experiments . . . 31
1.5 Meta-analysis by type of experiment . . . 32
1.6 Random effects meta-analysis (all studies) . . . 32
1.7 Mixed effects meta-analysis with survey experiment moderator . . . 32
1.8 Meta-analysis (all field experiments excluding Banerjee et al. (2010) and Banerjee et al. (2011)) . . . 33
1.9 Random effects meta-analysis (all studies excluding Banerjee et al. (2010) and Banerjee et al. (2011)) . . . 33
1.10 Mixed effects meta-analysis with survey experiment moderator (excluding Banerjee et al. (2010) and Banerjee et al. (2011)) . . . 33
1.11 Meta-analysis (all survey experiments including De Figueiredo, Hidalgo and Kasahara (2011) . . . 34
1.12 P-values by study . . . 36
1.13 Do p-values predict publication status? . . . 37
1.14 Regression tests for funnel plot asymmetry . . . 37
1.15 Trim and fill estimates by subgroup . . . 37
1.16 PET-PEESE estimates by subgroup . . . 38

2.1 Examination of public comments in treated council meetings . . . 81
2.2 Covariate balance and difference in means test: treatment vs. placebo . . . 85
2.3 Covariate balance across all treatment groups . . . 86
2.4 Covariate predictiveness of compliance by treatment group . . . 90
2.5 Intent-to-treat effects . . . 91
2.6 Complier average causal effects . . . 92
2.7 Conditional complier average causal effect . . . 93
2.8 ITT and CACE estimates from penalized maximum likelihood . . . 95

3.1 Amakudari dataset example 103
3.2 (Log of) financials of firms hiring top ranking officials vs. firms hiring lower ranked officials 117

4.1 List of stock indices 126
4.2 Regime changes 127
4.3 Abnormal returns following coups 134
4.4 Abnormal returns following assassinations 135
4.5 Abnormal returns following resignations 137
4.6 Abnormal returns following authoritarian regime changes 140
4.7 Abnormal returns following democratic regime changes 141
4.8 Non-parametric tests of the impact of regime changes 146
4.9 List of public protests preceding resignations 149
4.10 Effect of public protests on stock prices 152

chapter 1

Corruption information and vote share: A meta-analysis and lessons for experimental design

Abstract: Debate persists on whether voters hold politicians accountable for corruption. Numerous experiments have examined if informing voters about corrupt acts of politicians decreases their vote share. Meta-analysis demonstrates that corrupt candidates are punished by zero percentage points across field experiments, but approximately 32 points in survey experiments. I argue this discrepancy arises due to methodological differences. Small effects in field experiments may stem partially from weak treatments and noncompliance, and large effects in survey experiments from social desirability bias and the lower and hypothetical nature of costs. Conjoint experiments introduce hypothetical costly tradeoffs, but it may be best to interpret results in terms of realistic sets of characteristics rather than marginal effects of particular characteristics. These results suggest that survey experiments may provide point estimates that are not representative of real-world voting behavior. However, field experimental estimates may also not recover the "true" effects due to design decisions and limitations.

Introduction

Competitive elections create a system whereby voters can hold policy makers accountable for their actions. This mechanism should make politicians hesitant to engage in malfeasance such as blatant acts of corruption. Increases in public information regarding corruption should therefore decrease levels of corruption in government, as voters armed with information expel corrupt politicians (Kolstad and Wiig 2009; Rose-Ackerman and Palifka 2016). However, this theoretical prediction is undermined by the observation that well-informed voters continue to vote corrupt politicians into office in many democracies. Political scientists and economists have therefore turned to experimental methods to test the causal effect of learning about politician corruption on vote choice.

Numerous experiments have examined whether providing voters with information about the corrupt acts of politicians decreases their re-election rates. These papers often suggest that there is little consensus on how voters respond to information about corrupt politicians (Arias, Larreguy, Marshall and Querubin 2018; Botero, Cornejo, Gamboa, Pavao and Nickerson 2015; Buntaine, Jablonski, Nielson and Pickering 2018; De Vries and Solaz 2017; Klašnja, Lupu and Tucker 2017; Solaz, De Vries and de Geus 2019). Others indicate that experiments have provided us with evidence that voters strongly punish individual politicians involved in malfeasance (Chong, De La O, Karlan and Wantchekon 2014; Weitz-Shapiro and Winters 2017; Winters and Weitz-Shapiro 2015,1).

By contrast, meta-analysis suggests that: (1) In aggregate, the effect of providing information about incumbent corruption on incumbent vote share in field experiments is approximately zero, and (2) corrupt candidates are punished by respondents by approximately 32 percentage points across survey experiments. This suggests that survey experiments may provide point estimates that are not representative of real-world voting behavior. Field experimental estimates may also not recover the "true" effects due to design decisions and limitations.

I also examine mechanisms that may give rise to this discrepancy. I do not find systematic evidence of publication bias. I discuss the possibility that social desirability bias may lead survey respondents to under-report socially undesirable behavior. The costs of changing one's vote is also lower and more abstract in hypothetical environments. In field experiments, the magnitude of treatment effects may be small due to weak treatments and noncompliance. Field and survey experiments also may be measuring different causal estimands due to differences in context and survey design. Finally, surveys may not capture the complexity and costliness of real-world voting decisions. Conjoint experiments attempt to alleviate some of these issues, but are often analyzed in ways that may fail to illuminate the most substantively important comparisons. I suggest examining the probability of voting for candidates with specific combinations of attributes in conjoint experiments when researchers have priors about the conditions that shape voter decision-making, and using classification trees to illuminate these conditions when they do not.

I therefore (1) find that the "true" or average effect of voter punishment of revealed corruption remains unclear, but is likely to be small in magnitude in actual elections, (2) show that researchers should use caution when interpreting point estimates in survey experiments as indicative of real world behavior, (3) explore methodological reasons that estimates may be particular large in surveys and small in field experiments, and (4) offer suggestions for design and analysis of future experiments.

Corruption information and electoral accountability

Experimental support for the hypothesis that providing voters with information about politicians' corrupt acts decreases their re-election rates is mixed. Field experiments have provided some causal evidence that informing voters of candidate corruption has negative (but generally small) effects on candidate vote-share. This information has been provided by: randomized financial audits (Ferraz and Finan 2008), fliers revealing corrupt actions of politicians (Chong et al. 2014; De Figueiredo, Hidalgo and Kasahara 2011), and SMS messages

(Buntaine et al. 2018). However, near-zero and null findings are also prevalent, and the negative and significant effects reported above sometimes only manifest in particular subgroups. Banerjee et al. (2010) primed voters in rural India not to vote for corrupt candidates, and Banerjee et al. (2011) provided information on politicians' asset accumulation and criminality, with both studies finding near-zero and null effects on vote share. Boas, Hidalgo and Melo (2019) similarly find zero and null effects from distributing fliers in Brazil. Finally, Arias et al. (2018); Arias, Balán, Larreguy, Marshall and Querubín (2019) find that providing Mexican voters with information (fliers) about mayoral corruption actually *increased* incumbent party vote share by 3%.[1]

By contrast, survey experiments consistently show large negative effects from informational treatments on vote share for hypothetical candidates. These experiments often manipulate moderating factors other than information provision (e.g. quality of information, source of information, partisanship, whether corruption brings economic benefits to constituents, etc.), but even so systematically show negative treatment effects (Anduiza, Gallego and Muñoz 2013; Avenburg 2019; Banerjee, Green, McManus and Pande 2014; Boas, Hidalgo and Melo 2019; Breitenstein 2019; Eggers, Vivyan and Wagner 2018; Franchino and Zucchini 2015; Klašnja and Tucker 2013; Klašnja, Lupu and Tucker 2017; Mares and Visconti 2019; Vera 2019; Weitz-Shapiro and Winters 2017; Winters and Weitz-Shapiro 2013,1,1,2). These experiments have historically taken the form of single treatment arm or multiple arm factorial vignettes, but more recently have tended toward conjoint experiments (Agerberg 2020; Breitenstein 2019; Chauchard, Klašnja and Harish 2019; Franchino and Zucchini 2015; Klašnja, Lupu and Tucker 2017; Mares and Visconti 2019).

Boas, Hidalgo and Melo (2019) find differential results in a pair of field and survey experiments conducted in Brazil—zero and null in field; large, negative, and significant in survey. They argue that norms against malfeasance in Brazil are constrained by other factors at the polls, but that "differences in research design are unlikely to account for much of the

[1]The authors theorize that this average effect stems from levels of reported malfeasance actually being lower than voters' no-information expectations of corruption.

difference in effect size."[2] Boas, Hidalgo and Melo identify moderating factors specific to Brazil—low salience of corruption to voters in municipal elections and the strong effects of dynastic politics—to explain the small effects in their field experiment. However, meta-analysis demonstrates that this discrepancy exists not only in Boas, Hidalgo and Melo's experiments in Brazil, but extends across a systematic review of all countries and studies conducted to date. This suggests that the discrepancy between field and survey experimental findings is driven by methodological differences, rather than Brazil-specific features. I therefore enumerate features inherent in the research designs of field and survey experiments that may drive the small effects in field experiments and large effects in survey experiments.

Lab experiments that reveal corrupt actions of politicians to fellow players, then measure vote choice also show large negative treatment effects. While recognizing that the sample size of studies is extremely small, a meta-analysis of the three lab experiments that meet this study's selection criteria reveal a point estimate of approximately -33 percentage points (Arvate and Mittlaender 2017; Azfar and Nelson 2007; Solaz, De Vries and de Geus 2019) (see Figure 1.6).[3] This discrepancy is worth noting as previous examinations of lab-field correspondence have found evidence of general replicability (Camerer 2011; Coppock and Green 2015).

Research Design and Methods

Selection criteria

I followed standard practices to locate the experiments included in the meta-analysis. This included following citation chains and searches of data bases using a variety of relevant terms ("corruption experiment," "corruption field experiment," "corruption survey experiment,"

[2]The specific design differences Boas, Hidalgo and Melo note are unlikely to cause the discrepancy are differences in the language used between the information in the vignette and flier, and timing of outcome measurement.

[3]See Valentine, Pigott and Rothstein (2010) for a discussion of statistical power in meta-analysis. Note that Valentine, Pigott and Rothstein conclude that the minimum number of studies needed to conduct a meta-analysis is "two studies."

"corruption factorial", "corruption candidate choice", "corruption conjoint", "corruption, vote, experiment", and "corruption vignette"). Papers from any discipline are eligible for inclusion, but in practice stem only from economics and political science. Both published articles and working papers are included so as to ensure the meta-analysis is not biased towards published results. In total, I located 10 field experiments from 8 papers, and 18 survey experiments from 15 papers.

Field experiments are included if researchers randomly assigned information regarding incumbent corruption to voters, then measured corresponding voting outcomes. This therefore excludes experiments that randomly assign corruption information, but use favorability ratings or other metrics rather than actual vote share as their dependent variable. I include one natural experiment, Ferraz and Finan (2008), as random assignment was conducted by the Brazilian government. Effects reported in the meta-analysis come from information treatments on the entire sample of study only, not subgroup or interactive effects that reveal the largest treatment effects.

For survey experiments, studies must test a no-information or clean control group versus a corruption information treatment group and measure vote choice for a hypothetical candidate. This necessarily excludes studies that compare one type of information provision (e.g. source) to another and the control group is one type of information rather than no information, or where the politician is always known to be corrupt (Anduiza, Gallego and Muñoz 2013; Botero et al. 2015; Konstantinidis and Xezonakis 2013; Muñoz, Anduiza and Gallego 2012; Rundquist, Strom and Peters 1977; Weschle 2016). In many cases, studies have multiple corruption treatments (e.g. high quality information vs. low quality information, co-partisan vs. opposition party, etc.). In these cases, I replicate the studies and code corruption as a binary treatment (0 = clean, 1 = corrupt) where *all* treatment arms that provide corruption information are combined into a single treatment. Studies that use non-binary vote choices are rescaled into a binary vote choice.[4]

[4]For example, a 1-4 scale is recoded so that 1 or 2 is equal to no vote, and 3 or 4 is equal to a vote.

Included studies

A list of all papers - disaggregated by field and survey experiments - that meet the criteria outlined above are provided in Table 1.1 and Table 1.2. A list of lab experiments (4 total) can also be found in and Table 1.3, although these studies are not included in the meta-analysis. A list of excluded studies with justification for their exclusion can be found in Table 1.4.

Table 1.1: Field experiments

Study	Country	Treatment
Arias et al. (2018)	Mexico	Fliers
Banerjee et al. (2010)	India	Newspapers
Banerjee et al. (2011)	India	Newspapers
Boas, Hidalgo and Melo (2019)	Brazil	Fliers
Buntaine et al. (2018)	Ghana	SMS
Chong et al. (2014)	Mexico	Fliers
De Figueiredo, Hidalgo and Kasahara (2011)	Brazil	Fliers
Ferraz and Finan (2008)	Brazil	Audits

Table 1.2: Survey experiments

Study	Country	Type of survey
Agerberg (2020)	Spain	Conjoint
Avenburg (2019)	Brazil	Vignette
Banerjee et al. (2014)	India	Vignette
Breitenstein (2019)	Spain	Conjoint
Boas, Hidalgo and Melo (2019)	Brazil	Vignette
Chauchard, Klašnja and Harish (2019)	India	Conjoint
Eggers, Vivyan and Wagner (2018)	UK	Conjoint
Franchino and Zucchini (2015)	Italy	Conjoint
Klašnja and Tucker (2013)	Sweden	Vignette
Klašnja and Tucker (2013)	Moldova	Vignette
Klašnja, Lupu and Tucker (2017)	Argentina	Conjoint
Klašnja, Lupu and Tucker (2017)	Chile	Conjoint
Klašnja, Lupu and Tucker (2017)	Uruguay	Conjoint
Mares and Visconti (2019)	Romania	Conjoint
Vera (2019)	Peru	Vignette
Weitz-Shapiro and Winters (2017)	Brazil	Vignette
Winters and Weitz-Shapiro (2013)	Brazil	Vignette
Winters and Weitz-Shapiro (2020)	Argentina	Vignette

Additional selection comments

Additional justification for the inclusion or exclusion of certain studies, as well as coding and/or replication choices may be warranted in some cases. Despite often being considered a form of corruption (Rose-Ackerman and Palifka 2016), I exclude electoral fraud experiments as whether vote buying constitutes clientelism or corruption is a matter of debate (Stokes, Dunning, Nazareno and Brusco 2013). The field experiment conducted by Banerjee et al. (2010) is included. However, the authors treated voters with a campaign not to vote for corrupt candidates in general, but did not provide voters with information on which candidates were corrupt. Similarly, the field experiment conducted by Banerjee et al. (2011) is included, but their treatment provided information on politicians' asset accumulation and criminality, which may imply corruption but is not as direct as other types of information provision. The point estimates remain approximately zero when these studies are excluded from the meta-analysis (see Figure 1.7 and Table 1.8).

With respect to survey experiments, Chauchard, Klašnja and Harish (2019) include two treatments—wealth accumulation and whether the wealth accumulation was illegal. The effect reported here is the illegal treatment only. This is likely a conservative estimate, as the true effect is a combination of illegality and wealth accumulation. Winters and Weitz-Shapiro (2016) and Weitz-Shapiro and Winters (2017) report results from the same survey experiment, as do Winters and Weitz-Shapiro (2013) and Winters and Weitz-Shapiro (2015). Each of these results are therefore only reported once. The survey experiment in De Figueiredo, Hidalgo and Kasahara (2011) is excluded from the analysis as it does not use hypothetical candidates, but instead asks voters if they would have changed their actual voting behavior in response to receiving corruption information. This study has a slightly positive and null finding. Including this study, the point estimates are 32 and 31 percentage points using fixed and random effects estimation, respectively (see Figure 1.8 and Table 1.11).

Results

Survey experiments estimate much larger negative treatment effects of providing information about corruption to voters relative to field experiments. In fact, the field-experimental results in Figure 1.1 reveal a precisely estimated point estimate of approximately zero and suggest that we cannot reject the null hypothesis of no treatment effect (the 95% confidence interval is -0.56 to 0.15 percentage points using fixed effects and -2.1 to 1.4 using random effects). By contrast, Figure 1.2 shows that corrupt candidates are punished by respondents by approximately 32 percentage points in survey experiments based on fixed and random effects meta-analysis (the 95% confidence interval is -32.6 to -31.2 percentage points using fixed effects and -38.2 to -26.2 using random effects). Of the 18 survey experiments, only one shows a null effect (Klašnja and Tucker 2013), while all others are negative and significantly different from zero at conventional levels.

Examining all studies together, a test for heterogeneity by type of experiment (field or survey) reveals that up to 68% of the total heterogeneity across studies can be accounted for by a dummy variable for type of experiment (0 = field, 1 = survey) (see Table 1.7). This dummy variable has a significant association with the effectiveness of the information treatment at the 1% level. In fact, with this dummy variable included, the overall estimate across studies is -0.007, while the point estimate of the survey dummy is -0.315.[5] This implies that the predicted treatment effect across experiments is not significantly different from zero when an indicator for type of experiment is included in the model. In other words, the majority of the heterogeneity in findings is accounted for by the type of experiment conducted.

[5]Using a mixed effects model with a survey experiment moderator (see Table 1.7). With Banerjee et al. (2010) and Banerjee et al. (2011) excluded from the model, the point estimate of the survey dummy is 0.31 and the heterogeneity accounted for by the survey experiment moderator is 65% (see Table 1.10 and Table 1.9).

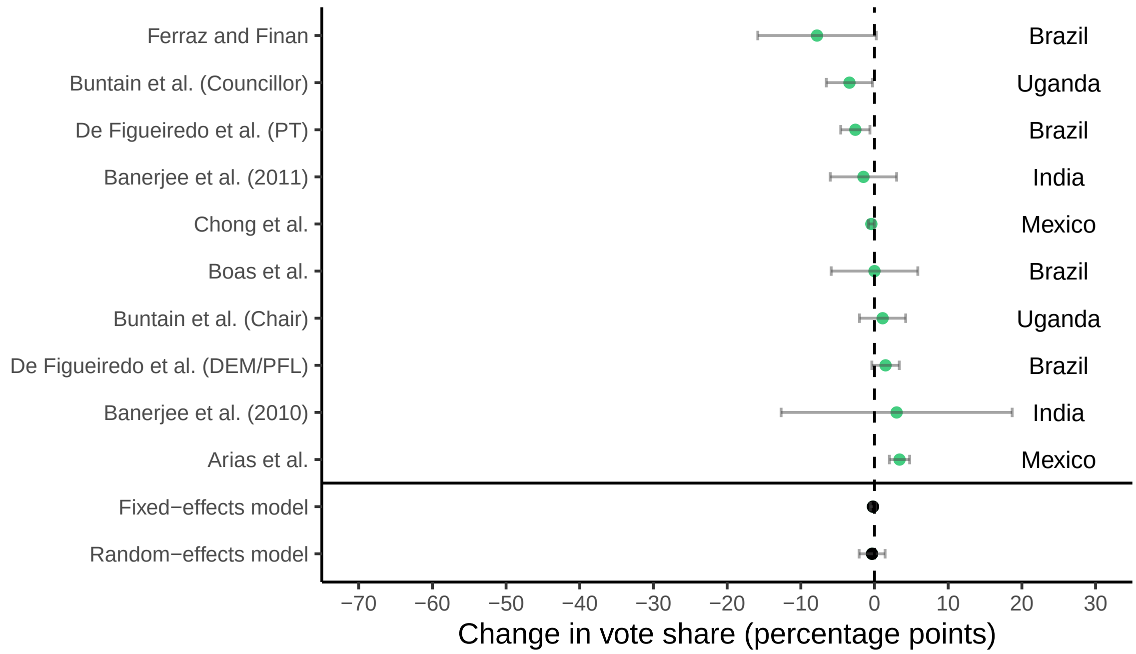


Figure 1.1: Field experiments: Average treatment effect of corruption information on incumbent vote share and 95% confidence intervals

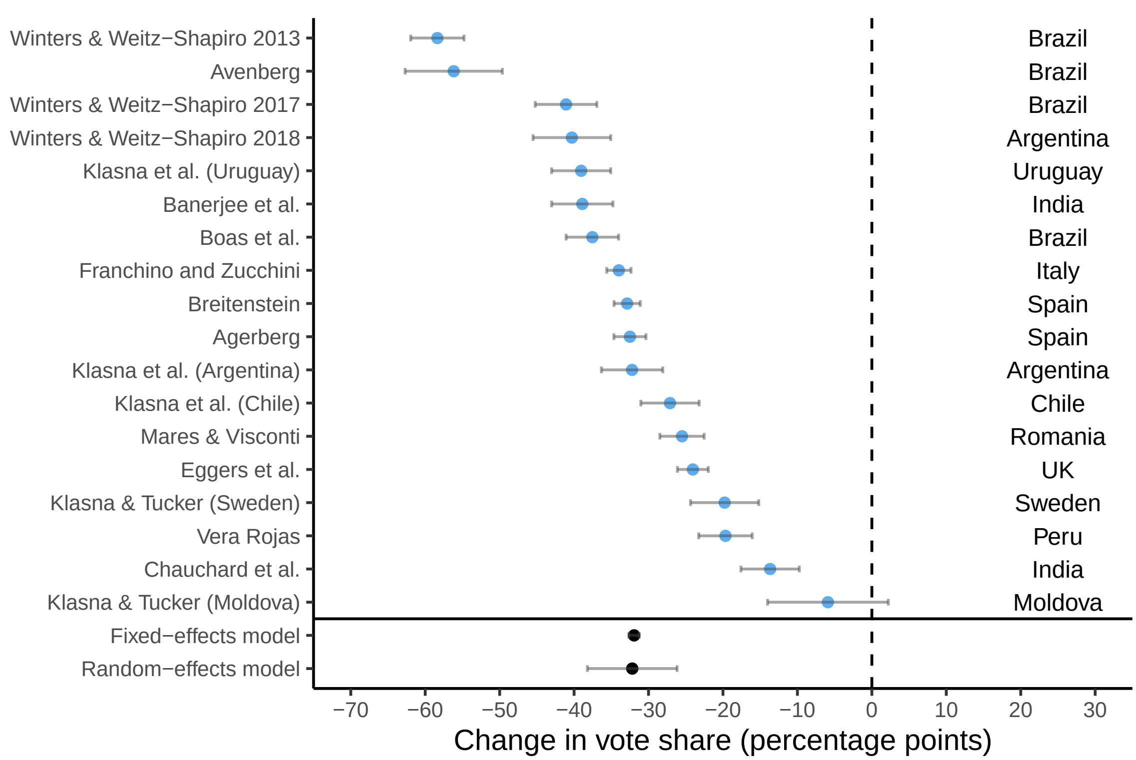


Figure 1.2: Survey experiments: Average treatment effect of corruption information on incumbent vote share and 95% confidence intervals

Exploring the discrepancy

What accounts for the large difference in treatment effects between field and survey experiments? One possibility is publication bias. Null results may be less likely to be published than significant results, particularly in a survey setting. A second possibility is social desirability bias, which may cause respondents to under-report socially undesirable behavior. Related is hypothetical bias, in which costs are more abstract in hypothetical environments. Survey and field experiments may also not mirror each other and/or real-world voting decisions. Potential ways in which the survey setting may differ from the field are: treatment salience and noncompliance, differences in outcome choices, and costliness/decision complexity. Weak treatments and noncompliance may decrease treatment effect sizes in field experiments. Design decisions may change the choice sets available to respondents. Finally, surveys may not capture the complexity and costliness of real-world voting decisions. It is possible that more complex factorial designs—such as conjoint experiments—may more successfully approximate real-world settings. However, common methods of analysis of conjoint experiments may not capture all theoretical quantities of interest.

Publication bias and p-hacking

Publication bias and p-hacking can lead to overestimated effects in meta-analysis (Carter, Schönbrodt, Gervais and Hilgard 2019; Duval and Tweedie 2000; Sterne, Egger and Smith 2001; Van Aert, Wicherts and Van Assen 2019). While I have identified heterogeneity stemming from the type of experiment performed as a potential source of overestimation, this may reflect that null results are less likely to be published than studies with large and significant negative treatment effects. I therefore now turn to the possibility of publication bias and/or p-hacking. In order to formally test for publication bias, I use the p-curve, examination of funnel plot asymmetry, trim and fill, and PET-PEESE[6] methods.[7]

[6] Precision effect test—precision effect estimate with standard error (Stanley and Doucouliagos 2014).

[7] Note that the "best" technique for assessing bias in meta-analysis varies by circumstance, and the proper test for each circumstance is a subject of active debate. See Carter et al. (2019) for a recent overview.

Of the eight field experimental papers located, only five are published. By contrast, only one of the 15 survey experimental papers remains unpublished, and this is a recent draft. This may reflect that the null results from field experiments are less likely to be published than their survey counterparts with large and highly significant negative treatment effects. While recognizing that the sample size of studies is small, OLS and logistic regression do not indicate that reported p-value is a significant predictor of publication status, although the directionality of coefficients is consistent with lower p-values being more likely to be published (Table 1.13). However, this simple analysis is complicated by the fact that the p-value associated with the average treatment effect across all subjects may not be the primary p-value of interest in the paper.

In order to more formally test for publication bias, I first use the p-curve (Simonsohn, Simmons and Nelson 2015; Simonsohn, Nelson and Simmons 2014*a*,1). The p-curve is based on the premise that only "significant" results are typically published, and depicts the distribution of statistically significant p-values for a set of published studies. The shape of the p-curve is indicative of whether or not the results of a set of studies are derived from true effects, or from publication bias. If p-values are clustered around 0.05 (i.e. the p-curve is left skewed), this may be evidence of p-hacking, indicating that studies with p-values just below 0.05 are selectively reported. If the p-curve is right skewed and there are more low p-values (0.01), this is evidence of true effects. All significant survey experimental results included in the meta-analysis are significant at the 1% level, implying that publication bias likely does not explain the large negative treatment effects in survey experiments.[8] For field experiments, there is not a large enough number of published experiments to make the p-curve viable.[9] Only six studies are published, and of these only four are significant at at least the 5% level.

Next, I test for publication bias by examining funnel plot asymmetry. A funnel plot

[8]See Figure 1.10 for a visual p-curve and formal test for right-skewness for survey experiments and Table 1.12 for a list of p-values associated with each study. There is also no indication of publication bias at the 1% level using this method.

[9]See Figure 1.11 for a visual p-curve and formal test for right-skewness for field experiments.

depicts the outcomes from each study on the x-axis and their corresponding standard errors on the y-axis. The chart is overlaid with an inverted triangular confidence interval region (i.e. the funnel), which should contain 95% of the studies if there is no bias or between study heterogeneity. If studies with insignificant results remain unpublished the funnel plot may be asymmetric. Both visual inspection and regression tests of funnel plot asymmetry reveal an asymmetric funnel plot when survey and field experiments are grouped together (see Figure 1.12 and Table 1.14). However, this asymmetry disappears when accounting for heterogeneity by type of experiment, either with the inclusion of a survey experiment moderator (dummy) variable or by analyzing field and survey experiments separately (see Table 1.14, Figure 1.14, Figure 1.15, and Figure 1.16.). Trim and fill analysis overestimates effect sizes and hypothesizes that three studies are missing due to publication bias when analyzing all studies together (see Figure 1.13 and Table 1.15). However, when trim and fill is used on survey experiments or field experiments as separate subgroups, estimates remain unchanged from random effects meta-analysis and no studies are hypothesized to be missing. Similarly, PET-PEESE estimates remain virtually unchanged when survey and field experiments are analyzed as separate subgroups.[10] [11]

In sum, while publication bias cannot be ruled out completely—particularly with such a small sample size of field experiments—there is no smoking gun. This implies that differences in experimental design likely account for the difference in the magnitude of treatment effects in field versus survey experiments, rather than publication bias.

[10] With all experiments grouped together, PET-PEESE estimates an effect of 0.8 percentage points (95% CI -4.5 to 6.2). See Table 1.16 for the results from PET-PEESE estimation.

[11] The results from this section are in accordance with the findings in Terrin, Schmid, Lau and Olkin (2003), Peters, Sutton, Jones, Abrams and Rushton (2007), Carter et al. (2019), and Van Aert, Wicherts and Van Assen (2019). Peters et al. (2007) show that trim and fill returns biased estimates under high between-study heterogeneity. Carter et al. (2019) find that both the trim and fill method and p-curve overestimate effect sizes and show high false positive rates in the presence of heterogeneity. van Aert, Wicherts and van Assen (2016) show similar findings with respect to p-curve estimation, which assumes homogenous effect sizes. PET-PEESE also assumes homogenous effect size and has been shown to be biased when between-study variance in effect sizes is large (Stanley and Doucouliagos 2017; Van Aert, Wicherts and Van Assen 2019). Reed, Florax and Poot (2015) show that random effects meta-analysis exhibits lower mean-squared error than PET-PEESE under high heterogeneity. Carter et al. (2019) recommend standard random effects meta-analysis (as performed here) if publication bias is unlikely.

Social desirability bias and hypothetical bias

A second possible explanation is social desirability or sensitivity bias, in which survey respondents under-report socially undesirable behavior. A respondent may think a particular response will be perceived unfavorably by society as whole, by the researcher(s), or both, and underreport such behavior. In the case of corruption, respondents are likely to perceive corruption as harmful to society, the economy, and their own personal well-being. They may therefore be more likely to choose the socially desirable option (no corruption), particularly when observed by a researcher or afraid of response disclosure.[12] However, a researcher is not the only social referent to whom a respondent may wish to give a socially desirable response. Respondents also may not wish to admit to themselves that they would vote for a corrupt candidate. Voting against corruption in the abstract may therefore reflect the respondents' actual preferences.

However, sensitivity bias is unlikely to account entirely for the difference in magnitude of treatment effects. A recent meta-analysis finds that sensitivity biases are typically smaller than 10 percentage points, and that respondents underreport vote buying by 8 percentage points on average (Blair, Coppock and Moor 2018). As vote buying is often considered a form of corruption, the amount of sensitivity bias present in corruption survey experiments may be similar.

A related but distinct source of bias is hypothetical bias. Hypothetical bias is often found in stated preference surveys in environmental economics, in which respondents report a willingness to pay that is larger than what they will actually pay using their own money as the costs are purely hypothetical (Loomis 2011). For corruption experiments, this would manifest as respondents reporting a willingness to punish corruption larger than in reality as the costs in terms of tradeoffs are purely hypothetical. There are few costs to selecting the socially

[12]Note, however, that social desirability bias differs from norms as norms reflect internalized values, whereas social desirability bias corresponds to misreporting due to fear of judgement by a social referent. Internalized norms would be reflected in both field and survey experimental studies. I would like to thank an anonymous reviewer for this insight. Also see Philp and David-Barrett (2015) for an in-depth discussion of how social norms interact with behavior surrounding corruption.

desirable option in a hypothetical survey experiment. By contrast, the cost of changing one's actual vote (as in field experiments) may be higher. Voters might have pre-existing favorable opinions of real candidates, discount corruption information, or have strong material and/or ideological incentives to stick with their candidate. As the informational treatment will only have an effect on supporters of the corrupt candidate who must change their vote—opponents have already decided not to vote for the candidate—these costs are particularly high. Where anticorruption norms are particularly strong—as in Brazil as highlighted by Boas, Hidalgo and Melo (2019)—the magnitude of hypothetical bias may be particularly large.

How might we overcome social desirability bias and hypothetical bias in survey experiments? For social desirability bias, one option is the use of list experiments. None of the survey experiments included here are list experiments. More complex factorial designs such as conjoint experiments have also been shown to reduce social desirability bias (Hainmueller, Hopkins and Yamamoto 2014; Horiuchi, Markovich and Yamamoto 2018). For hypothetical bias, an option is to eschew hypothetical candidates in favor of real candidates. In fact, the only corruption survey experiment to date to use real candidates found a null effect on vote choice (De Figueiredo, Hidalgo and Kasahara 2011), and McDonald (2019) elicits smaller effects in survey experiments using the names of real politicians vs. a hypothetical politician. Of course, for corruption experiments this limits researchers to having actual information regarding the corrupt actions of candidates for ethical reasons.

Do field and survey experiments mirror real-world voting decisions?

Even if subjects (voters), treatments (information), and outcome (vote choice) are similar, contextual differences between survey and field experiments may also offer fundamentally different choice sets to voters. These discrepancies between survey and field experimental designs, as well as between the designs of different survey experiments, may alter respondents' potential outcomes and thus capture different estimands. Some possible contextual differences are discussed below.

Treatment strength, noncompliance, and declining salience

Informational treatments may be weaker in field experiments in part because of their method of delivery. Survey treatments tend to be clear and authoritative, and often provide information on the challenger (clean or corrupt). By contrast, many of the informational treatments utilized in past information and accountability field experiments—e.g. fliers and text messages—provide relatively weak one-time treatments that may even contain information subjects are already aware of. If the goal is to estimate real world effects, interventions should attempt to match those conducted in the real world (e.g. by campaigns, media, etc.). In fact, the natural experiment conducted by Ferraz and Finan (2008)—which takes advantage of random municipal corruption audits conducted by the Brazilian government—may provide evidence of the effectiveness of stronger treatments. The results of the audits were disseminated naturally by newspapers and political campaigns, and their study provides the largest estimated treatment effect amongst real-world experiments. While not measuring specific vote choice, past experiments using face-to-face canvassing contact have also demonstrated relatively large effects on voter turnout (Green and Gerber 2019; Kalla and Broockman 2018), but these methods have not been used in any information and accountability field experiments to date.

Treatment effects in field experiments (fliers, newspapers, etc.) may also be weaker in part because they can be missed by segments of the treatment group. More formally, survey experiments do not have noncompliance by design and therefore the average treatment effect (ATE) is equal to the intent-to-treat (ITT) effect,[13] whereas field experiments present ITT estimates as they are unable to identify which individuals in the treatment area actually received and internalized the informational treatment. Ideally, we would calculate the complier average causal effect (CACE)—the average treatment effect among the subset

[13]It could be argued that survey experiments have noncompliance if a respondent fails to absorb the information in the treatment. However, if there is also noncompliance in survey experiments, the CACE estimates would be even larger than the ITT estimates reported here, and the level of noncompliance in field experiments would need to be correspondingly larger to generate equal treatment effects. I thank an anonymous reviewer for this point.

of respondents who comply with treatment—in field experiments, but we are unfortunately unable to observe compliance in any of the corruption experiments conducted to date.

A theoretical demonstration shows how noncompliance can drastically alter the ITT. The ITT is defined as $ITT = CACE \times \pi_c$ where π_C indicates the proportion of compliers in the treatment group. When $\pi_C = 1$, $ITT = CACE = ATE$. If the ITT = -0.0033—as random effects meta-analysis estimates in field experiments—but only 10% of treated individuals "complied" with the treatment by reading the flier sent to them, this implies that the CACE is $\frac{-0.0033}{0.1} = -0.033$, or approximately -3 percentage points. In other words, while the effect of receiving a flier is roughly 0.3 percentage points, the effect of *reading* the flier is -3 percentage points. As the $ITT = CACE \times \pi_c$, any noncompliance necessarily reduces the size of the ITT. However, for the CACE to be equal in both survey and field experiments, the proportion of treatments that would need to remain undelivered in field experiments would have to be approximately 99% (i.e. 99% of subjects in the treatment group did not receive treatment or were already aware of the corruption information), implying that noncompliance likely does not tell the whole story.

Finally, treatments may be less salient at the time of vote choice in a field setting. Survey treatments are directly presented to respondents who are forced to immediately make a vote choice. Kalla and Broockman (2018) note that this mechanism manifests in campaign contact field experiments, where contact long before election day followed by immediate measurement of outcomes appears to persuade voters, whereas there is a null effect on vote choice on election day. Similarly, Sulitzeanu-Kenan, Dotan and Yair (Forthcoming) show that increasing the salience of corruption can increase electoral sanctioning, even without providing any new corruption information. Weaker treatments or lower salience of corruption in field experiments will weaken the treatment effect even amongst compliers (i.e. the CACE), further reducing the ITT.

Weak treatments, noncompliance, and declining treatment salience over time therefore make it unclear if the zero and null effects observed in field experiments stem from methodological choices or an actual lack of preference updating. Future field experiments should therefore consider using stronger treatments (e.g. canvassing), performing baseline surveys to measure subgroups amongst whom effects may be stronger, utilizing placebo-controlled designs that allow for measurement of noncompliance, and performing repeated measurement of outcome variables over time to capture declining salience.

Outcome choice

While vote choice is the outcome variable across all of the experiments investigated here, the choice set offered to voters is not necessarily always identical. Consider a voter's choice between two candidates in a field experiment conducted during an election. A candidate is revealed to be corrupt to voters in a treatment group, but not to voters in control. The treated voter can cast a ballot for corrupt candidate A, or candidate B, who may be clean or corrupt. The control voter can cast a ballot for candidate A or candidate B, and has no corruption information. Now consider a survey experiment with a vignette in which the randomized treatment is whether the corrupt actions of a politician are revealed or not. The treated voter can vote for the corrupt candidate A or not, but no challenger exists. Likewise, the control voter can vote for clean candidate A or not, but no challenger exists. Conjoint experiments overcome this difference, but the option to abstain still does not exist in the survey setting.[14] These differences in design offer fundamentally different choice sets to voters, altering respondents' potential outcomes and thus capturing different estimands.

Complexity, costliness, and conjoint experiments

Previous researchers have noted that even if voters generally find corruption distasteful, the quality of the information provided or positive candidate attributes and policies may outweigh the negative effects of corruption to voters, mitigating the effects of information provision on vote share.[15] These mitigating factors will naturally arise in a field setting, but may only be salient to respondents if specifically manipulated in a survey setting.

[14]See Eggers, Vivyan and Wagner (2018) and Agerberg (2020) for exceptions.
[15]See De Vries and Solaz (2017) for a comprehensive overview.

A number of survey experiments have therefore added factors other than corruption as mitigating variables, such as information quality, policy, economic benefit, and co-partisanship. Studies have randomized the quality of corruption information[16] (Banerjee et al. 2014; Botero et al. 2015; Breitenstein 2019; Mares and Visconti 2019; Weitz-Shapiro and Winters 2017; Winters and Weitz-Shapiro 2020), finding that lower quality information produces smaller negative treatment effects (see Figure 1.18). Policy stances in line with voter preferences have also been shown to mitigate the impact of corruption (Franchino and Zucchini 2015; Rundquist, Strom and Peters 1977). Evidence also suggests that respondents are more forgiving of corruption when it benefits them economically (Klašnja, Lupu and Tucker 2017; Winters and Weitz-Shapiro 2013). Evidence of co-partisanship as a limiting factor to corruption deterrence is mixed.[17] Boas, Hidalgo and Melo (2019) posit that abandoning dynastic candidates is particularly costly in Brazil. This evidence suggests that voters punish corruption less when it is costly to do so, and that these costly factors differ by country.

The fact that moderating variables may dampen the salience of corruption to voters has clearly not been lost on previous researchers. However, in the field setting numerous moderating factors may be salient to the voter. While there is likely no way to capture the complexity of real-world decision making in a survey setting, conjoint experiments allow researchers to randomize many candidate characteristics simultaneously, and thus have become a popular survey method for investigating the relative weights respondents give to different candidate attributes. In addition, conjoints force respondents to pick between two candidates, better emulating the choice required in an election. Finally, conjoints may minimize social desirability bias as they reduce the probability that the respondent is aware of

[16] For example, accusations from an independent anti-corruption authority may be deemed more credible than those from an opposition party, and accusations may be deemed less credible than a conviction.

[17] Anduiza, Gallego and Muñoz (2013), Agerberg (2020), and Breitenstein (2019) show that co-partisanship decreases the importance of corruption to Spanish respondents in survey experiments, and Solaz, De Vries and de Geus (2019) find that in-group membership reduces sanction of "corrupt" participants in a lab-experiment of UK subjects. However, Klašnja, Lupu and Tucker (2017) find relatively small effects of co-partisanship in Argentina, Chile, and Uruguay, Rundquist, Strom and Peters (1977) find null effects in a lab experiment in the US in the 1970s, and Konstantinidis and Xezonakis (2013) find no significant relationship in a survey experiment in Greece.

the researcher's primary experimental manipulation of interest (e.g. corruption).[18]

Researchers often present the results of conjoint experiments as average marginal component effects (AMCEs), then compare the magnitude of these effect sizes. AMCEs represent the unconditional marginal effect of an attribute (e.g. corruption) averaged over all possible values of the other attributes. This measurement is valuable, and crucially allows researchers to test multiple causal hypotheses and compare relative magnitudes of effects between treatments. However, this may or may not be a measure of substantive interest to the researcher, and implies that the AMCE is dependent on the joint distribution of the other attributes in the experiment.[19] These attributes are usually uniformly randomized. However, in the real world, candidate attributes are not uniformly distributed, so external validity is questionable. When we have a primary treatment of interest, such as corruption, we want to see how a "typical candidate" is punished for corruption. However a typical candidate is not a uniformly randomized candidate, but rather a candidate designed to appeal to voters. The corruption AMCE is therefore valid in the context of the experiment—marginalizing over the distribution of all other attributes in the experiment—but would likely be much smaller for a realistic candidate.[20] This implies that AMCEs have more external validity when the joint distribution of attributes matches the real world and the experiment contains the entire universe of possible attributes.[21]

When researchers have strong theories about the conditions that shape voter decision-making, a more appropriate method may be to calculate average marginal effects in order

[18]This is explicitly mentioned by Hainmueller, Hopkins and Yamamoto (2014), who argue that conjoint experiments give respondents "various attributes and thus [they] can often find multiple justifications for a given choice." Note, however, that an experiment does not necessarily need to be a conjoint design to have this feature. Conjoint experiments encourage researchers to randomize more attributes and therefore typically contain more complex hypothetical vignettes. However, the same vignette complexity could be achieved without full randomization of these attributes.

[19]See De la Cuesta, Egami and Imai (2019) for additional discussion and empirical demonstration of the impact of choice of distribution on the AMCE.

[20]Abramson, Koçak and Magazinnik (2019) also point out that the AMCE represents a weighted average of both intensity and direction. It is therefore important to interpret conjoint results in terms of both intensity and direction of preferences.

[21]The uniform distribution may be reasonable when we are not attempting emulate real-world appearances of attributes—for example to find an optimal policy design from a menu of equally possible options.

to present predicted probabilities of voting for a candidate under these conditions.[22] For example, in a conjoint experiment including corruption information, the probability of voting for a candidate who is both corrupt and possesses other particular feature levels (e.g. party membership and/or policy positions), marginalizing across all other features in the experiment.[23]

To illustrate this point, I replicate the conjoint experiments conducted in Spain by Breitenstein (2019) and in Italy by Franchino and Zucchini (2015), and present both AMCEs and predicted probabilities. The Breitenstein (2019) re-analysis is presented in the main text, while the re-analysis of Franchino and Zucchini (2015) is in the appendix.[24] Note that I group all corruption accusation levels into a single "corrupt" level in my replications. The Breitenstein (2019) predicted probabilities are presented as a function of corruption, co-partisanship, political experience, and economic performance. The charts therefore show the probability of preferring a candidate who is always corrupt, but is a co-partisan or not, has low or high experience, and whose district experienced good or bad economic performance, marginalizing across all other features in the experiment. For Franchino and Zucchini (2015), the predicted probabilities are presented as a function of corruption and two policy positions—tax policy and same sex marriage—separately for conservative and liberal respondents. The charts therefore show the probability of preferring a candidate who is corrupt, but has particular levels of tax and same sex marriage policy, marginalizing across all other features in the experiment. Note that Franchino and Zucchini (2015) correctly conclude that their typical "respondent prefers a corrupt but socially and economically progressive candidate to a clean but conservative one," and Breitenstein (2019) presents certain predicted

[22]This method is utilized by Teele, Kalla and Rosenbluth (2018) to examine the probability of voting for female or male candidates holding other candidate attributes (marital status and number of children) constant, and in corruption experiments by Agerberg (2020), Breitenstein (2019), and Chauchard, Klašnja and Harish (2019). This method is discussed in more detail by Leeper, Hobolt and Tilley (2019).

[23]Note that standard errors will increase as a result of conditioning on certain combinations of attributes. However, this can be avoided by utilizing an experimental design that conditions on these features at the design stage.

[24]Additional predicted probability replications from Mares and Visconti (2019) and Chauchard, Klašnja and Harish (2019) can also be found in the appendix.

probabilities. While I therefore illustrate how predicted probabilities can be used to draw conclusions that may be masked by examination of AMCEs alone, the authors themselves do not make this mistake. I perform the same analysis including only cases where the challenger is clean in the appendix.

A casual interpretation of the traditional AMCE plots presented in Figure 1.3 and Figure 1.22 suggests that it is very unlikely a corrupt candidate would be chosen by a respondent. By contrast, the predicted probabilities plots presented in Figure 1.4, Figure 1.23, and Figure 1.24 show that even for corrupt candidates in the conjoint, the right candidate or policy platform presented to the right respondents can garner over 50% of the predicted hypothetical vote.[25] Further, the attributes included in these conjoints surely do not represent all candidate attributes relevant to voters, and indeed differ greatly across experiments. As in Agerberg (2020), the level of support for corrupt candidates also varies based on whether or not the challenger is clean (Figure 1.19, Figure 1.25, and Figure 1.26). In other words, respondents find it costly to abandon their preferences even if it forces them to select a corrupt candidate, and this costliness varies highly depending on contextual changes and choice of other attributes included in the experiments.

[25] Note that a negative corruption treatment effect is still present. See Figure 1.20 for a visual depiction of predicted probabilities for both a corrupt and clean candidate. The difference between the point estimates for the corrupt and clean candidate can be interpreted as a treatment effect. I thank an anonymous reviewer for suggesting this clarification.

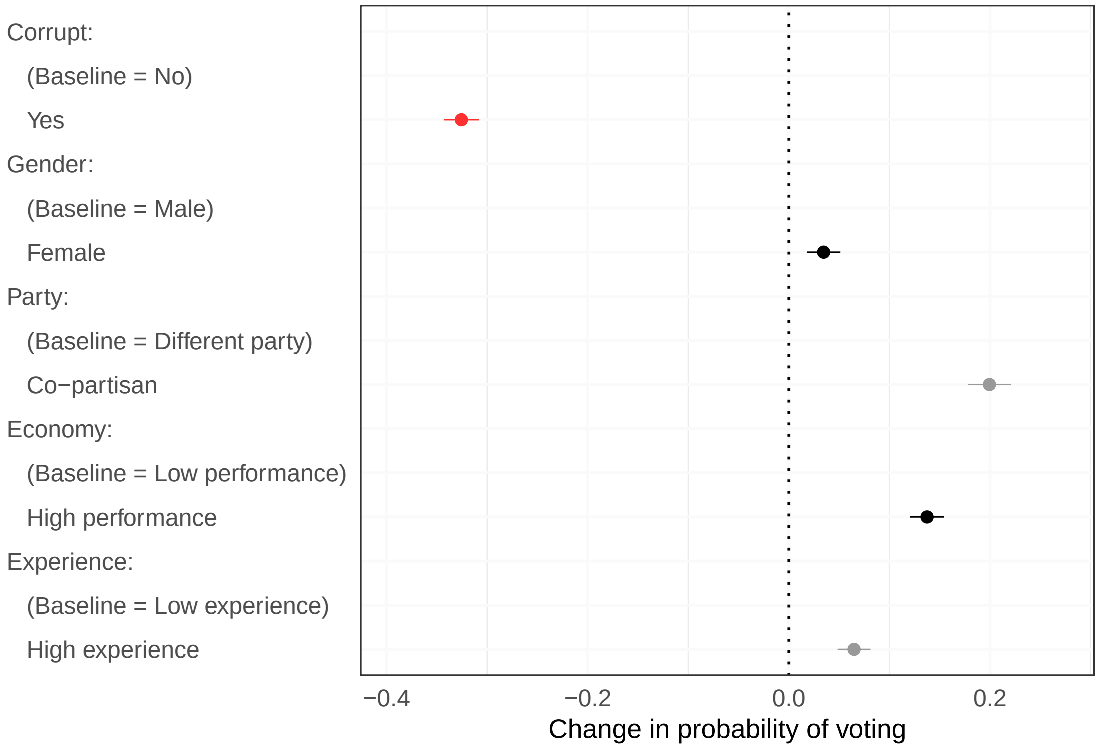


Figure 1.3: Breitenstein (2019) conjoint: average marginal component effects

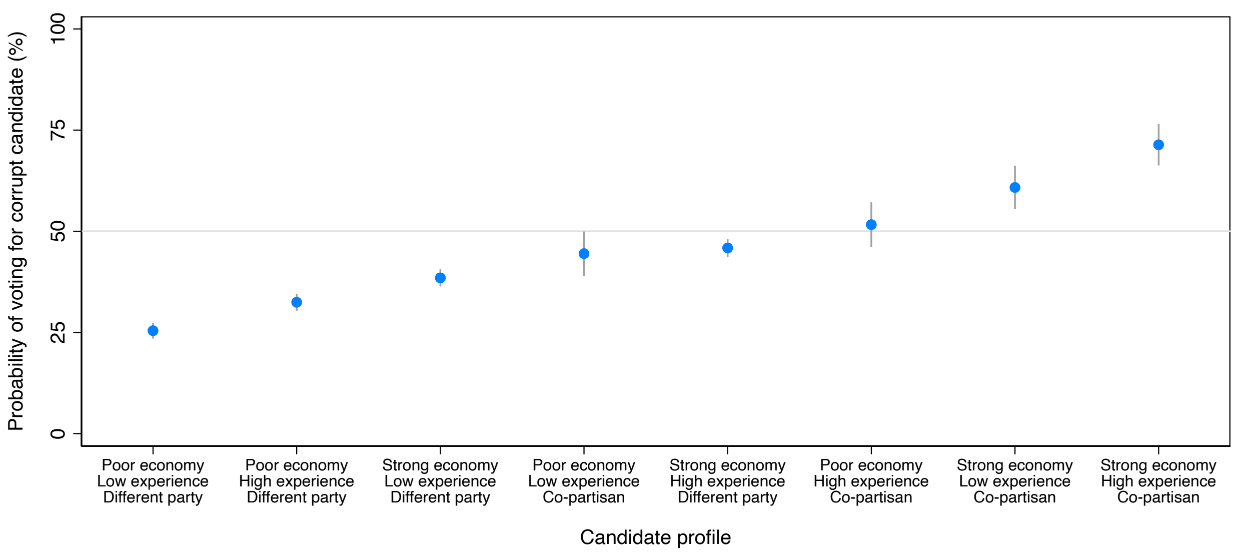


Figure 1.4: Breitenstein (2019) conjoint: can the right candidate overcome corruption?

Candidate or policy profiles that result in over 50% of voters selecting a corrupt candidate may not be outliers in real-world scenarios. Unlike in conjoint experiments, real-world candidates' attributes and policy profiles are not selected randomly, but rather represent choices designed to appeal to voters. Voters may also be unsure if the challenger is also corrupt or clean. It may therefore be preferable to analyze conjoint experiments as above, comparing outlier characteristics (e.g. corruption) to realistic candidate profiles that target specific voters, rather than fully randomized candidate profiles.

When the most theoretically relevant tradeoffs are unclear, we may be able to illuminate voter decision making processes through the use of decision trees.[26] The decision tree in Figure 1.5 was trained using all randomized variables in the Breitenstein (2019) conjoint, and the tree was pruned to minimize cross-validated classification error rate. Figure 1.5 draws similar conclusions as the predicted probabilities chart shown in Figure 1.4 with respect to what factors matter most to voters. A similar figure depicting corrupt candidates facing clean challengers only can be found in Figure 1.21.

[26]Decision trees offer a parsimonious way to model fundamental non-linearities in the conjoint data and will typically have lower bias than an OLS-based predicted probability estimator, but may exhibit higher variance.

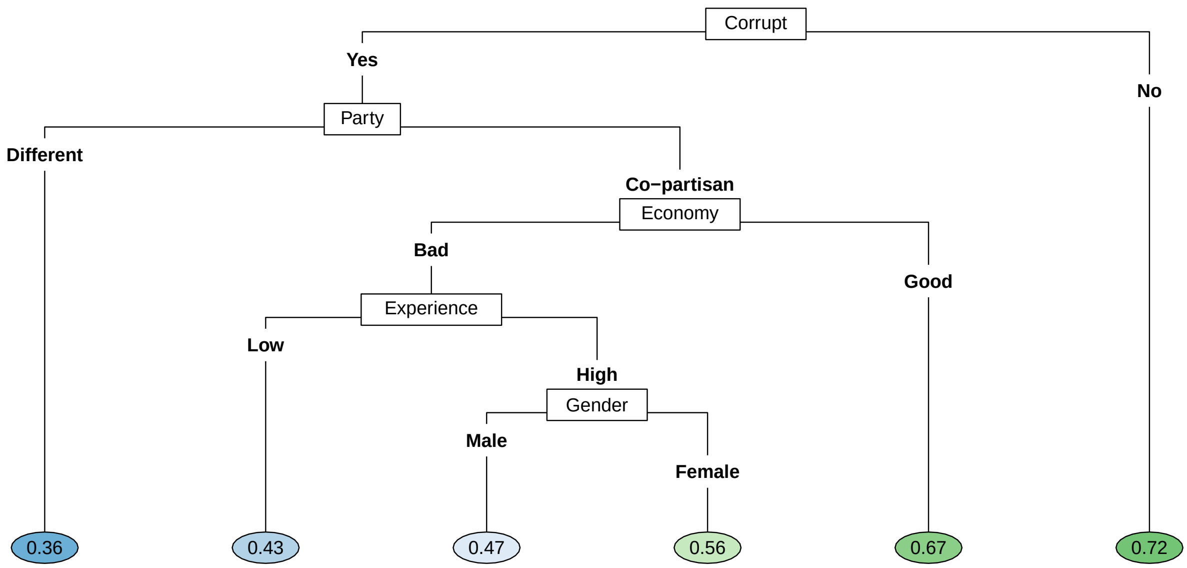


Figure 1.5: Breitenstein (2019) conjoint decision tree: predicted probabilities of voting for candidate

Discussion

The field experimental results reported here align with a growing body of literature that shows minimal effects of information provision on voting outcomes. The primary conclusion of the Metaketa I project—which sought to determine if politicians were rewarded for positive information and punished for negative information—was that "the overall effect of information [provision] is quite precisely estimated and not statistically distinguishable from zero" (Dunning, Grossman, Humphreys, Hyde, McIntosh and Nellis 2019), and a meta-analysis by Kalla and Broockman (2018) suggests that the effect of campaign contact and advertising on voting outcomes in the United States is close to zero in general elections.

However, we should be careful not to conclude that voters never punish politicians for malfeasance from these experiments, or that field experiments recover truth. Field and natural experiments in other domains have found effects when identifying persuadable voters prior to treatment delivery (Kalla and Broockman 2018; Rogers and Nickerson 2013), or when using higher dosage treatments (Adida, Gottlieb, Kramon and McClendon 2019; Ferraz and Finan 2008).[27] Combining stronger treatments, measurement of noncompliance, and pre-identification of subgroups most susceptible to persuasion should therefore be a goal of future field experiments.

Many of the survey experimental studies discuss how their findings may partially stem from the particular conditions of the experiment, claim that they are only attempting to identify tradeoffs or moderating effects, and/or acknowledge the limitations of external validity. However, other studies do not. A common approach is to cite Hainmueller, Hangartner and Yamamoto (2015), who show similar effects in a vignette, conjoint, and natural experiment. However, Hainmueller, Hangartner and Yamamoto (2015) use closeness in the magnitude of treatment effects between vignettes and the natural experiment as a justification for correspondence between the two methodologies. Their study therefore suggests that the relative importance *and magnitude* of treatment effects should be similar between hypothetical vignettes and the real world, which this meta-analysis shows is not the case with corruption voting. Further, the natural experimental benchmark takes the form of a survey/leaflet sent to voters containing the attributes of immigrants applying for naturalization in Swiss municipalities. The conjoint experiment is therefore able to perfectly mimic the amount of information voters posses in the real world, which is not the case for political candidates.[28] We should therefore be cautious when extrapolating the correspondence between these studies to cases such as candidate choice experiments.

[27] While an observational study, Chang, Golden and Hill (2010) also points to the effectiveness of higher dosage treatments.

[28] Hainmueller, Hangartner and Yamamoto (2015) acknowledge this directly, stating that "these data provide an ideal behavioral benchmark to evaluate stated preference experiments, because they closely resemble a real-world vignette experiment" and that "unlike many other real-world choice situations, in the referendums, the information environment and choice attributes are sufficiently constrained, such that they can be accurately mimicked in a survey experimental design."

Conclusion

In an effort to test whether voters adequately hold politicians accountable for malfeasance, researchers have turned to experimental methods to measure the causal effect of learning about politician corruption on vote choice. A meta-analytic assessment of these experiments reveals that conclusions differ drastically depending on whether the experiment was deployed in the field and monitored actual vote choice, versus hypothetical vote choice in a survey setting. Across field experiments, the aggregate treatment effect of providing information about corruption on vote share is approximately zero. By contrast, in survey experiments corrupt candidates are punished by respondents by approximately 32 percentage points.

I explore publication bias, social desirability bias, and contextual differences in the nature of the experimental designs as possible explanations for the discrepancy between field and survey experimental results. I do not find systematic evidence of publication bias. Social desirability bias may drive some of the difference if survey experiments cause respondents to under-report socially undesirable behavior, and hypothetical bias may cause respondents to not properly internalize the costs of switching their votes. The survey setting may differ from the field due to contextual differences such as noncompliance, treatment strength, differences in outcome choice sets, and costliness/decision complexity. Noncompliance necessarily decreases treatment effect sizes in field experiments. Weak treatments or lower salience of information to voters on election day versus immediately after treatment receipt will also reduce effect sizes. Previous survey experiments have also shown that treatment effects diminish as the costliness of changing one's vote increases, and these costs are likely to be much higher and more multitudinous in an actual election. The personal cost of changing one's vote may therefore be higher than accepting corruption in many real elections, but not in surveys.

High-dimension factorial designs such as conjoint experiments may better capture the costly tradeoffs voters make in the survey setting. However, it may be preferable to analyze candidate choice conjoint experiments by comparing the probability of voting for a realistic candidate with outlier characteristics (e.g. corruption) to the probability of voting for the same realistic candidate without this characteristic, rather than examining differences in AMCEs across fully randomized candidate profiles.

These findings suggest that while candidate choice survey experiments may provide information on the directionality of informational treatments in hypothetical scenarios, the point estimates they provide may not be representative of real-world voting behavior. More generally, researchers should exercise caution when interpreting actions taken in hypothetical vignettes as indicative of real world behavior such as voting. However, we should also be careful not to conclude that field experiments always recover generalizable truth due to design decisions and limitations.

Appendix

Lab experiments

Table 1.3: Lab experiments

Study	Country	ATE
Arvate and Mittlaender (2017)	Brazil	Negative
Azfar and Nelson (2007)	USA	Negative
Rundquist, Strom and Peters (1977)[1]	USA	Negative
Solaz, De Vries and de Geus (2019)	UK	Negative

[1] The candidate is always corrupt in the Rundquist, Strom and Peters (1977) experiment. A "corruption" point estimate is therefore not provided in the coefficient plot below.

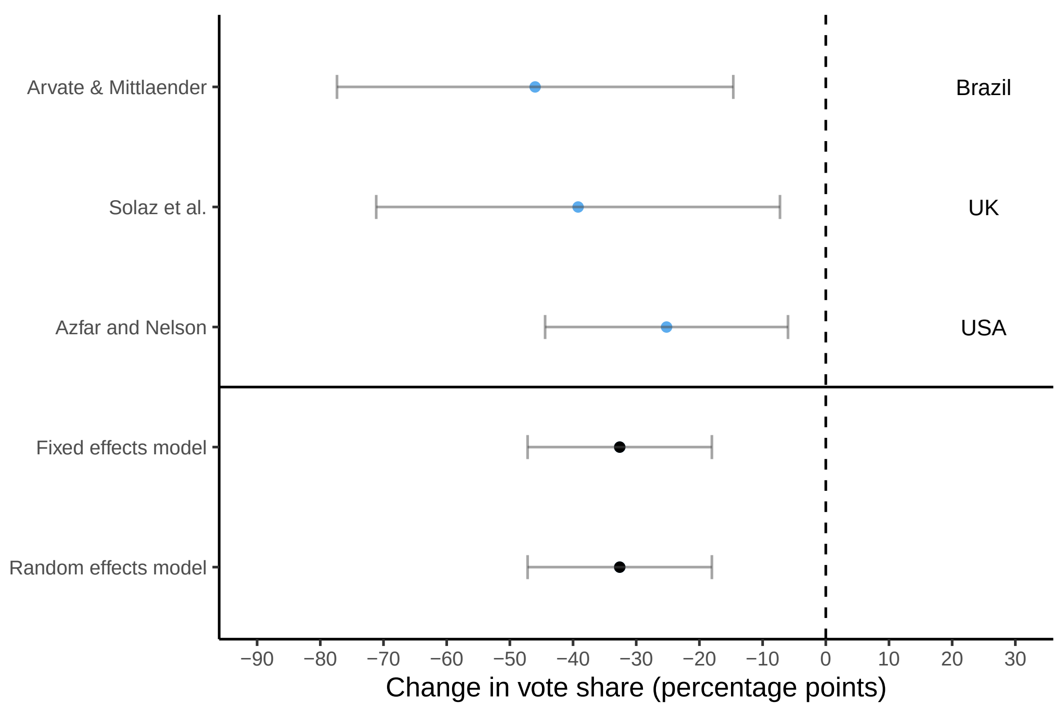


Figure 1.6: Lab experiments: Average treatment effect of corruption information on vote share

.

Excluded studies

Table 1.4: Excluded experiments

Study	Type	Reason for exclusion
Anduiza, Gallego and Muñoz (2013)	Survey	Lack of no-corruption control group
Botero et al. (2015)	Survey	Lack of no-corruption control group
De Figueiredo, Hidalgo and Kasahara (2011)	Survey	Outcome is hypothetically changing actual vote
Green, Zelizer, Kirby et al. (2018)	Field	Outcome is favorability rating, not vote share
Konstantinidis and Xezonakis (2013)	Survey	Lack of no-corruption control group
Muñoz, Anduiza and Gallego (2012)	Survey	Lack of no-corruption control group
Rundquist, Strom and Peters (1977)	Lab	Lack of no-corruption control group
Winters and Weitz-Shapiro (2016)	Survey	Data identical to Weitz-Shapiro and Winters (2017)
Winters and Weitz-Shapiro (2015)	Survey	Data identical to Winters and Weitz-Shapiro (2013)
Weschle (2016)	Survey	Lack of no-corruption control group

Meta-analysis and heterogeneity by type of experiment

Table 1.5: Meta-analysis by type of experiment

Value	Estimate	95% CI
Field: weighted fixed effects	-0.002	-0.006 to 0.001
	(0.002)	
Field: random effects	-0.003	-0.021 to 0.014
	(0.009)	
Survey: weighted fixed effects	-0.319	-0.326 to -0.312
	(0.004)	
Survey: random effects	-0.322	-0.382 to -0.262
	(0.031)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.6: Random effects meta-analysis (all studies)

Value	Estimate	95% CI
Estimate	-0.21	-0.279 to -0.141
	(0.035)	
Estimated total heterogeneity	0.034	0.016 to 0.053
	(0.009)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.7: Mixed effects meta-analysis with survey experiment moderator

Value	Estimate	95% CI
Constant	-0.007	-0.074 to 0.06
	(0.034)	
Survey experiment moderator	-0.315	-0.398 to -0.232
	(0.043)	
Residual heterogenity with moderator	0.011	0.005 to 0.017
	(0.003)	
Heterogenity accounted for	68.023%	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

"Heterogeneity accounted for" is calculated as: $\dfrac{(\text{Total heterogeneity} - \text{Residual heterogeneity})}{(\text{Total heterogeneity})}$

Robustness checks

Table 1.8: Meta-analysis (all field experiments excluding Banerjee et al. (2010) and Banerjee et al. (2011))

Value	Estimate	95% CI
Field: weighted fixed effects	-0.002	-0.006 to 0.002
	(0.002)	
Field: random effects	-0.003	-0.022 to 0.016
	(0.01)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.9: Random effects meta-analysis (all studies excluding Banerjee et al. (2010) and Banerjee et al. (2011))

Value	Estimate	95% CI
Estimate	-0.226	-0.296 to -0.155
	(0.036)	
Estimated total heterogeneity	0.033	0.015 to 0.052
	(0.01)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.10: Mixed effects meta-analysis with survey experiment moderator (excluding Banerjee et al. (2010) and Banerjee et al. (2011))

Value	Estimate	95% CI
Constant	-0.009	-0.086 to 0.067
	(0.039)	
Survey experiment moderator	-0.313	-0.405 to -0.221
	(0.047)	
Residual heterogenity with moderator	0.012	0.005 to 0.019
	(0.004)	
Heterogenity accounted for	64.751%	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.11: Meta-analysis (all survey experiments including De Figueiredo, Hidalgo and Kasahara (2011)

Value	Estimate	95% CI
Survey: weighted fixed effects	-0.317 (0.004)	-0.324 to -0.31
Survey: random effects	-0.305 (0.034)	-0.371 to -0.239

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

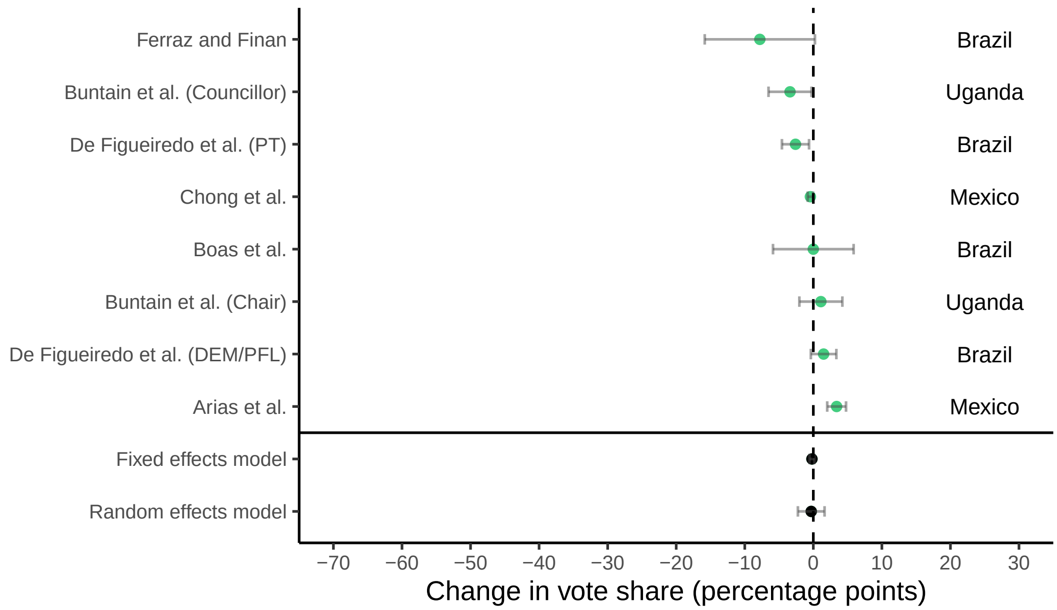


Figure 1.7: Field experiments: Average treatment effect of corruption information on incumbent vote share (excluding Banerjee et al. (2010) and Banerjee et al. (2011))

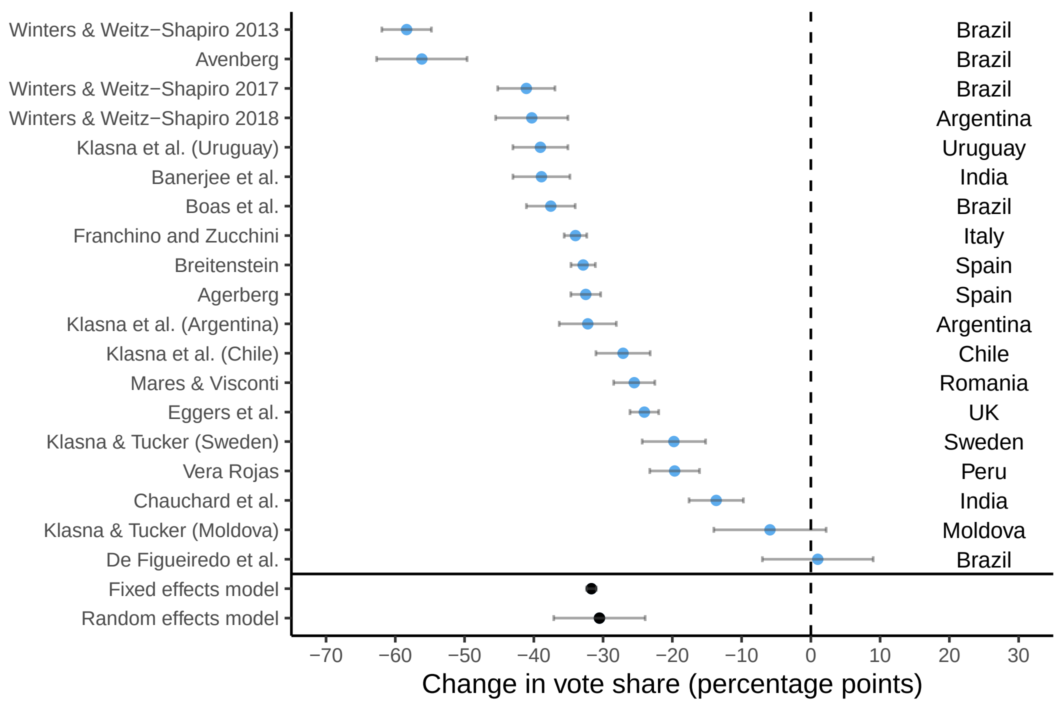


Figure 1.8: Survey experiments: Average treatment effect of corruption information on incumbent vote share (including De Figueiredo, Hidalgo and Kasahara (2011))

.

Publication bias

Table 1.12: P-values by study

Study	Experiment Type	Published	Reported p-value	Replicated p-value
Winters and Weitz-Shapiro 2013	Survey	Yes	0.000	0.000
Avenberg	Survey	Yes	0.000	
Winters and Weitz-Shapiro 2017	Survey	Yes	0.000	0.000
Winters and Weitz-Shapiro 2018	Survey	Yes	0.000	0.000
Klasna et al. (Uruguay)	Survey	No	0.000	0.000
Banerjee et al.	Survey	Yes	0.000	
Boas et al.	Survey	Yes	0.000	0.000
Franchino and Zucchini	Survey	Yes	0.000	0.000
Breitenstein	Survey	Yes	0.000	0.000
Agerberg	Survey	Yes	0.000	
Klasna et al. (Argentina)	Survey	No	0.000	0.000
Klasna et al. (Chile)	Survey	No	0.000	0.000
Mares and Visconti	Survey	Yes	0.000	0.000
Eggers et al.	Survey	Yes	0.000	0.000
Klasna and Tucker (Sweden)	Survey	Yes	0.000	
Vera Rojas	Survey	Yes	0.000	
Chauchard et al.	Survey	Yes	0.000	0.000
Arias et al.	Field	Yes	0.000	
De Figueiredo et al. (PT)	Field	No	0.011	
Chong et al.	Field	Yes	0.032	
Buntain et al. (Councillor)	Field	Yes	0.034	
Ferraz and Finan	Natural	Yes	0.058	
De Figueiredo et al. (DEM/PFL)	Field	No	0.116	
Klasna and Tucker (Moldova)	Survey	Yes	0.155	
Banerjee et al. (2011)	Field	No	0.268	
Banerjee et al. (2010)	Field	No	0.708	
Buntain et al. (Chair)	Field	Yes	0.754	
Boas et al.	Field	Yes	1.000	

Notes: Publication status as of November 2019. All p-values rounded to the nearest thousandth decimal place. Reported p-value is the p-value associated with the corruption ATE directly reported in the paper if available. If not available, p-values are reconstructed from point estimates, standard errors, and sample size in regression tables. Replicated p-values are shown for all studies which were fully replicated.

Table 1.13: Do p-values predict publication status?

	Dependent variable:	
	Published	
	OLS	Logit
Reference: P less than 0.01	0.80***	1.39**
	(0.12)	(0.65)
P less than 0.05	−0.13	−0.69
	(0.29)	(1.38)
P less than 0.1	0.20	15.18
	(0.47)	(2, 399.54)
P greater than 0.1	−0.13	−0.69
	(0.19)	(0.96)
Observations	28	28
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 1.14: Regression tests for funnel plot asymmetry

Studies included	p value
All	0.0004
All with moderator	0.923
Field	0.954
Survey	0.862

Table 1.15: Trim and fill estimates by subgroup

Value	Estimate	95% CI
All experiments: random effects	-0.237	-0.307 to -0.168
	(0.035)	
Field: random effects	-0.003	-0.021 to 0.014
	(0.009)	
Survey: random effects	-0.322	-0.382 to -0.262
	(0.031)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

Table 1.16: PET-PEESE estimates by subgroup

Value	Estimate	95% CI
All experiments	0.008	-0.045 to 0.062
	(0.027)	
Field experiments	-0.002	-0.013 to 0.009
	(0.006)	
Survey experiments	-0.317	-0.38 to -0.254
	(0.032)	

Note: Standard errors in parenthesis. Figures rounded to nearest thousandth decimal place.

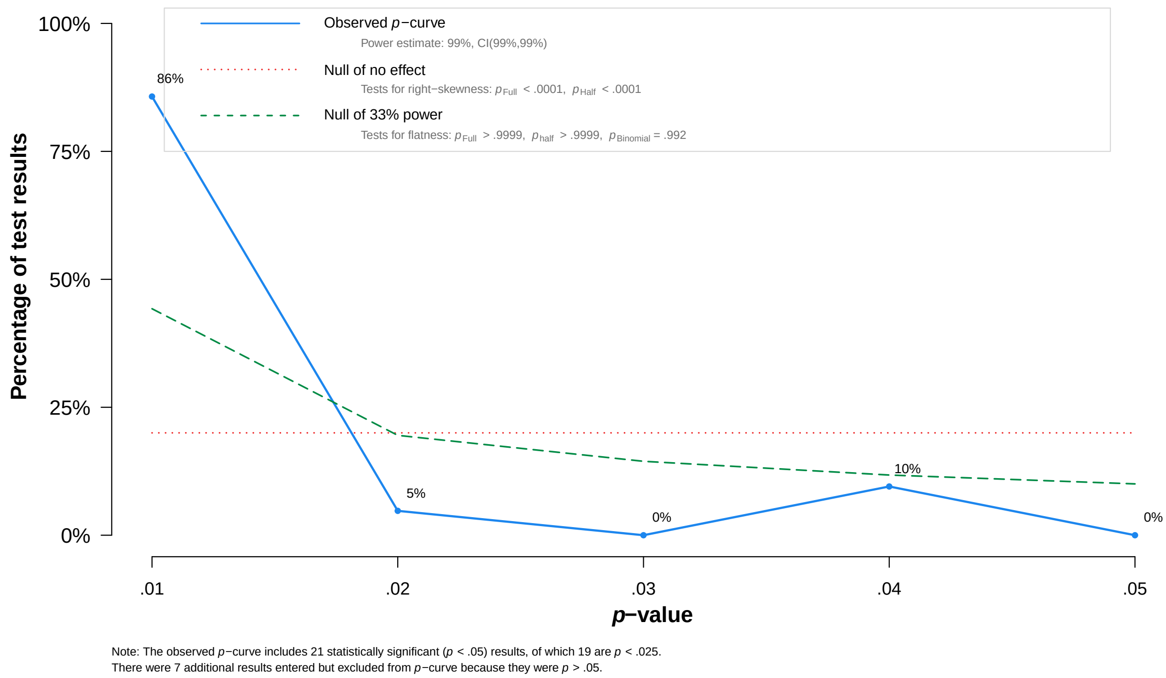


Figure 1.9: P-curve: all experiments

.

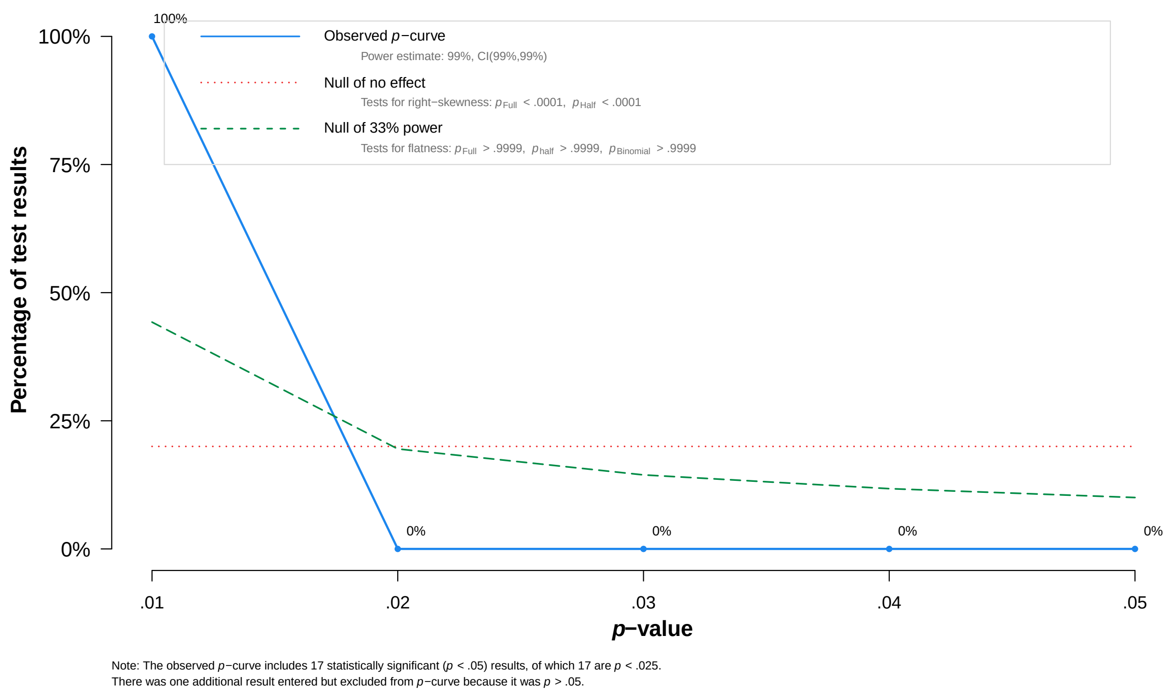


Figure 1.10: P-curve: survey experiments

.

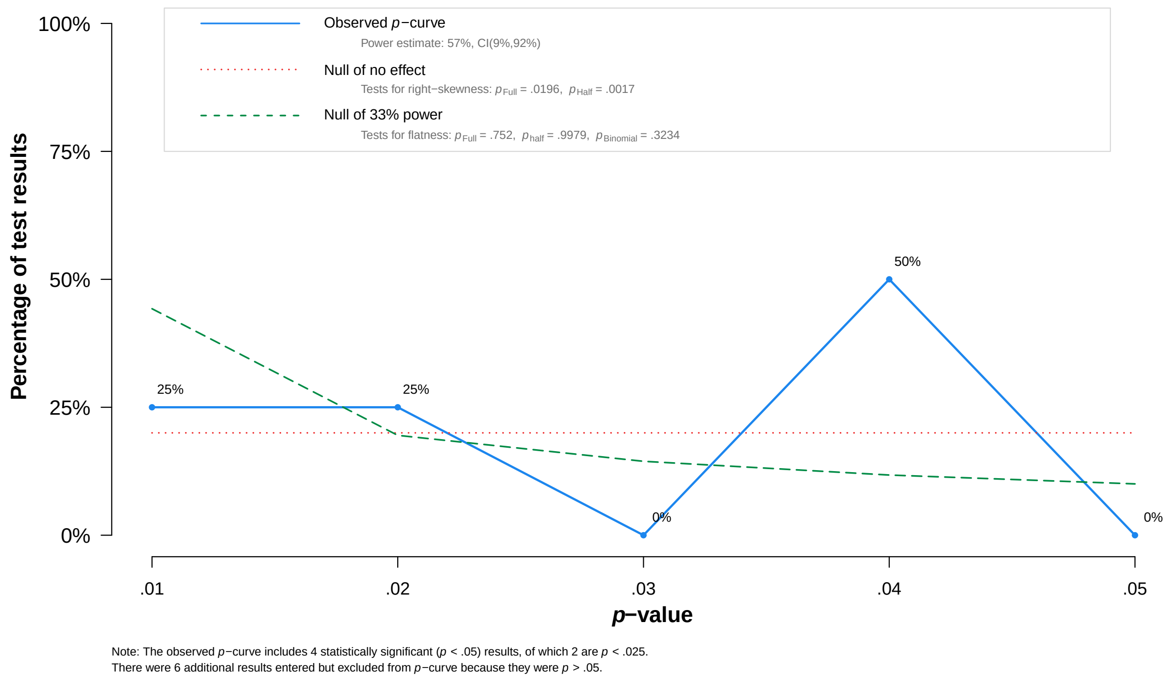


Figure 1.11: P-curve: field experiments

.

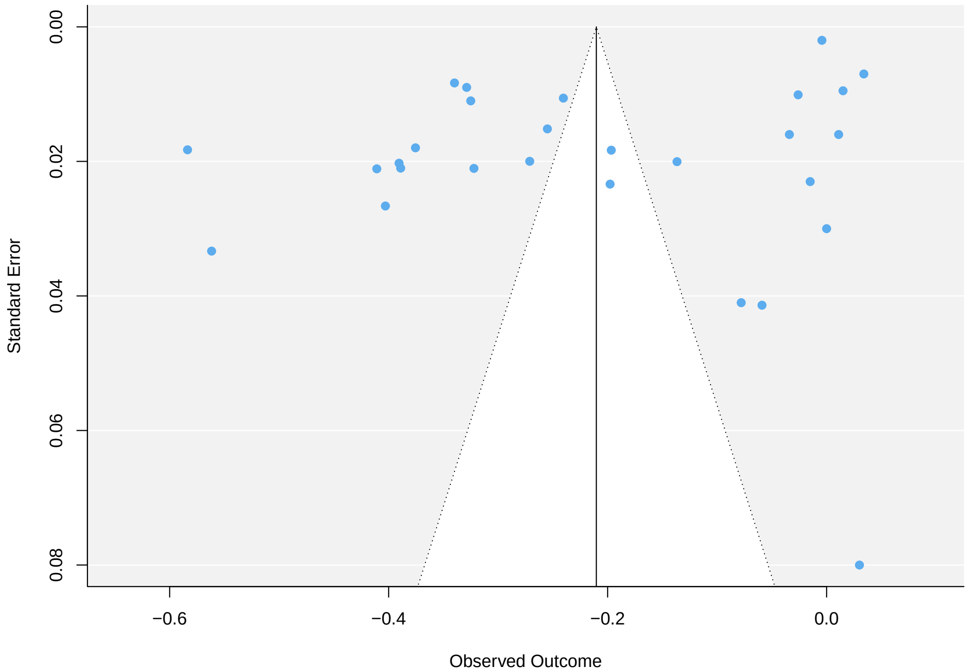


Figure 1.12: Funnel plot: all experiments

.

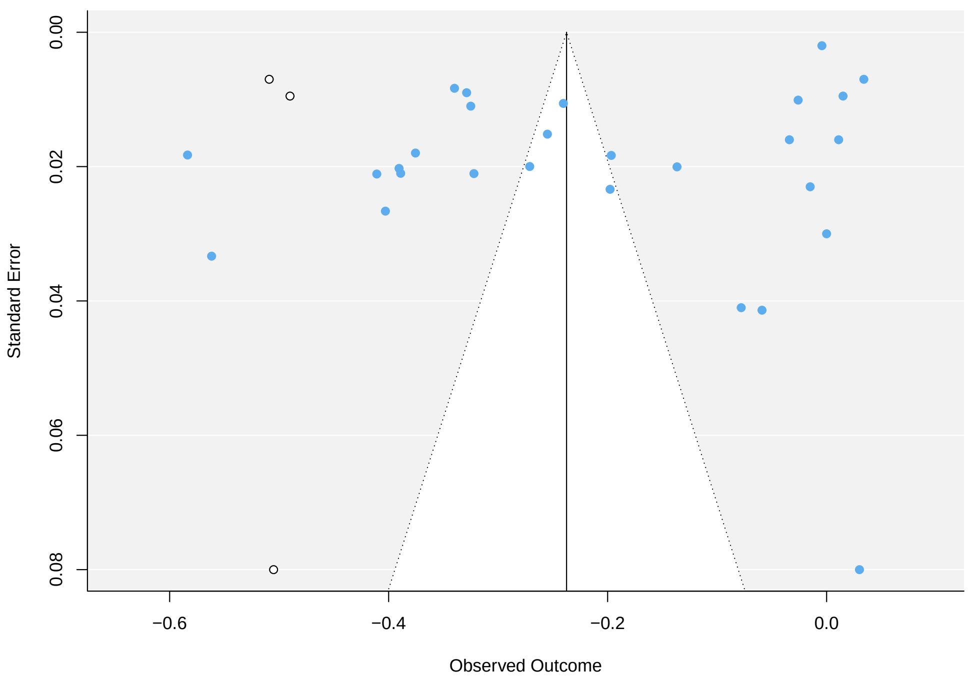


Figure 1.13: Funnel plot including trim and fill "missing" studies: all experiments

Note: Actual studies in blue and estimated missing studies in white.

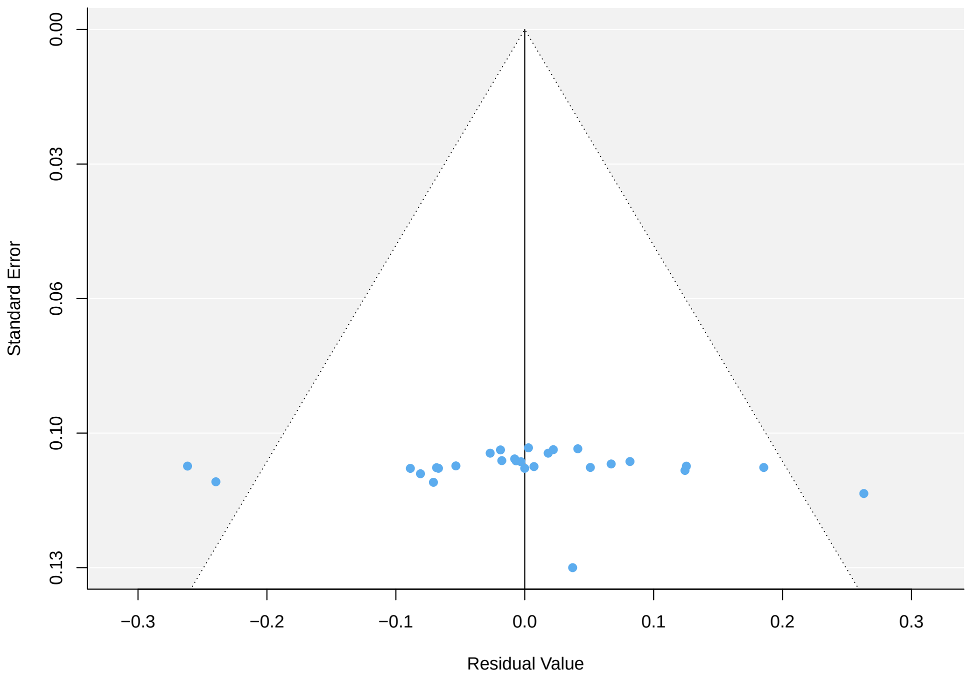


Figure 1.14: Funnel plot: all experiments with field experiment moderator

.

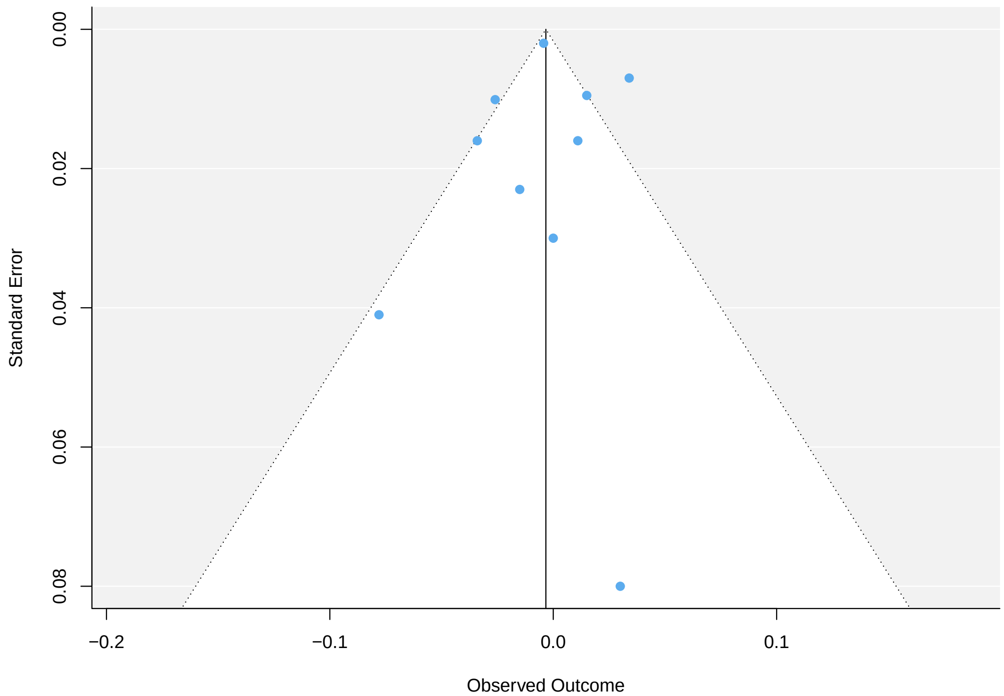


Figure 1.15: Funnel plot: field experiments

.

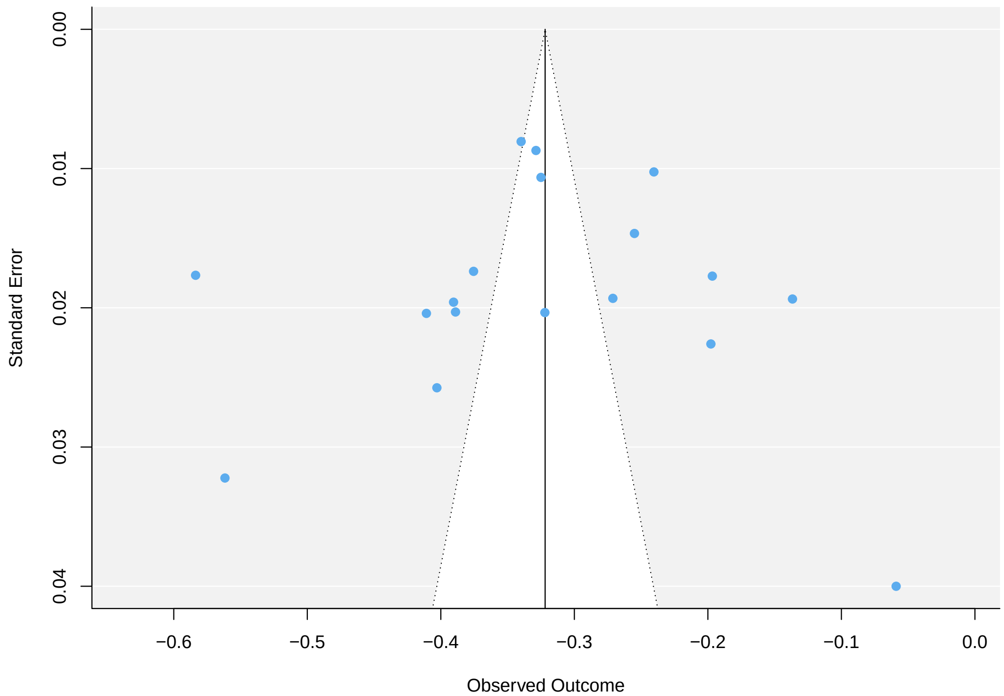


Figure 1.16: Funnel plot: survey experiments

.

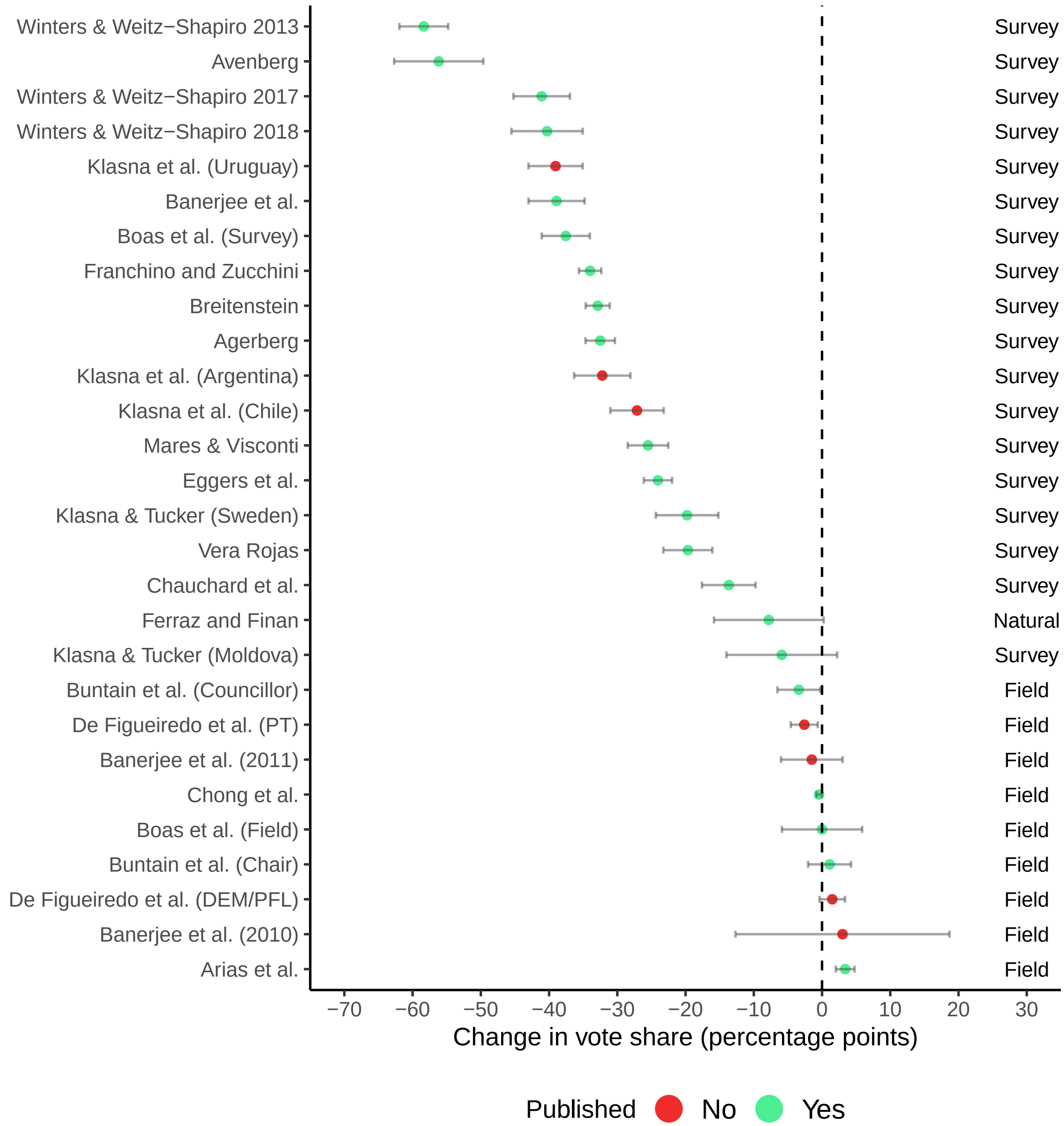


Figure 1.17: All experiments by publication status: Average treatment effect of corruption information on vote share and 95% confidence intervals

.

Information quality

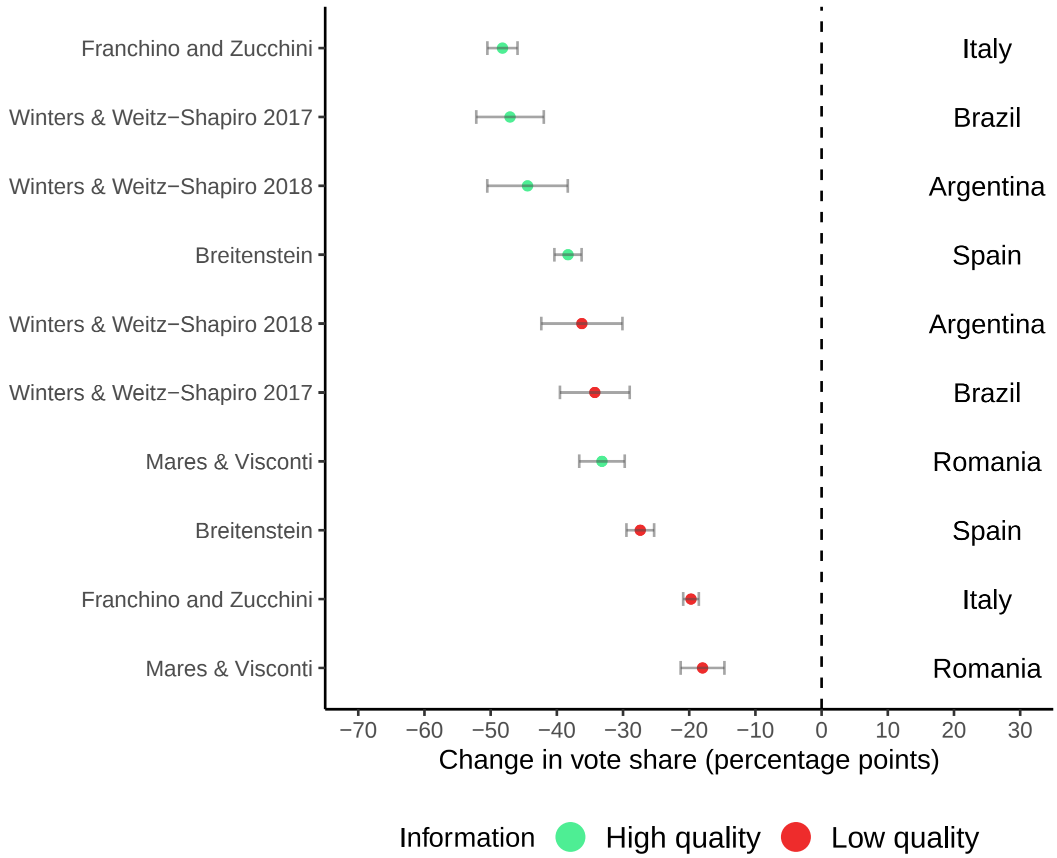


Figure 1.18: Survey experiments by information quality: Average treatment effect of corruption information on vote share and 95% confidence intervals

.

Additional conjoint replications using predicted probabilities

Breitenstein (2019)

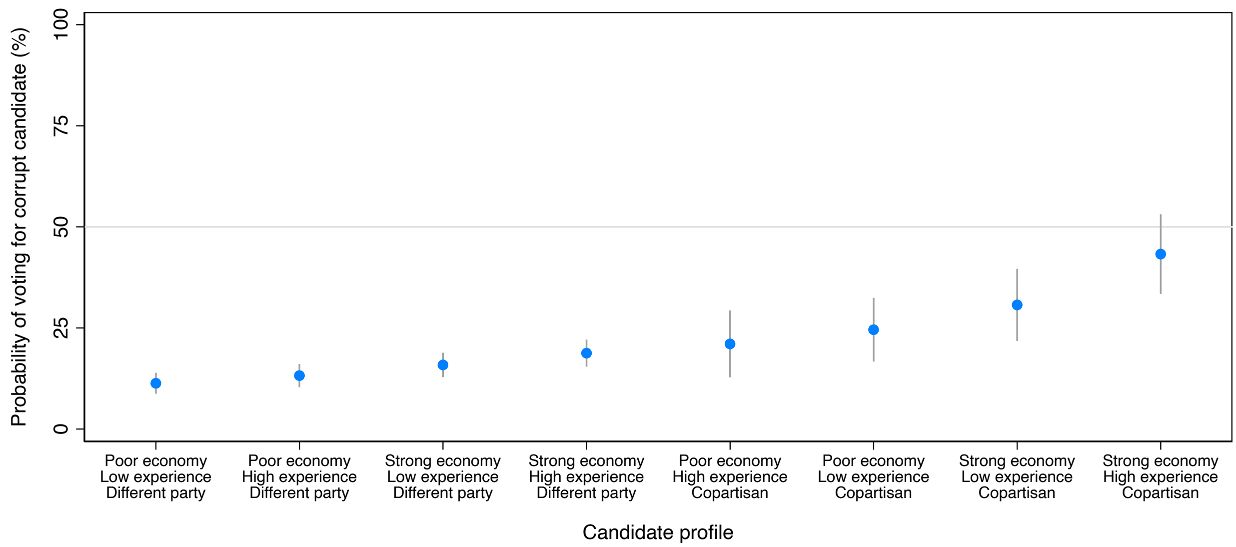


Figure 1.19: Breitenstein (2019) conjoint: can the right candidate overcome corruption (clean challenger)?

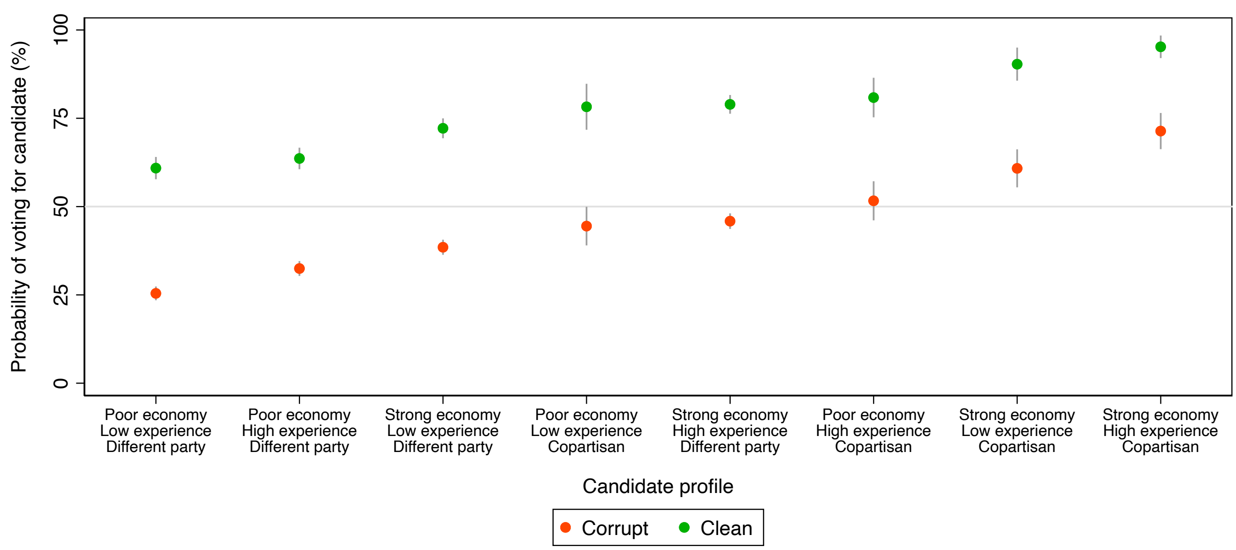


Figure 1.20: Breitenstein (2019) conjoint: probability of choosing candidate (by clean or corrupt)

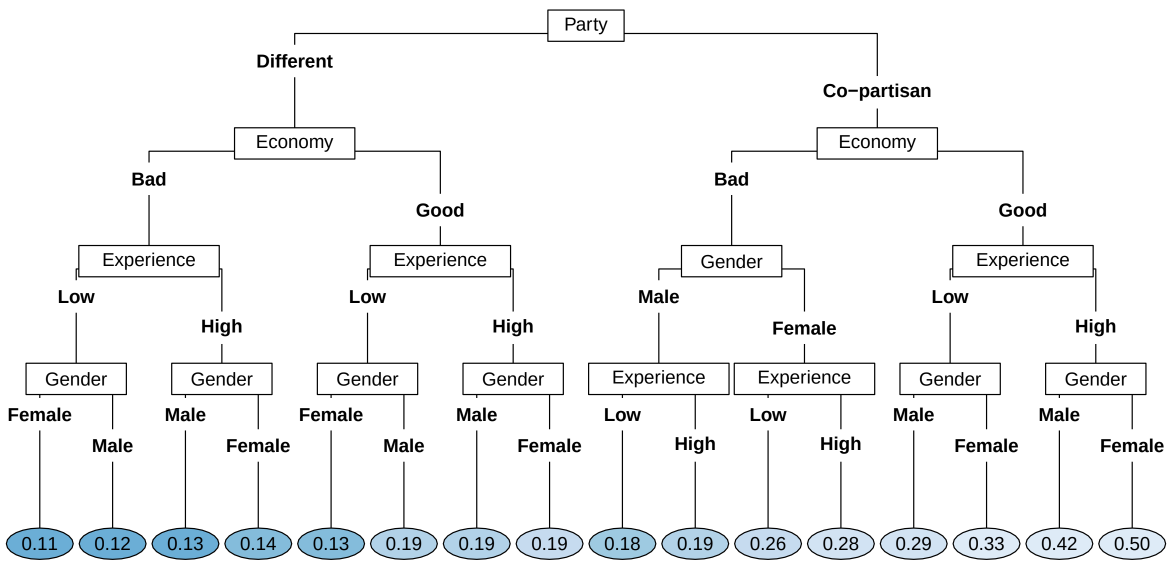


Figure 1.21: Breitenstein (2019) conjoint decision tree: predicted probabilities of voting for corrupt politician with clean challenger

Franchino and Zucchini (2015)

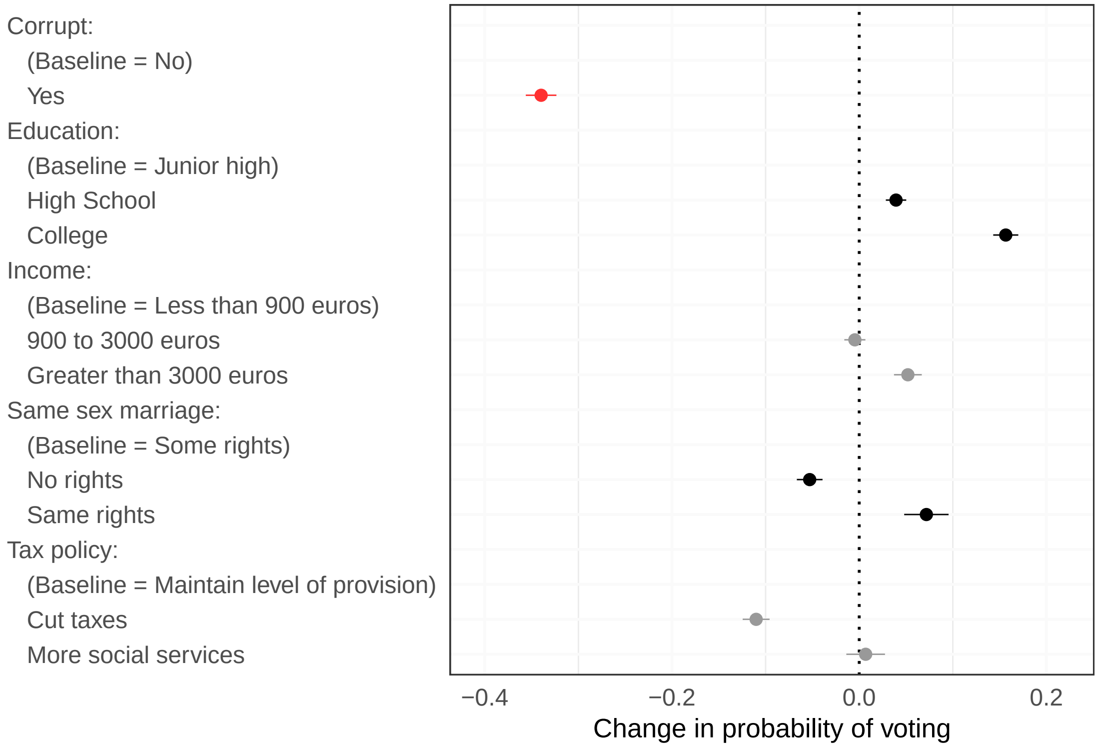


Figure 1.22: Franchino and Zucchini (2015) conjoint: AMCEs

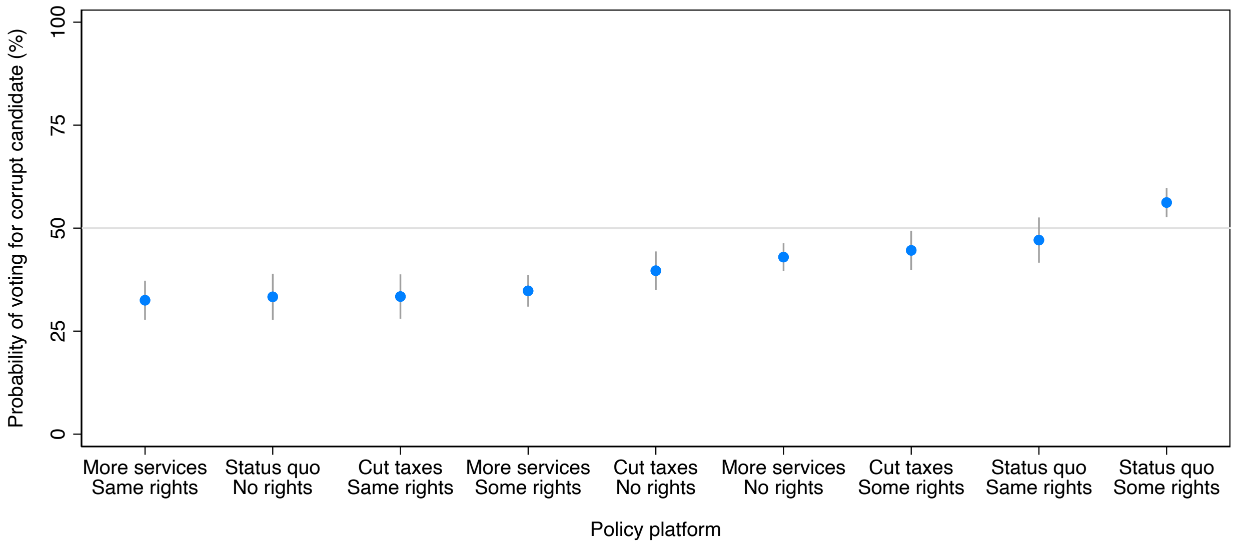


Figure 1.23: Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (conservative respondents)?

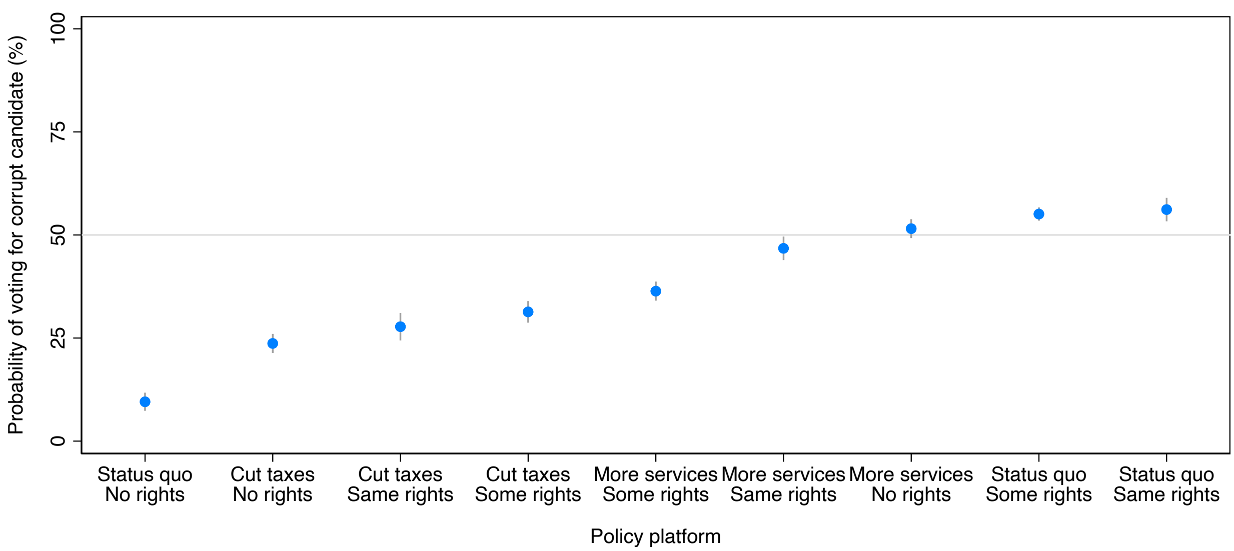


Figure 1.24: Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (liberal respondents)?

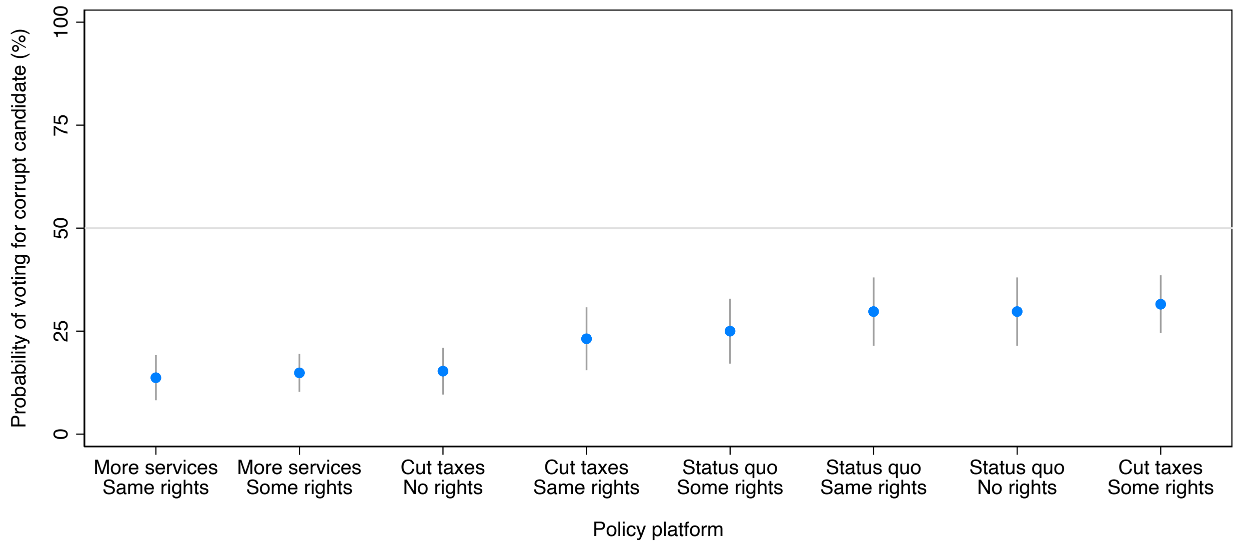


Figure 1.25: Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (conservative respondents and clean challenger)?

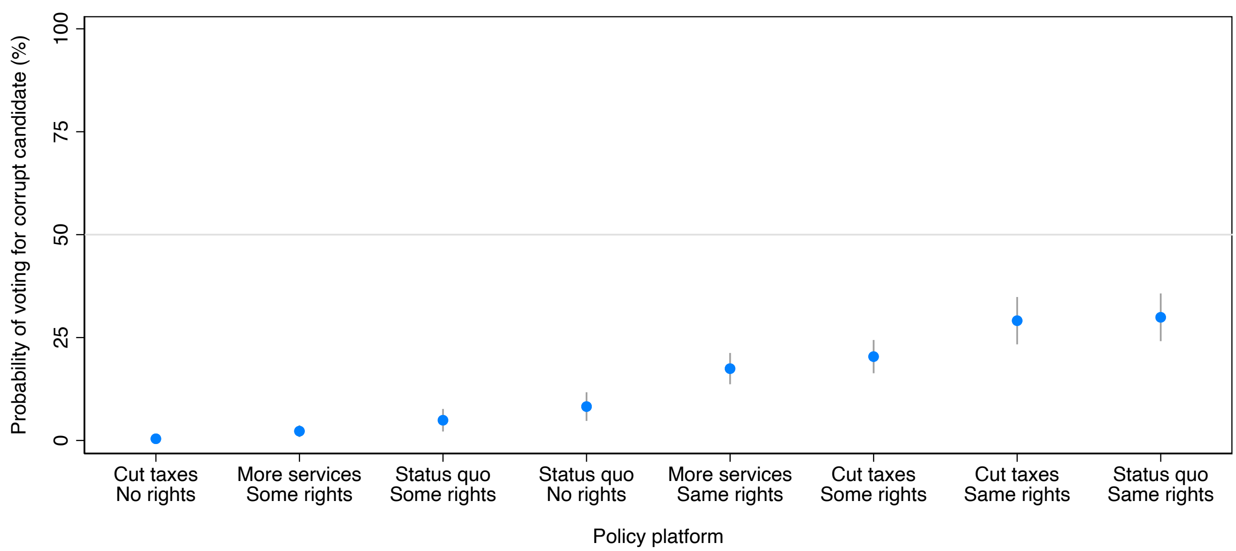


Figure 1.26: Franchino and Zucchini (2015) conjoint: can policy positions overcome corruption (liberal respondents and clean challenger)?

Mares and Visconti (2019)

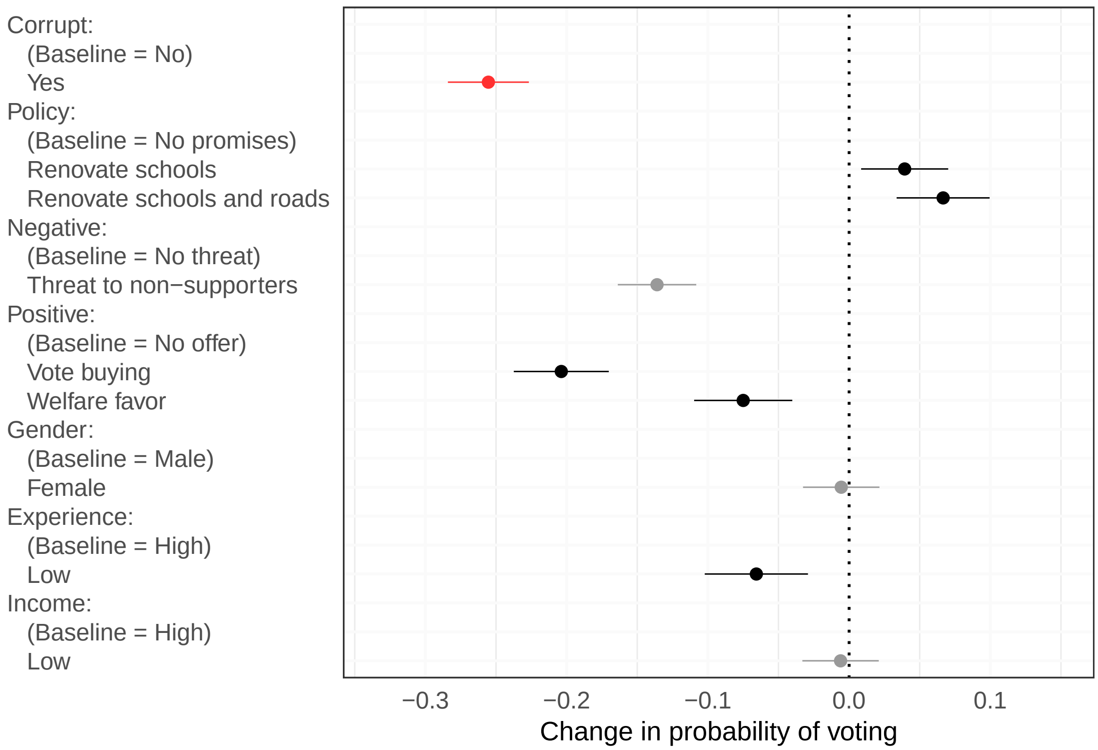


Figure 1.27: Mares and Visconti (2019) conjoint: AMCEs

Note that the primary goal of Mares and Visconti (2019) is to determine the degree to which respondents punish various illicit electoral activities. The experiment therefore includes a number of other negative attributes in addition to corruption, such as vote buying, clientelistic offerings, and threats of violence against political opponents. Due to uniform randomization, calculating predicted probabilities that do not include these attributes therefore marginalizes over a number of other illicit activities that respondents view negatively and reduces overall vote probability. Conditioning on the candidate not engaging in illicit activities other than corruption reveals probabilities of voting for corrupt candidates over 50%.

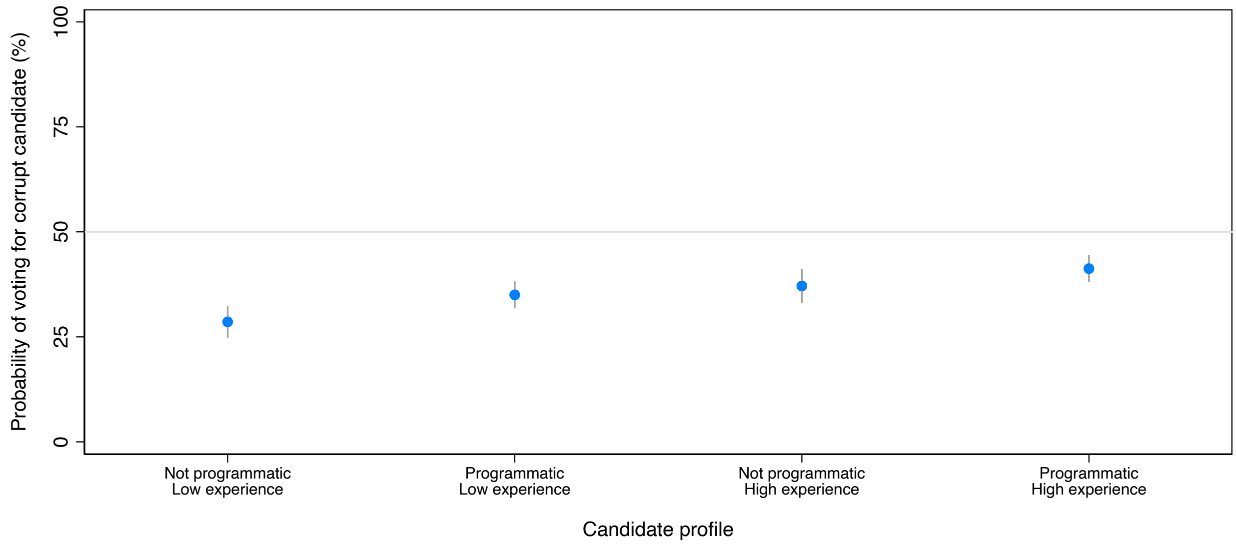


Figure 1.28: Mares and Visconti (2019) conjoint: can programmatic offerings and experience overcome corruption?

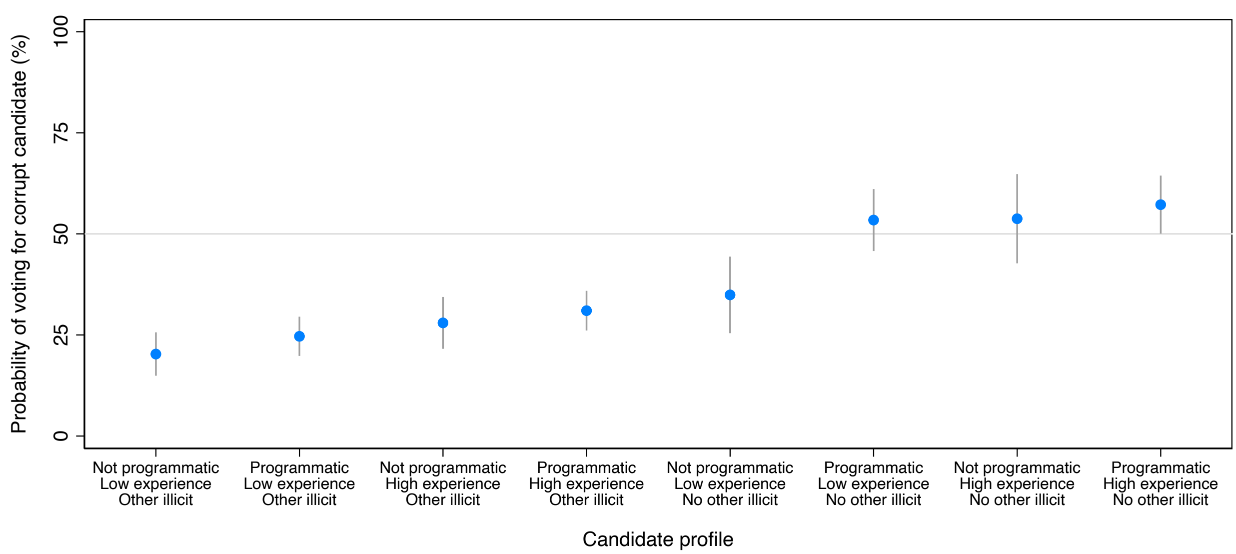


Figure 1.29: Mares and Visconti (2019) conjoint: can programmatic offerings and experience overcome corruption (conditional on other illicit activities)?

Chauchard, Klašnja and Harish (2019)

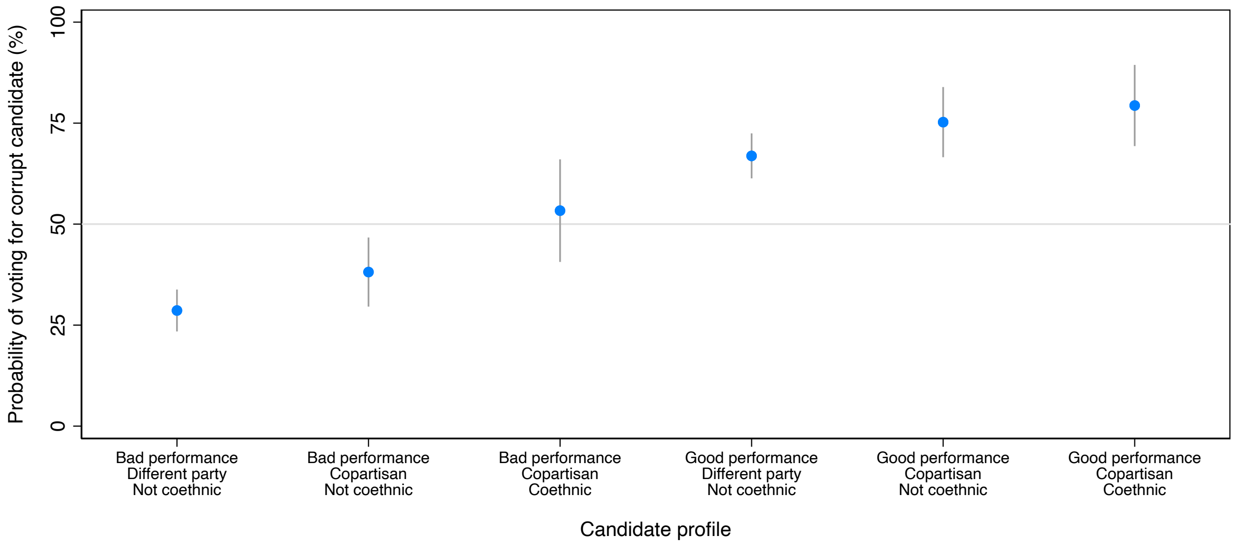


Figure 1.30: Chauchard, Klašnja and Harish (2019) conjoint: can performance, partisanship, and coethnicity overcome corruption?

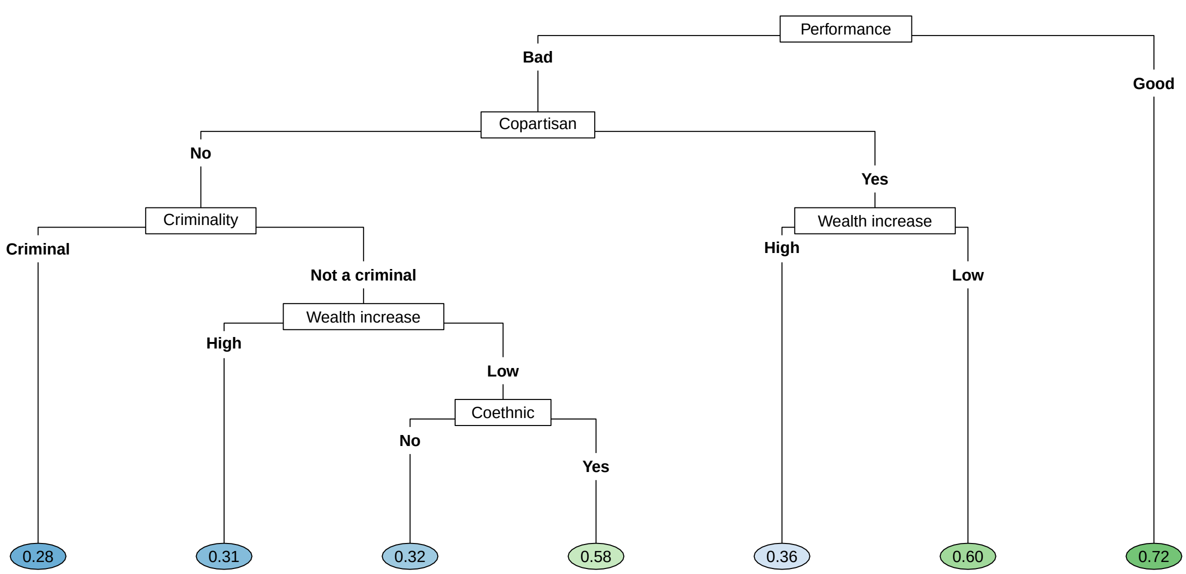


Figure 1.31: Chauchard, Klašnja and Harish (2019) conjoint decision tree: predicted probabilities of voting for corrupt politician

chapter 2

Combatting capture in local politics: Evidence from eight field experiments

Abstract: Understanding how to motivate individuals with long-term collective interest to engage in costly political behavior is an enduring question in political economy. While renters have an economic incentive to participate in local politics and encourage housing growth, their participation lags homeowners who yield immediate financial returns from participation. I conducted 8 email field experiments to investigate how to motivate renters (n=19,951) to comment at city council meetings in opposition to regulations that harm them. Opening a message highlighting high costs of abstention caused a 1.4 percentage point increase in public comments, with voters particularly responsive to treatment. Treatment-induced comments represented 8% of total comments and 46% of pro-housing comments across all treated meetings. These results suggest that increasing the perception that abstention is costly is an effective motivator of collective action, and that outreach can make civic bodies greater reflect the broader public where increases in accessibility alone do not.

Homeowners are more likely than renters to participate in local government, and participate in ways consistent with protecting their property values (Hall and Yoder Forthcoming; Marble and Nall 2021; McCabe 2016; Yoder 2020). Evidence suggests homeowners are more likely than renters to oppose new housing development (Einstein, Palmer and Glick 2019; Hankinson 2018), increase their turnout in elections when zoning rules are on the ballot (Yoder 2020), and participate more often in city council, planning, and zoning meetings (Yoder 2020). Yet the same policies that benefit homeowners often harm renters through decreased access to housing and higher rents (Charette, Herbert, Jakabovics, Marya and McCue 2015; Ganong and Shoag 2017; Glaeser and Gyourko 2018,0; Glaeser, Gyourko and Saks 2005*a*; Lens and Monkkonen 2016; Quigley and Rosenthal 2005; Reeves 2018). Renters therefore also have a monetary incentive to participate in local politics and oppose these policies, but their participation lags that of homeowners. How to translate renter economic self interest into participation in local politics remains unclear.

This lack of renter participation exhibits classic symptoms of the collective action problem—those with only a long-term collective interest in a political decision are less likely to participate in politics than those who realize direct, short-term private gains (Olson 1965; Ostrom 2000). Previous research indicates that self-interest is only a strong motivator of political behavior when immediate costs or benefits are "tangible, large, visible, and certain" (Citrin, Reingold and Green 1990). Homeowners receive tangible, large, and certain benefits from blocking a neighboring development in the form of preserved property value. Benefits are far less tangible and certain for renters, as large-scale new development will only depress rents and increase access to homeownership throughout a diffuse geographic region in the long term. How then can those who only benefit through long-term and uncertain gains such as renters be motivated to engage in personally costly political behavior?

Literature in political economy, behavioral economics, and political psychology offer suggestions. First, lack of residential stability may make renters less connected to their local political system (Ansolabehere 2012; McCabe 2016). Providing instructions and lowering the costs of attendance may therefore encourage participation (Milkman, Beshears, Choi, Laibson and Madrian 2011; Nickerson and Rogers 2010). Second, priming rational economic self interest may encourage participation. Third, renters may be unaware that lack of housing supply is driven by local regulations reflecting homeowner preferences. Pointing out that lack of participation is costly and pressing renters to participate to counteract these regulations may therefore also be a highly effective motivator (Aytaç and Stokes 2019).

I test these instructional, economic cost, and costly abstention theories through a series of email outreach field experiments among approximately 20,000 renter households in 8 cities in Los Angeles (LA) County encouraging individuals to participate in their city council meetings by making public comments. Renter households in LA County were identified by geo-matching addresses in the LA County voter file to Department of City Planning records of multi-unit apartment buildings. Due to COVID-19, public comment was limited to online participation via email, telephone, or videoconference. The experiments were designed and deployed in partnership with a local non-profit organization with years of experience advocating for increased housing supply. Three treatment groups (and a placebo control) tested the three theories outlined above of how to encourage renters to translate their economic self interest into costly political behavior.

Overall, receipt of any treatment increased public comments by 1 percentage point (pp), while highlighting the cost of abstention increased comments by 1.4pp. Individuals already engaged in local politics in the form of voting in local elections were more responsive to treatment (2.3pp) than non-voters (0.9pp). Including all treatment groups, treatment-induced comments represented 8% of total comments and 46% of pro-housing comments across all city council meetings. Pro-housing comments made up a majority of comments in over 50% of treated meetings. This contrasts sharply with previous findings that pro-housing comments typically are in the minority in most council meetings in equilibrium (Einstein, Glick, Puig and Palmer 2021; Yoder 2020).

The results support three primary theoretical and substantive conclusions. First, when abstention is perceived as highly costly, highlighting its consequences is an effective motivator of political participation. Second, the large change in the composition of comments caused by the treatments shows that outreach can change the representation of civic bodies to be more reflective of the broader public where increases in accessibility alone do not. Finally, unlike voting, email is able to effectively increase political participation when participation is also conducted online, particularly amongst those already engaged in politics.

Motivation

Lack of renter participation in local politics

Homeowners are more likely than renters to participate in local politics across numerous dimensions. Renters are less likely to run for office (Einstein, Ornstein and Palmer 2019), vote in local elections, donate to local political candidates, or participate in city council meetings (Yoder 2020).

Moreover, there is evidence that homeowners actively participate in politics to oppose housing development. Homeowner voter turnout roughly doubles in elections when zoning regulations are on the ballot (Hall and Yoder Forthcoming), and those who participate in city council, planning, and zoning board meetings are much more likely to oppose new housing construction than the general public (Einstein, Palmer and Glick 2019; Fischel 2005). In addition, homeowner participation rates are an increasing function of home value (Hall and Yoder Forthcoming; Marble and Nall 2021).

The makeup of local political participation therefore does not typically reflect general public opinion. This is visible in the difference between the percentage of public comments in support of additional housing and the percentage of votes in favor of additional housing on ballot measures. Einstein, Palmer and Glick (2019) show that in Massachusetts, over 50% of voters in the majority of townships supported a ballot measure in favor of affordable housing, while public comments overwhelmingly oppose new developments. Public comments have been shown to have an effect on the substance of policy documents (Judge-Lord 2022), and are likely to affect local regulatory policy outcomes.

The ability of residents to block new housing construction is regularly cited as a key cause of decreases in housing supply (Glaeser, Gyourko and Saks 2005*a*), and supply restrictions are estimated to be a net welfare loss for society (Glaeser and Gyourko 2018). This lack of supply is not primarily due to natural geographic scarcity or construction costs, but government regulation (Brueckner 2009; Glaeser, Gyourko and Saks 2005*a*,0; Gyourko and Molloy 2015; Molloy et al. 2020; Ortalo-Magné and Prat 2014). In fact, Glaeser, Gyourko and Saks (2005*b*) estimate that the effective "regulatory tax"[1] on home prices in Los Angeles and San Francisco were roughly 1/3 and 1/2 respectively in 1999, and as of 2014 home prices were more than double that of production costs (Glaeser and Gyourko 2018).

Impact of lack of housing supply

While economic growth and housing growth used to occur in concert, these processes have decoupled (Glaeser and Gyourko 2018). While new home construction in California was an average of 0.011 houses per capita per year in the late 1960s, this rate declined to 0.002 in the late 2010s (United States Census Bureau 2020). Instead of leading to large increases in home construction and encouraging low-skill migration, economic booms in coastal cities now primarily increase housing prices (Ganong and Shoag 2017).

This decrease in housing supply increases rents, reduces real income for renters, and keeps homeownership out of reach for an increasing number of Americans. Real housing prices in the top quintile of the price distribution doubled in Los Angeles and New York, and tripled in San Francisco since 1970 (Glaeser, Gyourko and Saks 2005*a*; Hankinson 2018). A quarter of renters in the United States currently spend over half their incomes on housing, and this number is expected to grow (Charette et al. 2015). High housing costs also constrain worker mobility, reducing real US economic growth by an estimated 36% from 1964 to 2009 (Hsieh and Moretti 2019) and real GDP by at least 2% (Glaeser and Gyourko 2018).

[1] Note that this is not a literal tax levied by the government, but rather the increased cost of housing caused by local regulations that restrict supply.

By contrast, higher housing prices benefit current homeowners, increasing their net worth and exacerbating income inequality. This housing wealth is concentrated amongst individuals in coastal regions who purchased homes in the decades before regulatory and political constraints on housing development were imposed (Glaeser and Gyourko 2018). For example, between 1983 and 2013 housing net worth increased by an average of 57% for those 65 and older in the 90th percentile of the wealth distribution (Glaeser and Gyourko 2018).[2]

Lack of housing supply also increases energy use and greenhouse gas (GHG) emissions. Some studies estimate that increasing urban infill—for example, replacing surface parking lots with apartment buildings—would have a larger effect on GHG emissions reductions than mass adoption of electric vehicles (Wheeler, Jones and Kammen 2018). Others are less bullish, but still highlight the importance of increasing urban infill on reducing GHG emissions (Cervero and Murakami 2010).

Theory and hypotheses

Lack of collective action by renters

Economic self-interest has been shown not to be a strong behavioral motivator of collective action except when the costs and benefits are "immediate," "tangible," "large," "visible," and "certain" (Citrin, Reingold and Green 1990; Sears and Funk 1991; Sears and Citrin 1982). Benefit to renters are far from immediate or certain, as their rents are only expected to decline after years of as large-scale development. By contrast, blocking (or failing to block) a neighboring development can have a large, certain, and immediate impact on property values for homeowners. However, additional theories also may offer explanations of why renters in particular tend not to participate in politics, and these theories can be evaluated on the basis of the extant empirical evidence.

First, lack of residential stability may make renters less connected to their local political

[2]For those in the 95th and 99th percentiles these numbers are 65% and 112%, respectively. See Figure 2.7 for a visualization of changes in all age groups and income percentiles.

system (Ansolabehere 2012; McCabe 2016). Simple increases in accessibility offered by moving city council meetings online during the COVID-19 pandemic did not meaningfully alter renter participation in 2020 (Einstein et al. 2021),[3] suggesting that high costs of attendance are likely not the primary barrier.[4] Fewer local connections may lead to a lack of information on how to participate and contribute to low renter participation. Providing information about the content of council meeting agenda items and instructions of how to participate is therefore a necessary minimum to spur collective action.

Second, lack of renter participation does not appear to be caused by general opposition to development, unlike homeowners. While homeowners consistently oppose new housing across all geographies, renters do not (Hankinson 2018; Marble and Nall 2021; Monkkonen and Manville 2019). Renters in high-cost cities sometimes oppose market-rate housing at the neighborhood level, but do not at the city level (Hankinson 2018). Highlighting "affordable" (i.e., government subsidized) or "missing middle" housing developments has been shown to increase support for more housing (Doberstein, Hickey and Li 2016), while highlighting developer profits is met with backlash (Monkkonen and Manville 2019). Messages encouraging collective action should therefore focus on the city-wide public benefits of increased housing, and highlight affordable housing.

Encouraging political participation: general evidence

Research in political economy, behavioral economics and experimental psychology offer lessons for encouraging participation inside and outside of the ballot box. I first provide a brief overview of the evidence these literatures provide on which mode of delivery is most effective in order to maximize participation, how instructions should be given, and which behavioral motivators should be used. I then show how my treatments are consistent with best practices from these literatures.

[3]Homeowners still made up 78% of commenters, and anti-housing comments comprised the majority in 35 out of 36 towns examined in Einstein et al. (2021)'s study in the Boston area.

[4]I recognize that homeowners are on average older, more likely to be retired, and wealthier (Yoder 2020), and therefore may have a lower opportunity cost of attendance, even in an online setting. While I do not test this theory directly here, it is likely a constant presence across all of my treatment arms.

In terms of instructions and message structure, past studies suggest that merely overcoming information costs should not have a large impact on political behavior (Green and Gerber 2019; Riker and Ordeshook 1973). By contrast, giving individuals a detailed plan of how and when to participate has been shown to be an effective method of increasing both voter turnout (Nickerson and Rogers 2010) and vaccination rates (Milkman et al. 2011). These studies suggest that clear, concrete instructions of how and when to participate lower the cost of participation.

In terms of mode of delivery, get-out-the-vote (GOTV) experiments find email has proven largely ineffective at increasing voter turnout (Green and Gerber 2019; Malhotra, Michelson, Valenzuela et al. 2012; Nickerson et al. 2007).[5] However, the efficacy of email at increasing political participation that can itself be conducted electronically—lowering the cost of participation substantially—is less clear. Due to COVID-19, public comment in 2021 was limited to online participation via email, telephone, or videoconference. Green and Gerber (2019) note that "it is one thing to present recipients with options that they can choose from immediately from the comfort of their chair, such as visiting a website that tells them where to vote. More difficult is to motivate the recipient to get out of that chair on some future date in order to cast a ballot." We possess scarce evidence of the effectiveness of email on participation that does not require the recipient to get out of the chair.

In terms of message content and behavioral motivations, economic self-interest alone is typically not the most effective motivator of collective action mobilization (Citrin, Green, Muste and Wong 1997; De Rooij, Green and Gerber 2009; Ostrom 2000; Sears and Funk 1991), except when the benefits are large and certain. By contrast, psychological motivators such as highlighting social norms and comparing treated individuals' behavior to neighbors and peers has been shown to induce costly pro-social behavior such as voting and energy saving (Allcott 2011; Gerber, Green and Larimer 2008). Relatedly, Aytaç and Stokes (2019) posit that there are often high psychological costs to abstention from political participation,

[5] Even in a low salience election, emails sent from the county registrar of San Mateo, California only increased turnout by 0.56 percentage points (Malhotra et al. 2012).

and that messages that elicit emotional responses (e.g., shame, anger, anxiety, etc.) draw people to collective action. These costs relate to neighbor and peer comparisons as abstention in the housing context is costly due to a lack of abstention by homeowner neighbors. Formally, Aytaç and Stokes (2019) posit a model with rewards of participation $P = A - C + D_E$ where P is rewards from participation, A is the cost of abstention, C is the cost of participation, and D_E is social pressure. A treatment that increases the costs of abstention A, decreases the cost of participation C, and provides social pressure D_E should therefore maximize the rewards from participation.

In sum, theory and lessons from prior research can inform the design of a campaign encouraging individuals to participate in local politics. Messaging strategies that: (1) lower costs of participation with concrete reminders of dates and simple but detailed instructions of how to participate, (2) increase social pressure by comparing non-participants to their neighbors, and (3) emphasize the high (economic and psychological) costs of abstention should be particularly effective. While emails are likely less effective than direct conversations, emails may be more effective in an online public comment context compared to GOTV due to the lower costs of electronic participation.

Tying theory to treatments

The observations in Lack of collective action by renters and Encouraging political participation: general evidence lead to three distinct but related theories of motivation to collective action, which I test with three distinct treatment arms. The relative efficacies of each treatment arm hypothesized below were pre-registered.

First, a treatment (T1) that lowers costs of participation with simple but detailed instructions of how to participate should increase attendance, but the effects should be small in magnitude (Green and Gerber 2019; Riker and Ordeshook 1973). To lower costs of participation, all treatment messages include a link that opens an email message with a pre-filled sample public comment[6] that is pre-signed with the respondent's name, while also noting

[6]See Sample comment for the wording of the sample message.

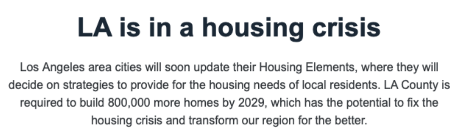
that individuals may draft their own comment. All treatment messages also include the phone number or Zoom link needed to submit a spoken comment.[7]

Second, a treatment (T2) providing information that lack of housing supply increases rents should increase attendance more than attendance instructions only by also priming economic self interest.

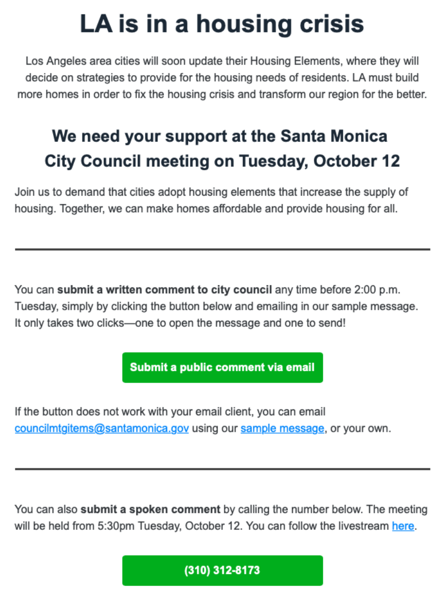
Third, a treatment (T3) that not only lowers costs of participation, but also pressures renters to participate and points out that a lack of participation is costly (i.e., costly abstention theory) should increase attendance more than lower costs of participation or economic self-interest alone. This provides the first experimental test of Aytaç and Stokes (2019) costly abstention theory in a real-world setting, and extends their theory into an understudied vein of civic participation—city council meetings.

This paper therefore provides three major empirical advancements. First, I provide the first real-world test of how to motivate renters to participate in local politics by comparing three theoretically-motivated treatment arms that allow me to examine the efficacy of three distinct potential motivators of collective action. Second, I examine if this motivation can meaningfully change the makeup of participation in city councils. Third, I test whether email is an effective motivator of online political participation.

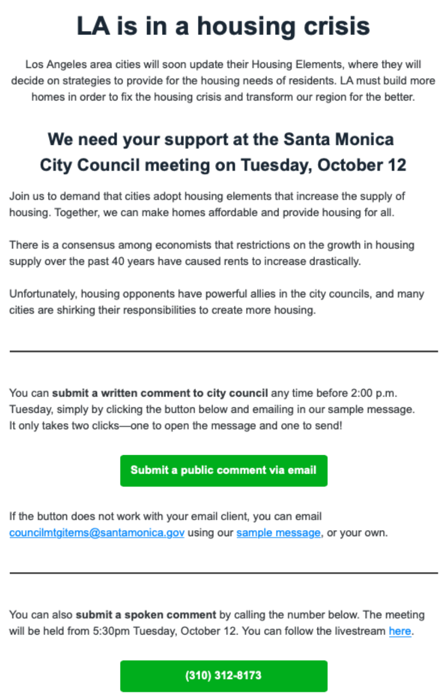
[7]Individuals were not encouraged to attend council meetings in person (even if possible) due to the COVID-19 pandemic.



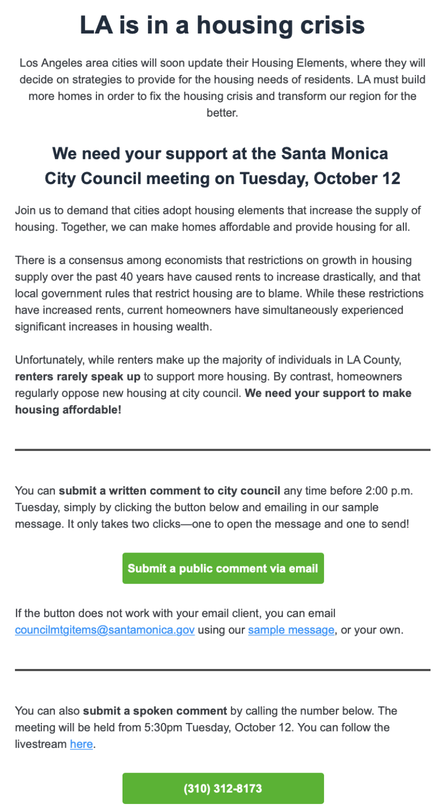
(a) Placebo treatment message



(b) Instructions only treatment message



(c) Economic treatment message



(d) Incentive treatment message

Figure 2.1: Treatment groups

Research design

Context

The COVID-19 pandemic moved city council meetings online. Some hypothesized that the more accessible nature of online meetings would diversify participation (Hernández 2021; Los Angeles Times 2021; The Boston Globe 2021). However, the participation gap between renters and homeowners did not decrease (Einstein et al. 2021).

Public comments can be made in spoken or written format. Written comments can be submitted either by emailing the city clerk, or in some cases using an online "e-comment" system. Depending on the city in question, written comments are either read aloud during the council meeting or distributed to council members and posted online prior to the meeting. City council members should therefore be aware of the sentiments expressed in public comments, whether spoken or written.

The experiment was fielded in Los Angeles (LA) County in partnership with a pro-housing NGO, and to the best of our knowledge is the first field experiment to examine participation in city council meetings. LA County cities were in the process of updating their 2021-2029 "Housing Elements." Housing Elements are updated once per decade as required by California state law, and are an analysis of a city's housing needs and strategies to be used to meet those needs. State law requires that each city accommodate its fair share of affordable housing, and requires local governments to adopt land-use plans and regulations that provide opportunities "for, and do not unduly constrain," housing development by the private sector. The experiment therefore targets city council meetings in which the updated Housing Element is on the agenda, and treatments encourage individuals to advocate for historically high housing growth.

While there is a vocal anti-development contingent in Los Angeles, the general voting public appears to support additional housing as a number of local anti-development ballot measures have recently failed. Measure S, which would have drastically curbed high-density development in the city, failed with only 30% support. Meanwhile, Measure JJJ—which grants zoning changes to developments that include affordable housing—passed. Measure H, which instituted a sales tax increase to fund (among other things) affordable housing, also passed. Further, only 28% of respondents in a survey of LA county residents oppose a hypothetical local development Monkkonen and Manville (2019).

The geographic and regulatory landscape in Los Angeles leads to a majority of new housing developments replacing parking lots or commercial buildings, not existing housing stock. Roughly 14% of land, or over 200 square miles, is currently dedicated to parking (Chester, Fraser, Matute, Flower and Pendyala 2015). In addition, the city of Los Angeles requires affordable housing in exchange for increases in density above current zoning limits. For example, the transit-oriented-communities program (part of Measure JJJ passed by voters in 2016) requires at least 25% of units to be set aside at the low income level in exchange for the largest density bonuses near major transit stops.

Ethics

Any intervention motivating individuals to change their behavior should be held to high ethical standards, particularly when the intervention involves participation in and effects on governmental processes. Beyond IRB approval, I argue this project falls within ethical bounds for the reasons outlined below.

First, the interventions are designed to minimize a pre-existing imbalance in representation by increasing representation amongst a historically underrepresented group. Treatments are designed to encourage renters to participate (albeit not coercively) and make local governance more reflective of the general population.

Second, the interventions do not directly effect electoral outcomes (as highlighted by Slough (2019) and McDermott and Hatemi (2020)). I recognize that local officials may change their votes based on perceived changes in support levels that the experiment might cause. However, ultimate decisions and votes still rest with local elected officials.

Third, the interventions focus on increasing the supply of housing generally across the LA region, not on particular developments or neighborhoods. Treatment and sample messages also specifically encourage individuals to advocate for *affordable* (i.e., government subsidized) housing developments. We should therefore expect the targeted groups to benefit from the research through decreased rents and increased access to affordable housing.

Fourth, in social-welfare enhancing interventions such as "green nudges," Bovens (2009) and Schubert (2017) argue that it should be possible "for everyone who is watchful to unmask the manipulation." The interventions meet this criteria, as the messages come from an advocacy group that is transparent in their motivations and involve no deception.

Experiment overview

The experiment proceeded in the following steps: (1) renters in the voter file were identified using LA city planning records, (2) city council meetings were monitored for agenda items discussing their Housing Element, and these council meetings were selected for the messaging campaign, (3) renters were randomly assigned to one of three email treatments asking them either to turn out to support increases in housing supply or a placebo control, (4) names in all treatment groups were matched with names of individuals who submitted a public comment, (5) analysis was performed using pre-registered outcomes and estimators. More detailed explanations of the processes follows below.

Identifying renters

I identified renters in Los Angeles County by geo-matching addresses in the voter file with Los Angeles County Department of City Planning records of multi-unit apartment buildings. This process was conducted using the FastLink probabilistic linkage algorithm developed in Enamorado, Fifield and Imai (2019). Only records with a 99.2% or greater posterior probability of a correct match were kept,[8] resulting in 641,184 matched renters, 266,057 of

[8] Manual checking of a random sample of 100 records indicated that 98% of matches with posterior probability above 99.2% were correct, while 96% of matches below posterior probability 99.2% were false positives.

whom listed their email addresses in the voter file.

Identifying council meetings

Partner organizations monitored city council meetings in LA County for agenda items discussing the Housing Element throughout fall and winter 2021. The specific timeline (i.e., recruitment starting and stopping dates) was pre-registered. Renters identified in the voter file as living in these cities were then targeted to receive emails encouraging them to submit a public comment on the Housing Element agenda item at their city council meeting. Ultimately, one council meeting in Santa Monica and two council meetings in Long Beach were selected for pilot studies, followed by pre-registration and treatment of individuals targeting council meetings in the cities of (in chronological order) Beverly Hills, Santa Monica, Whittier, Rancho Palos Verdes, Manhattan Beach, Norwalk, Sierra Madre, and Culver City.

Treatment assignment

Likely renters in the voter file were randomly assigned to an email treatment encouraging them to submit a public comment at their city council meeting, or a placebo control. Individuals were block randomly assigned by city[9] and cluster randomly assigned by address.[10] Treatment assignment probabilities were as follows: 10% probability of assignment to a placebo message with no information on how to attend a meeting, 30% probability of assignment to T1, 30% probability of assignment to T2, and 30% probability of assignment to T3. Balance tables by treatment or placebo status, as well as for each treatment group can be found in Balance, and a map of all cities that received treatment can be found in Figure 2.8. All treatments included identical subject lines and preview texts in order to ensure equal compliance rates across treatment arms.

[9]While random assignment took place simultaneously for all cities, treatments were launched at different points in time for each city. For this reason it is also reasonable to think each city as a separate experiment, rather than as blocks in a single simultaneous experiment.

[10]If a unit number was available, clustering took place at the unit level. If a unit number was not available, clustering took place at the building level.

Outcomes

The primary, pre-registered outcome of interest is a binary indicator of whether an individual submitted a spoken *or* written comment. As participation in a public hearing is a matter of public record, I match the names of those in the treatment group(s) with spoken or written comments. I also investigate *how* individuals commented through the creation of separate binary indicators for: spoken comments, written comments, comments that used our pre-written messages, custom comments, pro-housing comments, and anti-housing comments. In addition, I investigate whether the treatments changed the overall makeup of council meeting comments by comparing the number of pro-housing comments that were likely treatment induced with those that were not. I define "likely treatment induced" comments as those submitted by individuals in one of the three treatment groups. This definition seems reasonable, as no comments were made by compliers in the placebo group.

Analytical procedures

Analytical procedures were pre-registered with the Center for Open Science Open Science Framework (OSF) prior to data collection or analysis. The primary (and pre-registered) estimand of interest is the complier average causal effect (CACE),[11] of opening an email on submission of a public comment. In other words, I estimate the average treatment effect for only the subset of individuals who opened the emails (i.e., compliers). I employ a placebo-controlled design in order to mitigate statistical uncertainty (Broockman, Kalla and Sekhon 2017; Nickerson 2008). By randomly assigning individuals to a placebo control with no mention of council meetings, but featuring the same subject line and preview text as the treatment emails, I am able to observe the outcomes of a random sample of compliers (email openers) in the placebo group.[12] I can then compare email openers in treatment directly to

[11] Also commonly known as local average treatment effect (LATE).

[12] Tests for differential compliance by treatment group and differential covariate predictiveness of compliance can be found in Figure 2.9 and Table 2. While some covariates are predictive of compliance, they tend to be similarly predictive of compliance across treatment groups.

email openers in placebo.

I therefore monitor if an email was opened as a measure of compliance, and estimate the CACE using the estimator outlined by Lin (2013). I include the following pre-registered pre-treatment covariates in the regression specification: *city, number of units in the building, gender, age, building age, primary language spoken, vote history, and party affiliation.*[13] Missing covariates are mean imputed. As units were cluster randomly assigned by address, standard errors are clustered at the address level. Results are also reported in Robustness without covariate adjustment. I therefore estimate the OLS specifications below:

$$Y_i = \alpha + \beta_1 Z_i + \beta_2 X_i^c + \gamma X_i^c Z_i + \delta_{city} + \epsilon_i \qquad \text{(With Lin (2013) covariate adjustment)}$$

$$Y_i = \alpha + \beta_1 Z_i + \delta_{city} + \epsilon_i \qquad \text{(Without covariate adjustment)}$$

where Y_i is the individual-level comment outcome, Z_i is an indicator for the treatment group, X_i^c is a vector of pre-treatment covariates for unit i that have been centered to have mean zero, and δ_{city} are city (block) fixed effects.

As the outcome data take the form of "rare event" right-skewed binomial distributions (see Figure 2.14), I also calculate randomization inference based p-values (RI p)[14] free from distributional assumptions as an extra robustness test. In addition, I re-estimate all models using penalized maximum likelihood.[15]

Results are analyzed as above (i.e., as one large experiment with city fixed effects), as well as aggregated using precison-weighted[16] fixed effects and random effects meta-analysis.

[13] I show that these variables are balanced between the placebo and treatment groups in Balance.

[14] Specifically, I simulate a large number of "fake" random assignments for all units using the same procedure as the real random assignment, and estimate a treatment effect for each fake random assignment. I then calculate a p-value as the proportion of times fake treatment assignments resulted in an effect size larger than the actual treatment effect. For the CACE, I make the additional assumption that observed compliance would exist regardless of treatment status and hold compliers constant across simulations. I conduct 10,000 simulations for the CACE and 1000 simulations for the ITT. All simulations were performed without covariate adjustment due to high computational demands.

[15] I do not use traditional logistic regression due to the skewed nature of my outcome variable (i.e., becuase the comments in my sample represent are "rare events"). See King and Zeng (2001) and Cook, Hays and Franzese (2020) for discussions of maximum likelihood estimation in the case of rare events.

[16] With weights equal to the inverse of the variance.

This pre-registered, prospective multi-site study can be shown to be a valid application of meta-analysis.[17] Fixed effects meta-analysis—which assumes estimates vary across studies only due to having just a sample of observations from the total population—is often not appropriate in the social sciences. However, as identical pre-registered experiments were administered to different populations and measured the same outcome, it may be appropriate in this context. Nevertheless, I also include estimates using random effects meta-analysis, as well as excluding results from three pilot studies for robustness purposes.[18] For council meetings where no comments are reported in treatment or placebo, I estimate standard errors according to the procedure described in Gelman and Hill (2006).[19]

I also examine pre-registered heterogeneous treatment effects by the density of the building in which an individual lives, median area income, and turnout in the most recent local election.[20] Individuals who live in high-density buildings may be pre-disposed to a more dense urban environment. Income may correlate with a desire for, in particular, affordable housing development. Voters in local elections are pre-engaged in local politics, and may therefore perceive abstention as more costly than others and/or vote in part due to pre-existing opinions about development. I analyze heterogeneous treatment effects in two ways. First, I take the traditional (and pre-registered) experimental approach of regressing the outcome variables on treatments and the interaction between the treatment and the covariate, sometimes referred to as a conditional average treatment effect (CATE).[21] I also

[17]Borrowing the framework, language, and notation of Slough and Tyson (2021), the constituent studies $\mathcal{E}_i$ of the meta-study contain a measurement strategy m_i (a binary comment indicator), contrasts (ω_i', ω_i'') where ω_i' is the control condition and ω_i'' is the treatment condition, and setting/city (θ_i) specific treatment effect $\tau_{mi}(\omega_i', \omega_i'' \mid \theta_i)$, where all studies in the meta-analysis are constituent comparable ($\tau_{m1}(\omega_1', \omega_1'' \mid \theta_1) = \tau_{mi}(\omega_i', \omega_i'' \mid \theta_i)\ \forall\ i \in \{1, ..., n\}$) and that all studies are measurement harmonized ($m_1 = m_i\ \forall\ i \in \{1, ..., n\}$) and contrast harmonized ($\omega_1' = \omega_i'\ \forall\ i \in \{1, ..., n\}$ and $\omega_1'' = \omega_i''\ \forall\ i \in \{1, ..., n\}$).

[18]Meta-analysis excluding the pilot studies is performed for robustness purposes.

[19]See p. 17, footnote 1: "Consider a survey of size n with y Yes responses and $n - y$ No responses. The estimated proportion of the population who would answer Yes to this survey is $\hat{p} = y/n$, and the standard error of this estimate is $\sqrt{\hat{p}(1-\hat{p})/n}$. This estimate and standard error are usually reasonable unless $y = 0$ or $n - y = 0$, in which case the resulting standard error estimate of zero is misleading. A reasonable quick correction when y or $n - y$ is near zero is to use the estimate $\hat{p} = (y + 1)/(n + 2)$ with standard error $\sqrt{\hat{p}(1-\hat{p})/n}$."

[20]The most recent county-wide local elections at the time of data acquisition were the March 7, 2017 consolidated municipal and special elections.

[21]Note that this is analogous to conducting separate analyses by subgroup.

use randomization inference as a robustness check.[22]

Results

Overall

Across all council meetings,[23] the effect of opening any treatment email on submitting a public comment (i.e., the CACE) was 1.02 [RI p = 0.044; 95% CI 0.66, 1.38] percentage points (pp). The effect of being assigned to treatment (i.e., the ITT) on submitting a public comment was 0.19pp [RI p = 0.075, 95% CI 0.06, 0.31]. Both estimates are depicted graphically in Figure 2.2. Estimates in tabular form and without covariate adjustment are reported in the appendix. Compliance rates by treatment group were 17% in placebo, 17% in T1, 16% in T2, and 18% in T3 (see Figure 2.9 for a formal test of differential compliance by treatment group).

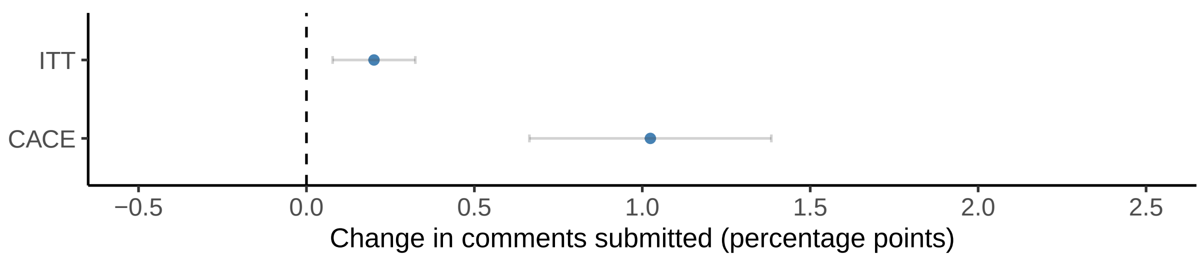


Figure 2.2: Intent-to-treat effect and complier average causal effect, all cities

CACEs for individual council meetings can be found in Figure 2.3. In addition to depicting the CACE in individual cities, Figure 2.3 also contains estimates of the CACE across council meetings using fixed and random effects meta-analysis. Figure 2.3 contains individual and meta-analytic estimates from three pilot studies, increasing the sample size to over 27,000 households. The point estimate using fixed effects meta-analysis including the pilot

[22]Specifically, I generate the full schedule of potential outcomes under the null hypothesis that the true treatment effect is constant and equal to the estimated CACE. Then, I simulate random assignment 10,000 times and calculate the proportion of instances the simulated estimate of the interaction effect is at least as large (in absolute value) as the actual estimate.

[23]Not including the aforementioned pilot studies.

studies is 0.78 [95% CI 0.51, 1.06], and excluding the pilot studies is 0.91 [95% CI 0.56, 1.25] (see Figure 2.13).

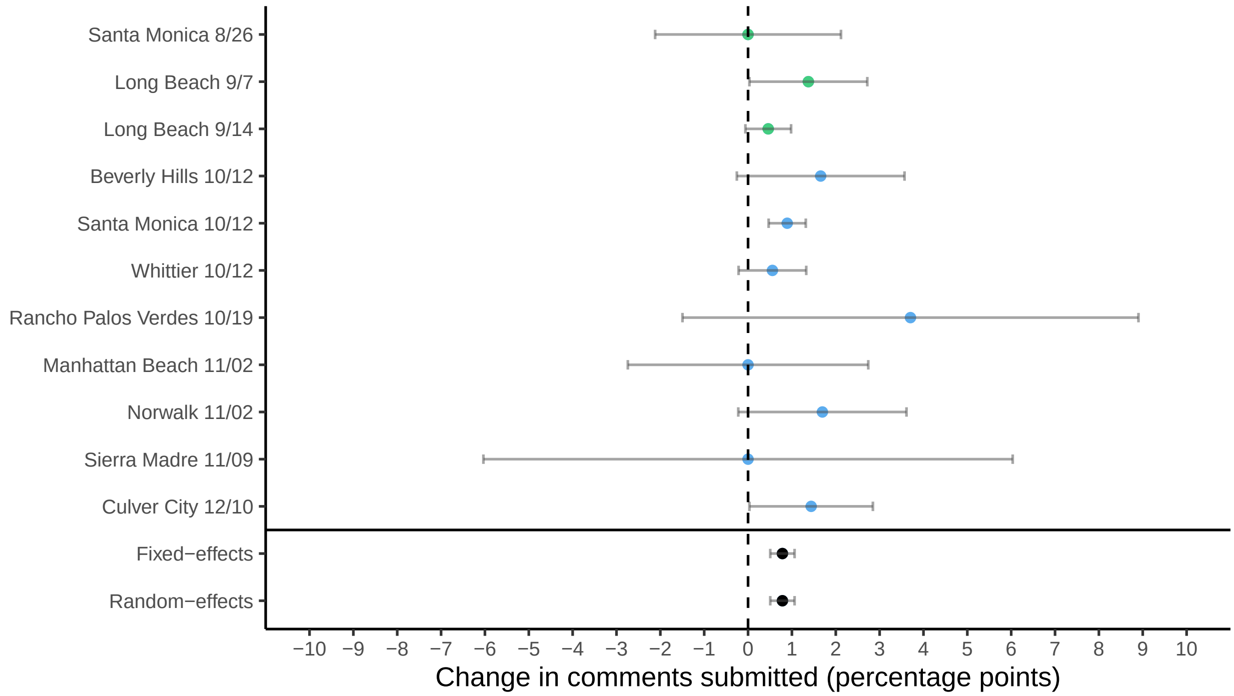


Figure 2.3: Meta-analysis of complier average causal effects, by council meeting

Note: Pilot studies in green.

By treatment group

Figure 2.4 depicts CACEs and ITTs by individual treatment group. In line with pre-registered hypotheses, Figure 2.4 shows that highlighting the costs of abstention had the largest effect on turnout (CACE = 1.44pp; RI p = 0.011; 95% CI [0.73, 2.15]), priming economic self interest was the second most effective (CACE = 1.01pp; RI p = 0.071; 95% CI [0.39, 1.63]), and the instructions-only treatment was the least effective (CACE = 0.54pp; RI p = 0.386; 95% CI [0.06, 1.03]).[24] Using the pre-registered Analytical procedures, the instructions-only treatment was significantly different from zero at the 5% level, while the economic cost and costly abstention treatments were each significantly different from zero at the 1% level. The estimates for the costly abstention and instructions-only treatments are

[24]ITT randomization inference p-values are: 0.380 for T1, 0.089 for T2, and 0.039 for T3.

significantly different from each other at the 5% level.[25]

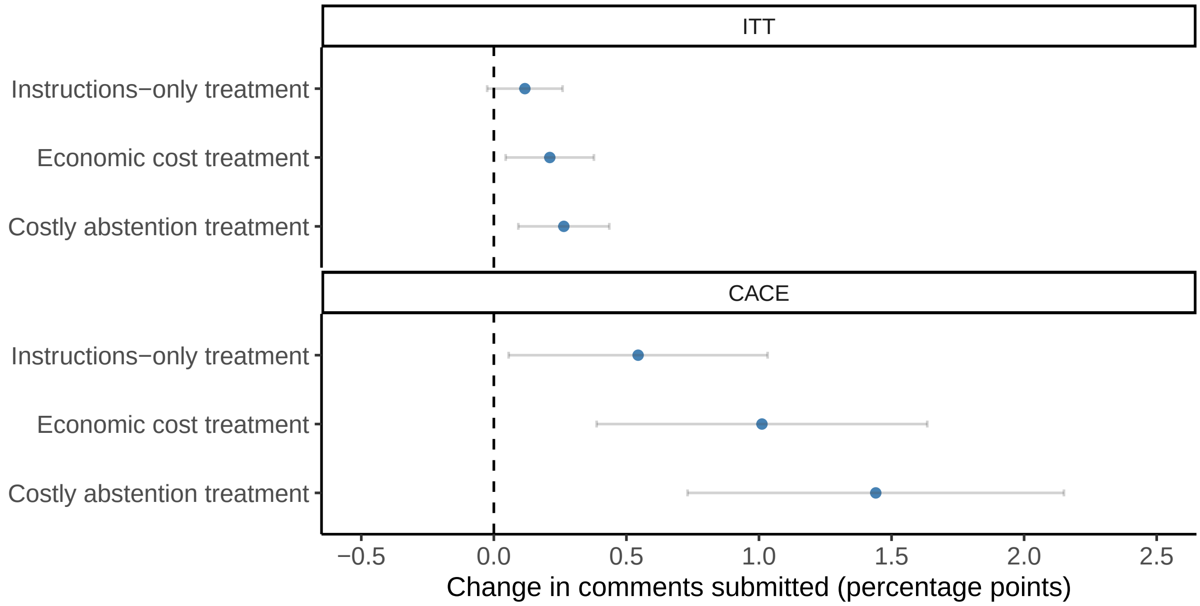


Figure 2.4: Effects by treatment group, all cities

To further asses confidence the costly abstention treatment was the most effective and aid interpretation, I fit a Bayesian linear multilevel model using prior distributions from the power analysis in my pre-registration. Coefficient estimates and posterior distributions can be found in Figure 2.15. Figure 2.16 provides a visualization of the posterior distributions of each coefficient and the posterior distributions of the differences between each coefficient, finding strong evidence that the null hypothesis they are equivalent can be rejected. Next, I compute Bayes factors[26] for hypotheses that the differences between treatments are greater than zero (e.g., costly abstention treatment - instructions only treatment > 0) and its alternative using the Savage-Dickey density ratio method. The Bayes factors are 97 and 5 for the costly abstention treatment vs. the instructions only treatment and costly abstention treatment vs. economic cost treatment, respectively.[27] This provides strong evidence that

[25] Based on a two-tailed linear hypothesis test.

[26] The posterior odds of one hypothesis when the prior probabilities of the two hypotheses are equal. Or more colloquially, the ratio of the likelihood of one particular hypothesis to the likelihood of another hypothesis. A Bayes factor of 5 implies the alternative hypothesis is 5 times as likely as the null hypothesis given the data.

[27] The posterior probability exceeds 95% for a one-sided hypothesis test in both comparisons, and exceeds

the costly abstention treatment was more effective than the instructions only treatment, and moderate evidence that it was more effective than the economic cost treatment.

These results confirm the (pre-registered) theoretical predictions. Lowering costs of participation with simple but detailed instructions of how to participate may have increased attendance, but only marginally. Priming economic concerns appears to be more effective than lowering participation costs alone. Finally, the strongest evidence supports Aytaç and Stokes (2019) theory that highlighting the perceived costs of abstention is more effective than lowering costs or economic self-interest alone.

Heterogeneous treatment effects

I find suggestive evidence that turnout in local elections is associated with an increased likelihood of being persuaded to make a public comment, and that the magnitude of this association is sizable.[28] OLS including a treatment-by-covariate interaction suggests that voters in local elections who opened the messages were 1.4pp more likely to comment than those who did not vote (see Figure 2.5). However, this association is only significant at the 10% level ($p = 0.086$).[29] A randomization inference based hypothesis test returns a p-value of 0.06. Voters were also more likely to open the emails across all treatment groups (see Table 2), suggesting greater engagement in local politics in general.

There is therefore suggestive evidence that participation in local politics in the form of voting begets willingness to participate in other forms, such as attending council meetings, submitting public comments, and engaging with outreach campaigns. This information is potentially relevant to practitioners on a limited budget, as they may see higher returns to participation by targeting likely voters.

95% for a two-sided test in the first comparison. Given that the directionality and relative magnitudes of the treatment effects were pre-registered and negative treatment effects are theoretically implausible, a one-sided hypothesis test seems reasonable.

[28] I do not uncover evidence that the other pre-registered covariates of interest—building density or median area income—are strongly associated with commenting.

[29] The uncertainty of the estimates are a result of low turnout (9.4% amongst the sample population)

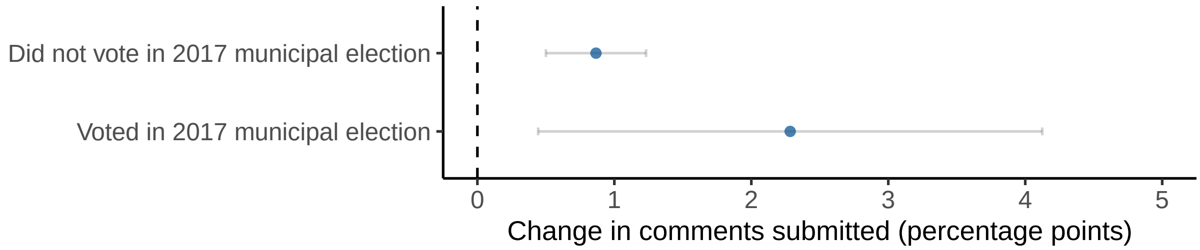


Figure 2.5: Complier average causal effects by turnout

How individuals commented

The vast majority of individuals (93%) submitted written public comments. In fact, the null hypothesis of no effect for spoken comments cannot be rejected. However, even written submissions were not purely costless. While the majority of written comments used the sample message included in the email, 29% represented custom, personal comments (see Figure 2.6).

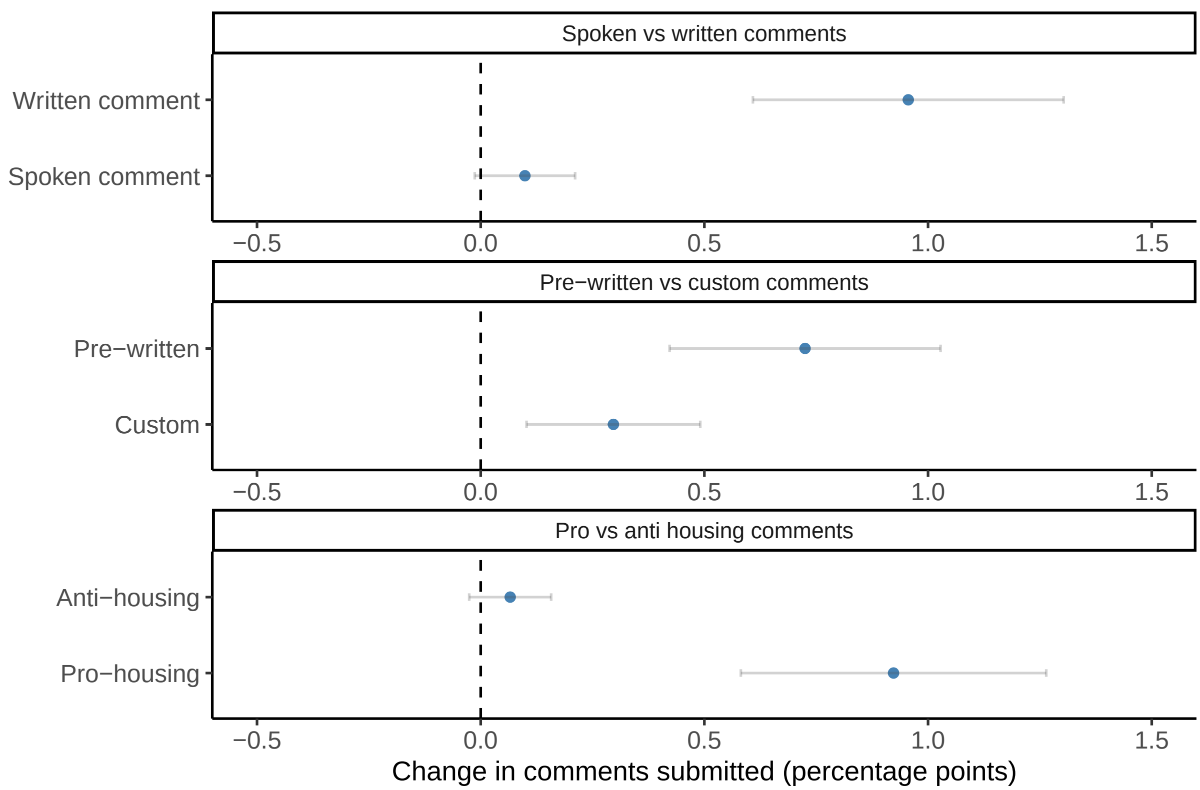


Figure 2.6: CACE by type of comment

Comment contents

I also examine the content of the comments in order to determine if the treatments provoked backlash in the form of anti-housing comments, if individuals submitted spoken or written comments, and if individuals submitted custom comments or used the pre-written comment supplied in the emails.

First, I investigate whether the treatments provoked backlash in the form of *anti*-housing comments. While some anti-housing comments were submitted (see Figure 2.6), they represented only 4% of total comments, and never comprised a majority of experimentally-induced comments in any council meetings.

As noted in How individuals commented, 29% of written commenters did not use our sample message, but instead drafted their own custom comments. I do not provide quotes of custom experimentally-induced comments as I did not ask for consent to re-print individuals' public comments. However, many of these custom comments were deeply personal and reflected individuals' lived experiences with high housing costs. For example, some discussed near experiences with homelessness, senior commenters discussed fear of being priced out of subsidized senior housing, and young renters lamented their inability to purchase a home like their parents.

Substantive impact of comments and changes in representation

In addition to the effect of contact on comment submission at the individual level, I also investigate the substantive impact of the campaigns on each council meeting. Table 2.1 shows that the treatments meaningfully changed the quantity and composition of comments in individual meetings. Overall, likely treatment-induced comments—i.e., comments made by individuals in one of the three treatment groups—represented 8% of total written public comments across all meetings. More significantly, likely treatment-induced comments represented 46% of pro-housing comments, and therefore swung the balance of pro-versus-anti housing comments toward a more equal footing.

The imbalances of comment makeup highlighted by Yoder (2020) that were not corrected merely be moving to an online setting (Einstein et al. 2021) appear to have been significantly altered by the treatments. This large change in the composition of comments caused by the treatments shows that outreach can change the representation of civic bodies to be more reflective of the broader public where simple increases in accessibility may not.

Non-experimental campaigns conducted by other groups, while not directly measurable, also appear to have had large impacts on comments in some of the observed meetings. For example, the abnormally large number of Manhattan Beach City Council meeting public comments on November 2 were related to an agenda item seeking agreement on language drafted by the local History Advisory Board for a plaque acknowledging the city's racially-motivated use of eminent domain to force the sale of beachfront property owned by Black families in 1927. This agenda item became the subject of "vitriolic public criticism backed by a viral, anonymous newsletter attacking [the History Advisory Board's] work" (McDermott 2021).

Meeting	Total comments (incl. treatment induced)	Pro-housing comments (not incl. treatment induced)	Pro-housing comments (incl. treatment-induced)	Anti-housing comments (incl. treatment-induced)
Beverly Hills 10/12	19	4	6	5
Santa Monica 10/12	67	15	38	11
Whittier 10/12	4	0	4	0
Rancho Palos Verdes 10/19	121	2	3	54
Manhattan Beach 11/02	225	0	0	0
Norwalk 11/02	7	0	3	0
Sierra Madre 11/09	20	0	0	8
Culver City 12/10	71	25	31	23
Total	**534**	**46**	**85**	**101**

Table 2.1: Examination of public comments in treated council meetings

These large marginal effects of contact on overall turnout contrast sharply with, e.g., GOTV. In voter turnout settings, the large number of individuals who regularly turn out to vote makes the change in overall turnout due to campaigns relatively small. By contrast, even a few new participants in city council meetings can drastically change the composition of comments due to generally low equilibrium participation rates.

Conclusion

Understanding how to motivate individuals to engage in personally costly collective action when their gains from mobilization are long-term and uncertain is an enduring and fundamental question in political economy. Well-established research indicates how and why homeowners with direct financial payoffs participate in local politics at disproportionately high rates. However, there is little evidence to suggest how to motivate renters—who face long-term and uncertain payoffs—to overcome the collective action problem.

I contribute to our understanding of how to motivate these groups to engage in costly political behavior using 8 email-outreach field experiments encouraging renters (N = 19,951 households) to participate in local politics in the form of commenting at city council meetings. In addition, I document how these motivational campaigns changed the balance of participation in civic bodies. Three treatment arms tested the effectiveness of messages that: (1) lowered the costs of participation only, (2) primed economic self-interest, or (3) highlighted the costs of abstention. Receipt of any treatment increased public comments by 1pp, while highlighting the cost of abstention increased comments by 1.4pp. Individuals already engaged in local politics were more responsive to treatment. Treatment-induced comments represented 8% of total comments and 46% of pro-housing comments across all city council meetings. The treatments therefore overcame many of the traditional barriers to renter collective action, and changed the representation of civic bodies to be more reflective of the broader public.

These results support three main theoretical and substantive conclusions. First, the high efficacy of a treatment arm applying social pressure and highlighting the high economic costs of abstention from local politics supports Aytaç and Stokes (2019)'s theory that abstention can be perceived as highly costly by individuals. While it is difficult to pin down if the specific mechanism for this larger effect hinged on highlighting economic costs of abstention versus queuing emotions such as anger towards an out-group, the results nevertheless show that highlighting financial harm by an out-group is effective at raising the perceived costs of abstention.

Second, the large change in the composition of comments caused by the treatments shows that strategic outreach can make representation more reflective of the broader public where simple increases in accessibility do not. Pro-housing comments made up a majority of comments in over 50% of treated meetings, contrasting sharply with previous research finding majority pro-housing comments in less than 5% of online council meetings in equilibrium (Einstein et al. 2021). As these civic bodies make regular decisions that directly impact the day-to-day lives of residents and the status quo of participation is highly unrepresentative in equilibrium, outreach may have the ability to change local officials' perceptions of resident preferences toward a more representative picture.

Finally, the results show that unlike e.g., voting, email is able to increase political participation when participation is also conducted online. It therefore appears possible to meaningfully increase political participation in under-appreciated and low-turnout settings such as city council meetings using relatively low cost strategies.

Appendix

Housing supply and housing prices

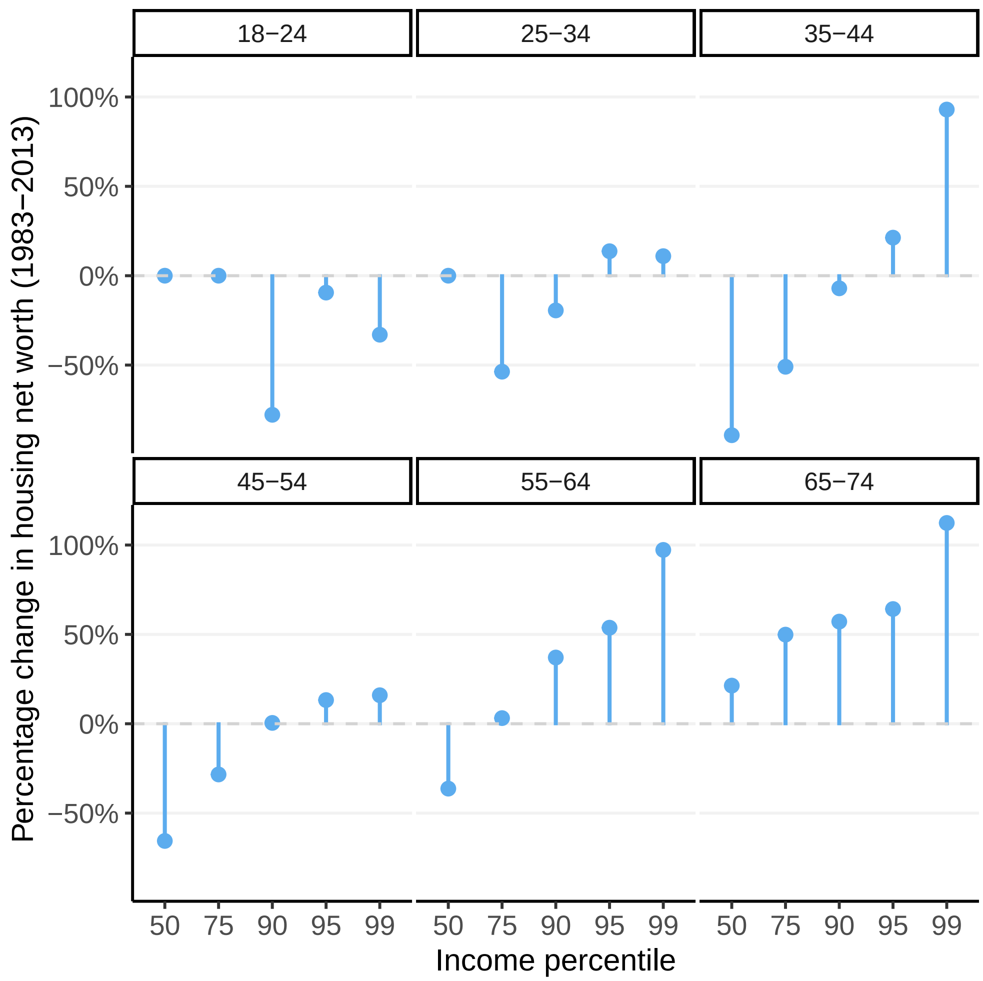


Figure 2.7: Change in housing net worth by age and income percentile

Source: Glaeser and Gyourko (2018)

Balance

	Placebo (N=2007)		Treatment (N=17944)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
gender	0.52	0.50	0.53	0.50	0.02	0.11
english	0.98	0.12	0.98	0.14	0.00	0.27
age	41.60	15.76	41.25	15.62	-0.37	0.31
yearbuilt	1964.93	18.63	1964.83	18.03	-0.14	0.75
units	34.25	64.90	34.39	66.40	0.08	0.96
dem	0.57	0.49	0.58	0.49	0.01	0.41
rep	0.13	0.33	0.11	0.32	-0.01	0.21
npp	0.24	0.43	0.24	0.43	0.00	0.73
vote_2020_general	0.79	0.40	0.81	0.40	0.01	0.28
vote_2017_municipal	0.10	0.30	0.09	0.29	-0.01	0.28
vote_2016_general	0.45	0.50	0.44	0.50	0.00	0.75

Table 2.2: Covariate balance and difference in means test: treatment vs. placebo

	Placebo (N=2007)		Treatment 1 (N=5984)		Treatment 2 (N=6002)		Treatment 3 (N=5958)	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
gender	0.52	0.50	0.52	0.50	0.54	0.50	0.54	0.50
english	0.98	0.12	0.98	0.14	0.98	0.13	0.98	0.14
age	41.60	15.76	41.16	15.61	41.35	15.63	41.23	15.62
yearbuilt	1964.93	18.63	1964.83	17.88	1964.83	18.33	1964.84	17.88
units	34.25	64.90	34.31	66.10	34.01	66.54	34.86	66.56
dem	0.57	0.49	0.58	0.49	0.60	0.49	0.58	0.49
rep	0.13	0.33	0.11	0.32	0.11	0.31	0.12	0.33
npp	0.24	0.43	0.25	0.43	0.24	0.43	0.24	0.43
vote_2020_general	0.79	0.40	0.80	0.40	0.81	0.40	0.81	0.39
vote_2017_municipal	0.10	0.30	0.09	0.29	0.10	0.30	0.09	0.29
vote_2016_general	0.45	0.50	0.45	0.50	0.45	0.50	0.43	0.50

Table 2.3: Covariate balance across all treatment groups

Treatment details

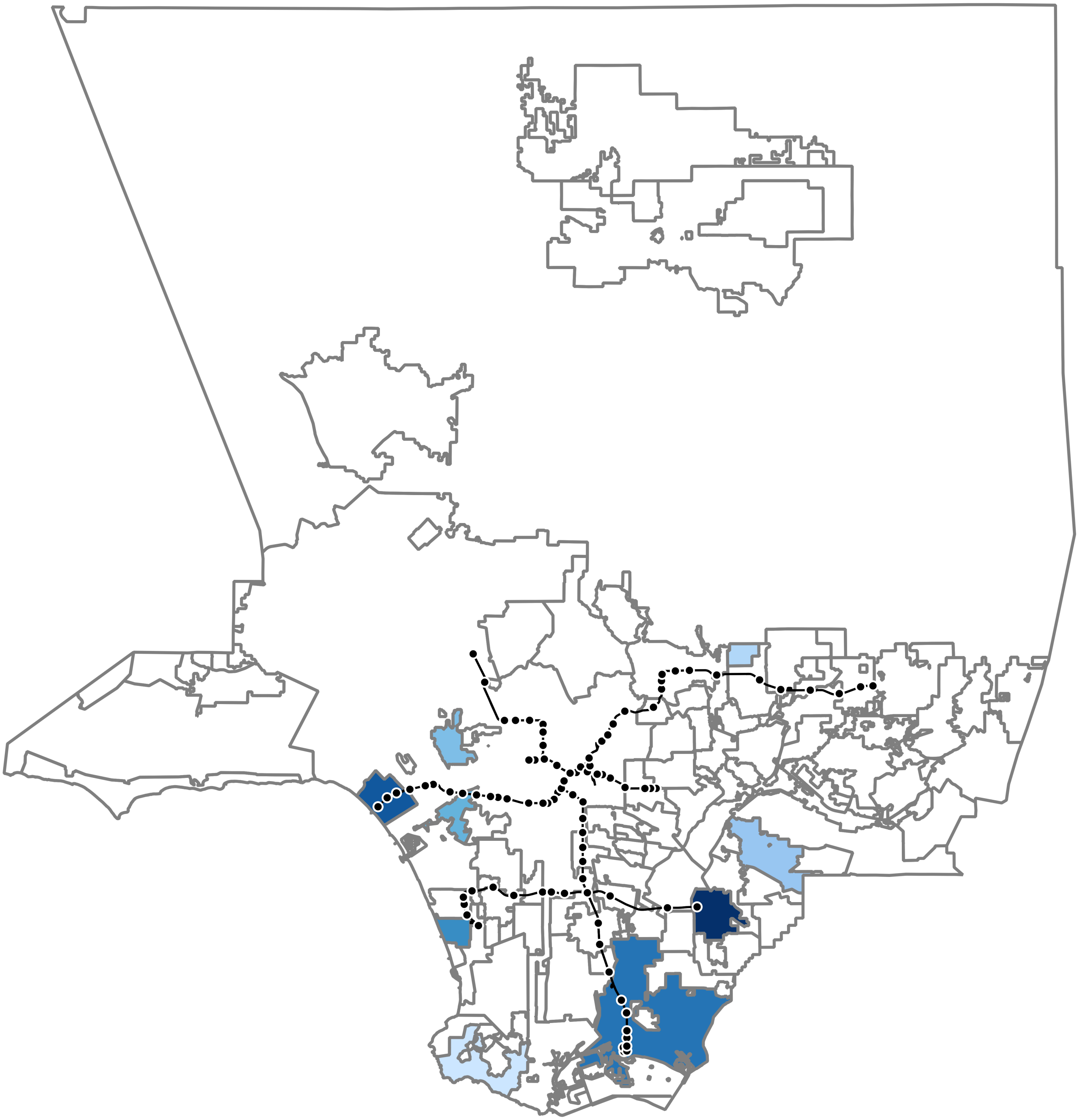


Figure 2.8: Map of cities in Los Angeles county by experiment status

Note: Cities in which an experiment was launched in blue. Cities shaded by population density. Los Angeles Metro rail lines and rail stations in black.

Sample comment

Subject:

Public comment for [DATE] council meeting agenda item [ITEM NUMBER]

Body:

Dear City Council,

I'm writing to express my concern about our affordable housing shortage and its impact on the future of our city. Exclusionary zoning and land use practices have led to an undersupply of affordable medium- and high-density housing near jobs and transit, and have perpetuated segregated living patterns and the exclusion of historically disadvantaged communities.

[CITY] has an opportunity to address the need for more housing in a way that furthers equity, environmental sustainability, and economic recovery in its housing element update. We should update the housing element in a way that encourages historically high housing growth, while furthering fair housing opportunities and undoing patterns of discrimination in housing. We can't miss this opportunity to fix our city's housing crisis.

I urge you to legalize more housing, make housing easier to build, fund affordable housing and end homelessness, and strengthen tenants' rights.

Sincerely,

FIRSTNAME LASTNAME

Tests for differential compliance

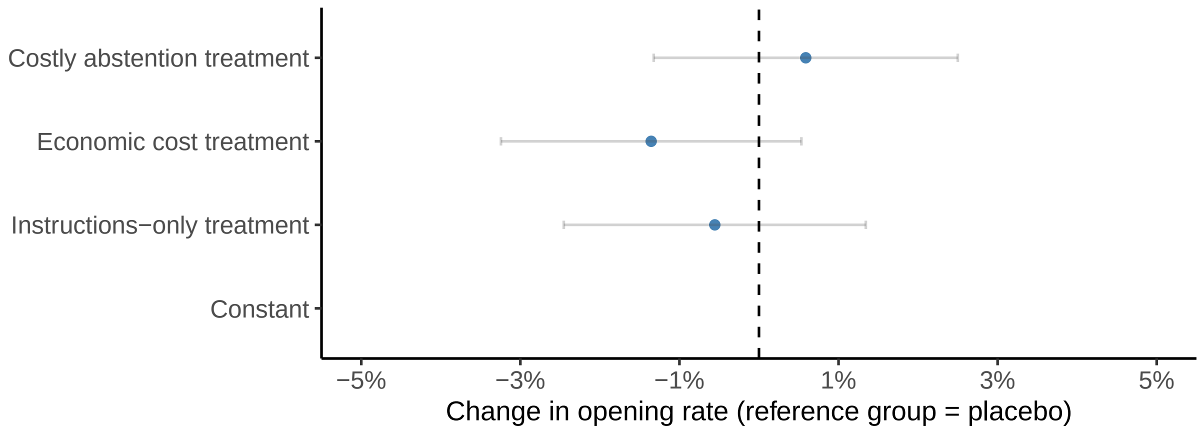


Figure 2.9: Average treatment effect on email opening, all cities

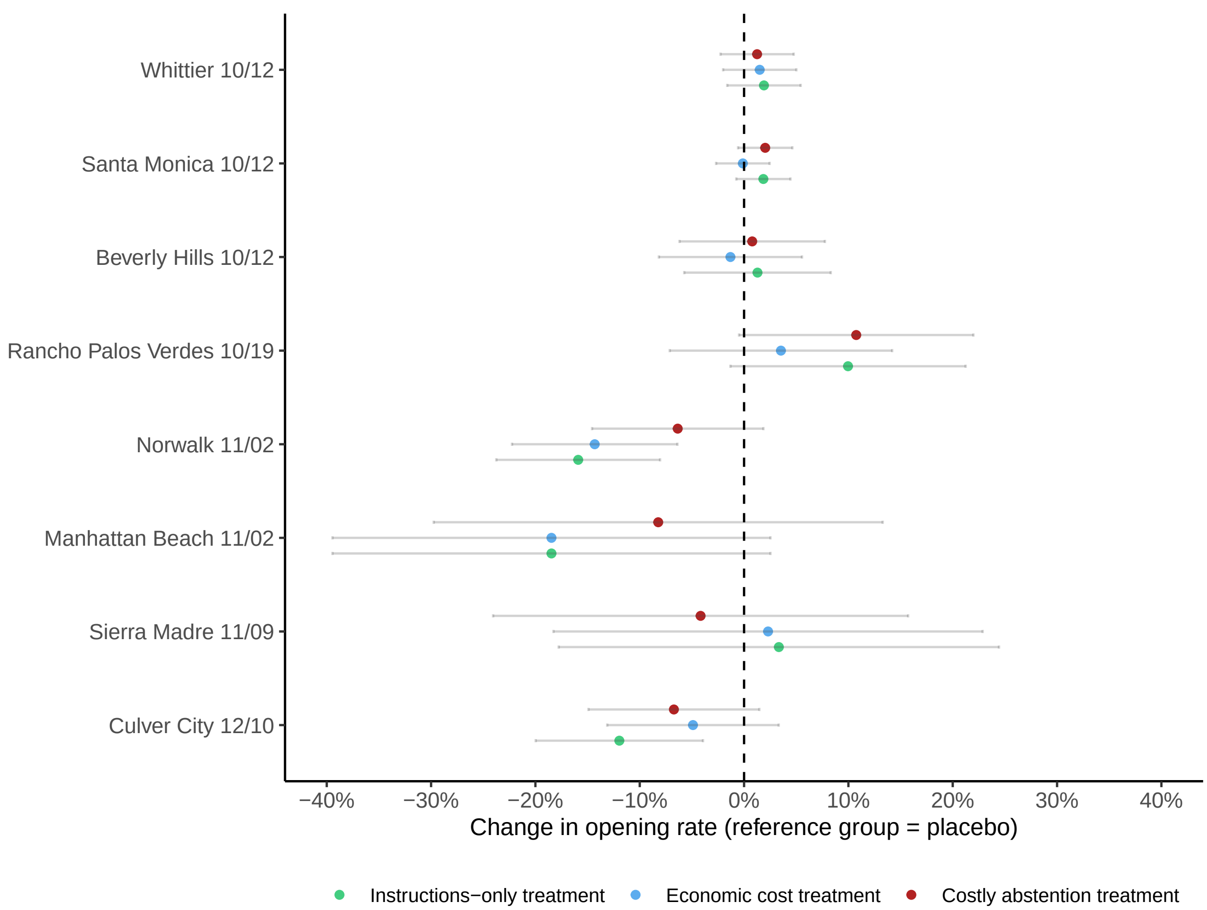


Figure 2.10: Average treatment effect on email opening, by city

	Placebo	Treatment 1	Treatment 2	Treatment 3
(Intercept)	−0.321	−0.535	−0.565	0.216
	(0.980)	(0.569)	(0.560)	(0.563)
gender	−0.028	0.004	−0.012	−0.004
	(0.017)	(0.010)	(0.010)	(0.010)
english	0.009	0.045	−0.020	−0.042
	(0.069)	(0.031)	(0.037)	(0.040)
age	0.000	0.000	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)
yearbuilt	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
units	0.000	0.000*	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
dem	0.033	0.012	0.033+	0.030
	(0.033)	(0.020)	(0.019)	(0.021)
rep	0.021	−0.008	0.003	−0.009
	(0.039)	(0.023)	(0.023)	(0.024)
npp	0.054	0.000	0.017	0.011
	(0.036)	(0.021)	(0.021)	(0.022)
vote_2020_general	0.028	0.031**	0.062***	0.030*
	(0.021)	(0.012)	(0.011)	(0.013)
vote_2017_municipal	0.041	0.057**	0.040*	0.035+
	(0.033)	(0.020)	(0.018)	(0.019)
vote_2016_general	−0.006	0.012	0.002	−0.019+
	(0.019)	(0.011)	(0.010)	(0.011)
Num.Obs.	2007	5984	6002	5958

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.4: Covariate predictiveness of compliance by treatment group

	All treatment groups vs. placebo		Individual treatments vs. placebo	
Constant	0.0005	0.0005	0.0005	0.0005
	(0.0005)	(0.0013)	(0.0005)	(0.0013)
	[−0.0005, 0.0015]	[−0.0022, 0.0031]	[−0.0005, 0.0015]	[−0.0022, 0.0031]
Treated	0.0020**	0.0020**		
	(0.0006)	(0.0006)		
	[0.0008, 0.0032]	[0.0007, 0.0032]		
Instructions-only treatment			0.0012	0.0011
			(0.0007)	(0.0007)
			[−0.0003, 0.0026]	[−0.0003, 0.0026]
Economic cost treatment			0.0021*	0.0021*
			(0.0008)	(0.0009)
			[0.0004, 0.0038]	[0.0004, 0.0038]
Costly abstention treatment			0.0026**	0.0027**
			(0.0009)	(0.0009)
			[0.0009, 0.0044]	[0.0009, 0.0044]
Covariate adjustment:	Yes	No	Yes	No
Num.Obs.	19 951	19 951	19 951	19 951

Notes: Standard errors clustered at the address level in parentheses. 95 percent confidence intervals in brackets.
+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Tabular results

Table 2.5: Intent-to-treat effects

	All treatment groups vs. placebo		Individual treatments vs. placebo	
Constant	0.0000	0.0061	0.0000	0.0063
	(0.0000)	(0.0086)		(0.0086)
	[0.0000, 0.0000]	[−0.0107, 0.0230]		[−0.0106, 0.0231]
Treated	0.0102***	0.0104***		
	(0.0018)	(0.0019)		
	[0.0066, 0.0138]	[0.0066, 0.0141]		
Instructions-only treatment			0.0054*	0.0052*
			(0.0025)	(0.0023)
			[0.0006, 0.0103]	[0.0006, 0.0098]
Economic cost treatment			0.0101**	0.0106**
			(0.0032)	(0.0033)
			[0.0039, 0.0163]	[0.0041, 0.0171]
Costly abstention treatment			0.0144***	0.0148***
			(0.0036)	(0.0037)
			[0.0073, 0.0215]	[0.0075, 0.0222]
Covariate adjustment:	Yes	No	Yes	No
Num.Obs.	3381	3381	3381	3381

Notes: Standard errors clustered at the address level in parentheses. 95 percent confidence intervals in brackets. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: Complier average causal effects

	CATE	Logit
Constant	0.006	-4.748^{***}
	(0.009)	(0.209)
Treated	0.009^{***}	
	(0.002)	
Voted in 2017 municipal election	0.000	0.973^{*}
	(0.001)	(0.414)
Treated x Voted	0.014+	
	(0.008)	
City fixed effects:	Yes	No
Num.Obs.	3381	3034
Log.Lik.		-170.607

Notes: CATE standard errors clustered at the address level.
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: Conditional complier average causal effect

Robustness

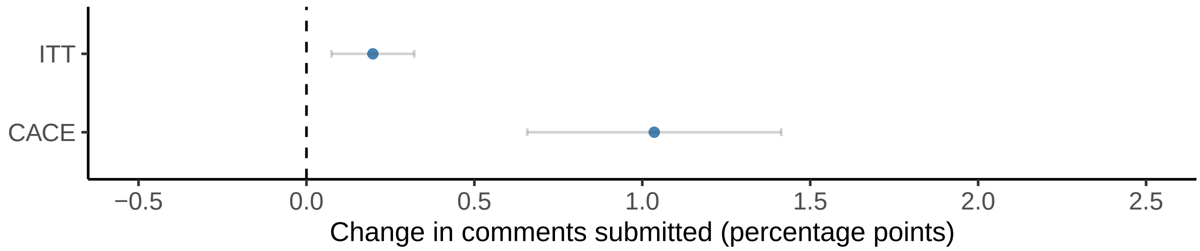


Figure 2.11: Intent-to-treat effect and complier average causal effect, all cities (without covariate adjustment)

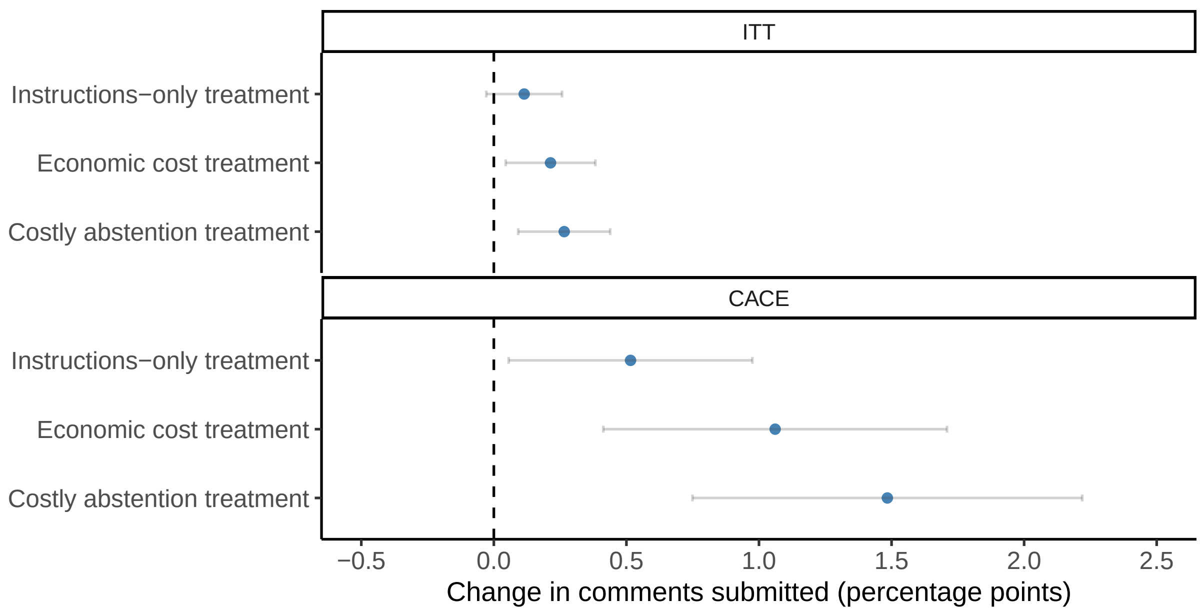


Figure 2.12: Effects by treatment group, all cities (without covariate adjustment)

	All treatment groups vs. placebo		Individual treatments vs. placebo	
	ITT	CACE	ITT	CACE
Constant	−7.1987***	−6.5439***	−7.1987***	−6.5439***
	(0.8170)	(1.4173)	(0.8170)	(1.4173)
	[−9.3648, −5.9318]	[−11.3781, −4.6301]	[−9.3648, −5.9318]	[−11.3781, −4.6301]
Treated	1.2239+	1.9864*		
	(0.8304)	(1.4285)		
	[−0.0850, 3.4045]	[0.0265, 6.8285]		
Instructions-only treatment			0.8548	1.3414
			(0.8735)	(1.4804)
			[−0.5931, 3.0816]	[−0.8391, 6.2197]
Economic cost treatment			1.3048+	2.0372+
			(0.8534)	(1.4509)
			[−0.0776, 3.5102]	[−0.0157, 6.8950]
Costly abstention treatment			1.4797*	2.3874*
			(0.8479)	(1.4388)
			[0.1150, 3.6792]	[0.3850, 7.2367]
Num.Obs.	19 951	3381	19 951	3381

Notes: Standard errors clustered at the address level in parentheses. 95 percent confidence intervals in brackets.
+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: ITT and CACE estimates from penalized maximum likelihood

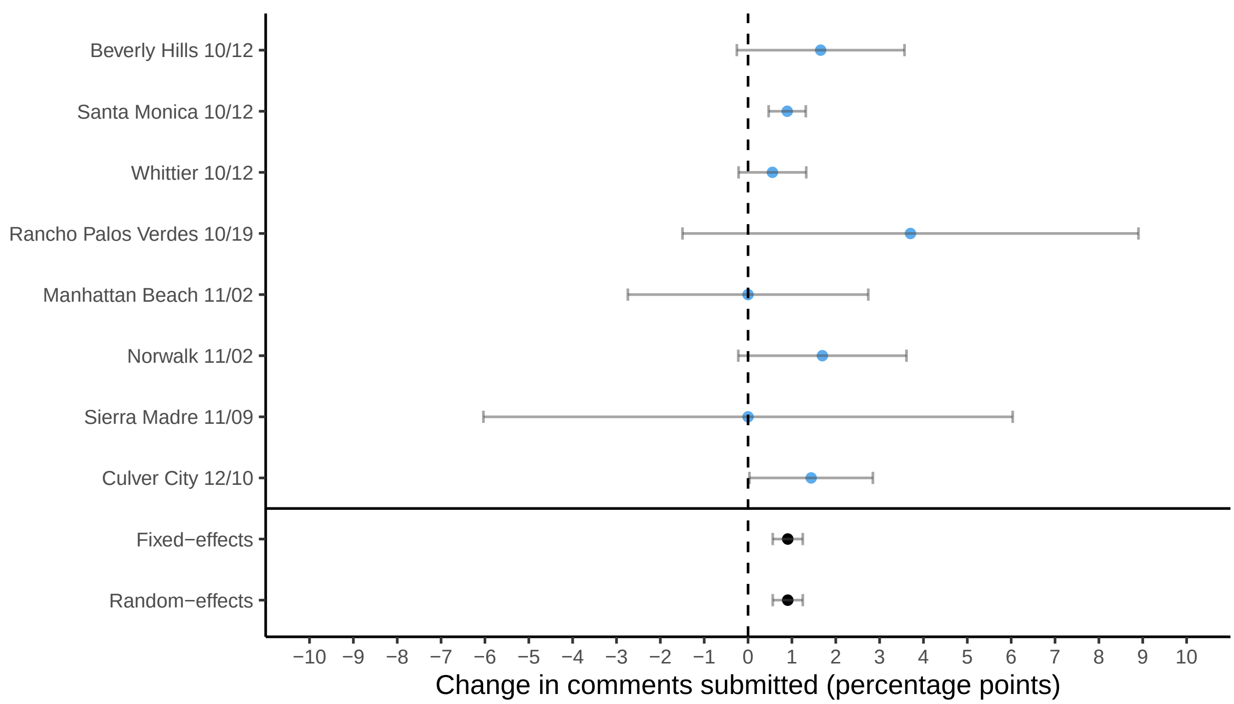


Figure 2.13: Meta-analysis of complier average causal effects by city, excluding pilot studies

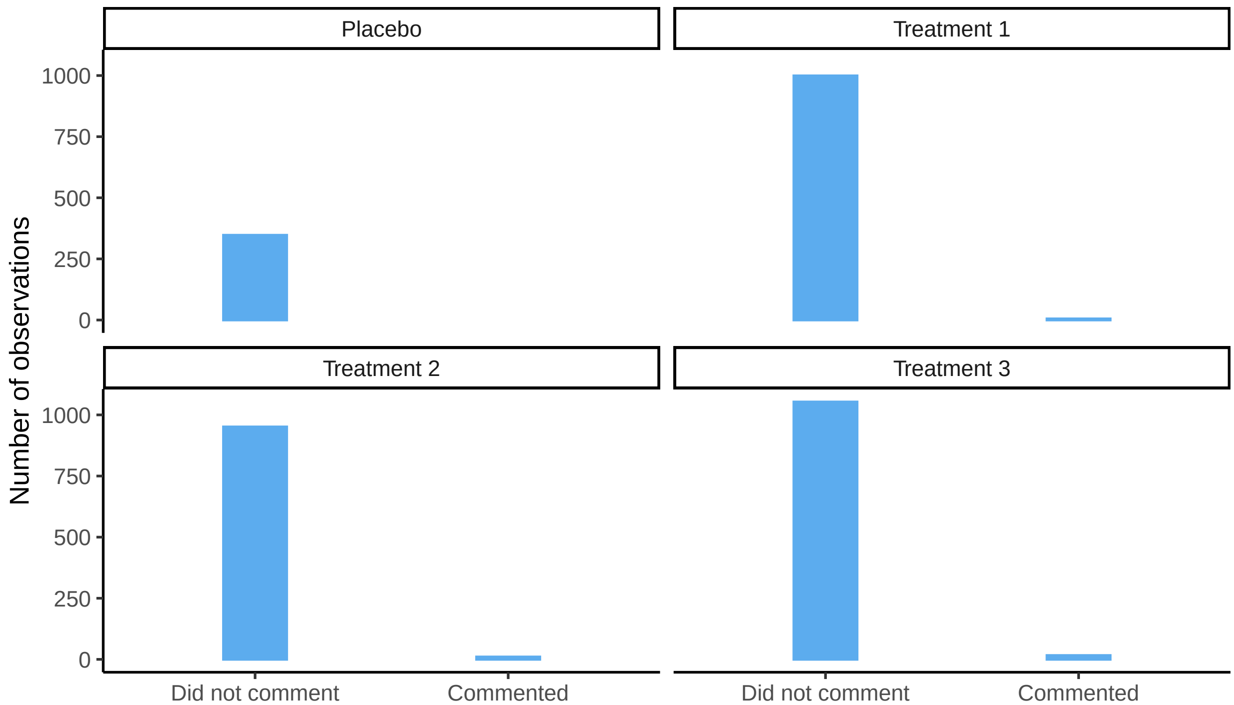


Figure 2.14: Distibution of outcomes by treatment group (compliers only)

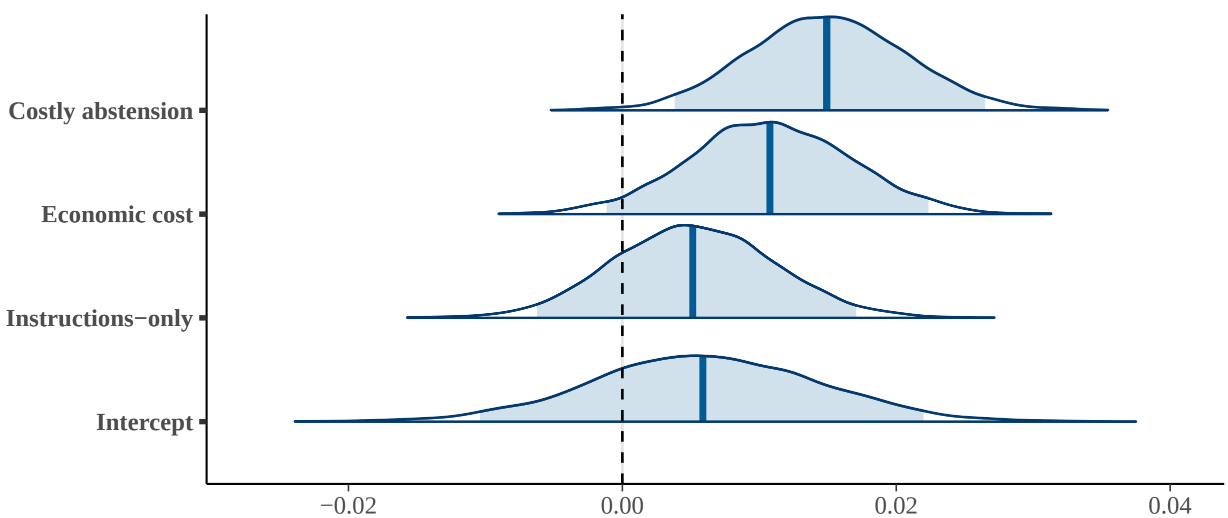


Figure 2.15: Bayesian multilevel model: coefficient estimates and posterior distributions (includes city fixed effects)

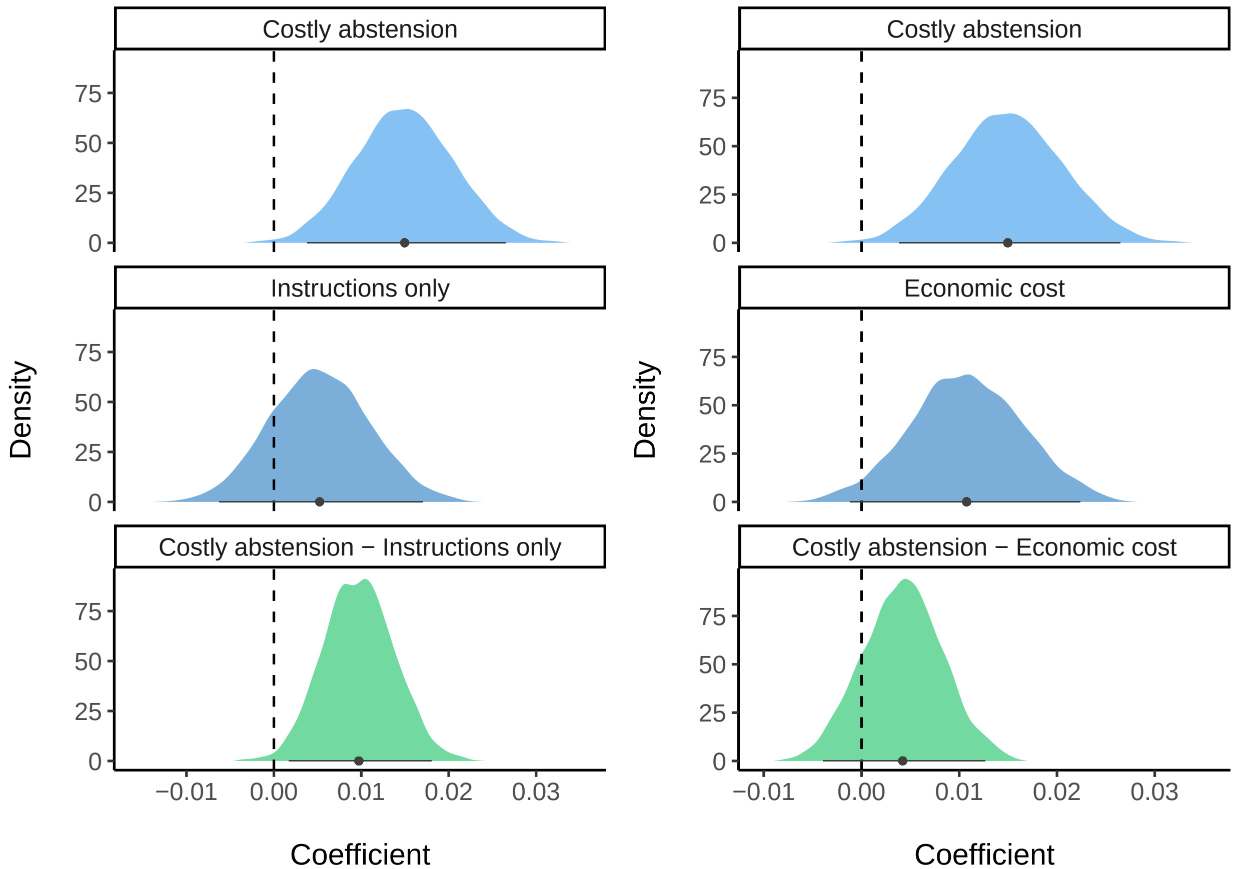


Figure 2.16: Posterior distributions of costly abstention treatment, instructions only treatment, and difference

chapter 3

Amakudata: a new dataset of bureaucratic revolving door hires

With Sayumi Miyano, Diana Stanescu, and Hikaru Yamagishi

Abstract: Political economists have long speculated about the effects of connections between bureaucracies and the private sector. However, data tracing flows of civil servants from the bureaucracy to the private sector remains rare. This article presents a new dataset, *Amakudata*, which contains individual-level data of virtually *all* Japanese bureaucrats retiring into positions outside of the bureaucracy from 2009 to 2019. We first present how the dataset was created and validated. Next, we describe what the data illuminates about the revolving door in Japan and beyond. We conclude by discussing how the data can be used to investigate empirical and causal questions in diverse subjects such as corruption and regulatory capture; procurement, pork, and government waste; bureaucratic representation; and international political economy.

Introduction

A growing literature examines the causes and consequences of revolving door hiring between the government and private sectors. However, a lack of data on bureaucratic personnel movements has hampered empirical analysis of bureaucratic connections. Past empirical work has therefore largely focused on legislators—for whom data is more readily available—or relied on survey samples for bureaucratic connections.

We leverage a 2008 reform to the Japanese National Public Service Act—which requires ministries to report bureaucrats' re-employment destinations (*National Public Service Act* 1947, Articles 106-23 & 106-25)—to create a dataset of all revolving door appointments in Japan from 2009 to 2020. To the best of our knowledge, this represents the first systematic dataset of all revolving door hires from the bureaucracy to industry in any country. The data includes information on former place of employment and former title, as well as subsequent position and title. The name of the dataset, *Amakudata*, is a reference to *amakudari*—Japanese for "descent from heaven"—the common practice of civil servants "retiring" from the bureaucracy into lucrative outside positions at advanced stages of their careers.

Amakudata was created by digitizing over 1,000 pages of PDF reports of bureaucratic rehiring to create a list of over 13,000 instances of bureaucratic re-employment over roughly one decade. These reports are compiled by the cabinet office and released annually, and are accessible from the website of the Japanese Cabinet Secretariat.[1] The data was then extensively computationally and manually cleaned, and unique identifiers added to facilitate merging with additional databases—including firm-level data—for empirical analysis.

We first present how the dataset was created. Next, we discuss how the data comports with previous theories of bureaucratic personnel flows within and outside of Japan. Finally, we examine how the data can be used to answer empirical questions on topics such as corruption, regulatory capture, pork barrel spending, bureaucratic representation, international

[1] https://www.cas.go.jp/jp/gaiyou/jimu/jinjikyoku/jinji_j.html

trade, and government waste. We also provide an interactive online dashboard which can be used to explore the data.

A brief review of revolving door literature

A large empirical literature shows how connections to political office provide benefits to firms (Blanes i Vidal, Draca and Fons-Rosen 2012; Boas, Hidalgo and Richardson 2014; Faccio 2006; Faccio, Masulis and McConnell 2006; Khwaja and Mian 2005; Truex 2014). Analyses of the political revolving door are notably aided by databases such as Lobbyview in the United States (Kim 2018), and more recently the Chinese Revolving-Door Officials Database (Li 2021).

A large theoretical literature models the impact of bureaucratic revolving door hiring practices on regulatory outcomes. Researchers have argued both that the revolving door leads to regulatory capture, and leads to improved regulatory outcomes as regulators attempt to signal competence to potential future employers. See Dal Bó (2006) for an excellent overview.

Empirically, analyses of the bureaucratic revolving door have been hampered by a lack of data. Exceptions include recent papers by Asai, Kawai and Nakabayashi (2021), Barbosa and Straub (2020), and Lee and You (2020), who conduct novel data collection efforts to examine bureaucratic ties to single agencies or industry sectors in Japan, Brazil, and the United States, respectively.

Within Japan, the revolving door is well-known in both academia and amongst the general public. The revolving door—or *amakudari*—has been blamed for multiple regulatory and policy failures,[2] as well as inability to enact structural economic reforms. However, despite decades of theoretical and qualitative research on *amakudari*, systematic analysis has been hampered by a lack of data. We note that *amakudari* is not typically "revolving"—retired bureaucrats do not often return to the bureaucracy, nor does the bureaucracy often hire

[2] E.g., the savings and loan bailout (Carlson and Reed 2018; Mishima 2013), the HIV-contaminated blood scandal (Carlson and Reed 2018; Mishima 2013), and the Fukushima Daiichi nuclear plant disaster (Diet of Japan 2012; Mishima 2013).

mid-career private sector employees. We therefore use the term "revolving door" to speak to the literature addressing connections between private and public sector positions.

Dataset creation

Why do these data exist?

Pressure to regulate *amakudari* culminated in reforms to the National Public Service Act drafted under the government of Prime Minister Junichiro Koizumi and enacted in 2008 (Kato 2017; Mishima 2013; Terada 2019). This reform requires civil servants above a certain rank to notify the Cabinet Office of their re-appointments (*National Public Service Act* 1947, Article 106-23). All notifications are therefore reported publicly at the end of each year (*National Public Service Act* 1947, as per Article 106-25) on the website of the Japanese Cabinet Secretariat.[3] While previous data linking *amakudari* bureaucrats to retirement destinations does exist, these data are not publicaly available and are limited to personnel sampled from individual firms (Horiuchi and Shimizu 2001) or ministries (Asai, Kawai and Nakabayashi 2021).

Dataset creation and validation

The Japanese Cabinet Secretariat uploaded over 1000 pages of PDFs reports documenting civil servant re-employment to their website over the past decade as per the regulations outlined above (see Figure 3.5 for an example). We collected these PDF documents, digitized them using OCR software, and conducted extensive manual and programmatic cleaning to ensure individuals were properly matched with their places of employment. This process resulted in a list of over 13,000 instances of bureaucratic re-employments over roughly one decade.

To validate the accuracy of the data, we randomly sampled 300 *amakudari* hires from executive positions in private sector companies, and ensured that they were listed in orga-

[3] https://www.cas.go.jp/jp/gaiyou/jimu/jinjikyoku/jinji_j.html

nization records in the appropriate position. As part of a separate analysis, we also verified the appointments of all vice-ministerial and assistant vice-ministerial appointments in the data, which are typically reported in the economic newspaper *Nikkei*.

Variables

The raw source PDFs contain the following information pertaining to the bureaucrat: full name, age, position title at retirement, date of retirement, date of re-employment, place of re-employment, duties at place of re-employment, and industry of re-employment. We retain all of these variables. A major advantage of this dataset is therefore that it contains information of each retiring bureaucrat as well as each private firm/organization in which they are employed. This allows us to retrieve information both at the individual bureaucrat level as well as those at the firm/organization level for each hire.

In order to facilitate empirical analysis, we add additional categorical variables pertaining to both bureaucratic and post-bureaucratic employment. Civil servants are required to indicate their position and place of government employment. However, this is not required to be in a standardized format. We therefore constructed variables identifying the government ministry and agency of employment with a list of roughly 400 regular expression string matches at the agency level, then mapped each agency to its respective ministry. We also created indicator variables for the highest level positions of vice-minster and assistant vice-minister using string matches.[4]

Next, we created variables identifying the type of post-bureaucratic employment. We indicate whether the new place of employment is: (1) a private corporation, public corporation, or government entity, (2) a for-profit or non-profit firm, (3) a stock, non-stock, intermediary, or public interest corporation, and (4) the specific firm type (e.g., stock company, LLC, foundation, credit cooperation, educational institution, etc.). All indicators were created with regular expression matching, followed by manual additions where multiple or

[4]We do not categorize below this level as position titles below this level are not standardized across ministries. Classifications across ministries would therefore be erroneous, and should be examined on a ministry-specific basis.

no matches were returned.

Connections to other data sets

We also facilitate connections with third-party databases such as the *Nikkei NEEDS* financial database and stock market databases such as *Yahoo Finance*. While Nikkei's financial data is proprietary, all firms that exist in both *Nikkei NEEDS* and *Amakudata* contain Nikkei's unique identifier code to facilitate ease of merging. This identifier was added through an exact match merge of firm names with the *Nikkei NEEDS* database, followed by manual cleaning and standardization of all remaining private firms to ensure that all possible matches were made. We also provide industry information and stock ticker information for all publicly traded firms that are also in the *Nikkei NEEDS* database.

Table 3.1 provides an example of some key variables and identifiers. An interactive online dashboard is also available that allows user to explore appointments by date, employer type, employer, ministry, agency, position, etc.

Table 3.1: Amakudari dataset example

date_ret	agency	ministry_short	firm_dest_en	firm_type1_en	tse_code
2015-04-01	Police	-99	Murata Manufacturing	Stock company	6981
2015-04-16	Labor Standard Bureau	MHLW	Nationwide Labour Insurance Association	Incorporated association	-99
2014-09-01	Technology Policy Coordination	MLIT	Nippon Steel	Stock company	5401
2011-05-01	Civil Aviation Bureau	MLIT	Kajima Road	Stock company	-99
2018-04-01	Police	-99	Japan Disaster Prevention Communication Association	Foundation	-99
2016-01-01	Minister's Secretariat	MLIT	Japan Ship Equipment Certification Association	Foundation	-99
2014-04-01	Police	-99	Taiko Bank	Stock company	8537
2012-10-01	Patent Office	METI	Pasona Group	Stock company	2168
2016-07-19	Minister's Secretariat	MOF	Keio University	Educational institution	-99
2016-09-01	Kanto Environmental Office	MOE	Hibiya Amenis	Stock company	-99
2018-11-01	Hokkaido Bureau	MLIT	Nippon Koei	Stock company	1954
2013-09-01	Japan Customs	MOF	Kobe Steel	Stock company	5406
2015-08-01	Regional Transportation Bureau	MLIT	Japan Automobile Dealers Association	Incorporated association	-99
2012-06-01	Minister's Secretariat	MHLW	Nippon Access	Stock company	-99
2018-06-16	Regional Transportation Bureau	MLIT	Kyushu Land Transportation Association	Foundation	-99
2012-07-01	Accounts Center	MOF	Organization For Promiting Urban Development	Foundation	-99
2014-06-02	Minister's Secretariat	MOF	Tokyo Shinkin Bank	Shinkin bank	-99
2015-10-15	Minister's Secretariat	MOF	Elecom	Stock company	6750
2015-06-01	Hokkaido Development Bureau	MLIT	Hokkaido Road Management Engineering Center	Foundation	-99
2013-06-14	Financial Services Agency	-99	Bank Of Japan	Bank of Japan	-99

Note: Does not include all variables.

Reexamining bureaucratic connections

Theories of the revolving door

Much research on the revolving door focuses on its impact on regulatory outcomes or the usefulness of state connections for obtaining lucrative government contracts. Our data provides suggestive evidence of the importance of both regulatory and contract connections, as highly regulated industries (e.g., banks, insurance, and utilities) and those reliant upon public contracts (e.g., construction and transportation) are overrepresented in our data relative to the overall economy (see Figure 3.1). Industries also tend to hire from the ministries that regulate them or generate their contracts (see Figure 3.6).

Industry	Count amakudari	Percent amakudari	Percent economy	Difference
Services	401	13.3	12.8	0.5
Banks	251	8.3	2.2	6.1
Construction	248	8.2	4.5	3.8
Land Transportation	237	7.9	1.7	6.2
Insurance	210	7.0	0.4	6.6
Electric Appliances	182	6.0	6.3	-0.3
Wholesale Trade	163	5.4	8.4	-3.0
Electric Power & Gas	129	4.3	0.6	3.7
Warehousing and Harbor transportation	118	3.9	1.0	2.9
Information & Communication	112	3.7	13.1	-9.4

Figure 3.1: Top 10 *amakudari* destinations vs. overall economy

Our data also illuminates aspects of revolving door hiring that are under-appreciated in previous literature. For example, most research on the revolving door focuses on private sector hiring. However, in the Japanese case the number of non-profit hires is roughly equal to for-profit hires (see Figure 3.2).[5] Many of these non-profits regularly hire large numbers of bureaucrats from a single ministry (Figure 3.4), suggesting that there are direct channels from the bureaucracy to non-profit organizations, many of which receive large amounts of government assistance.

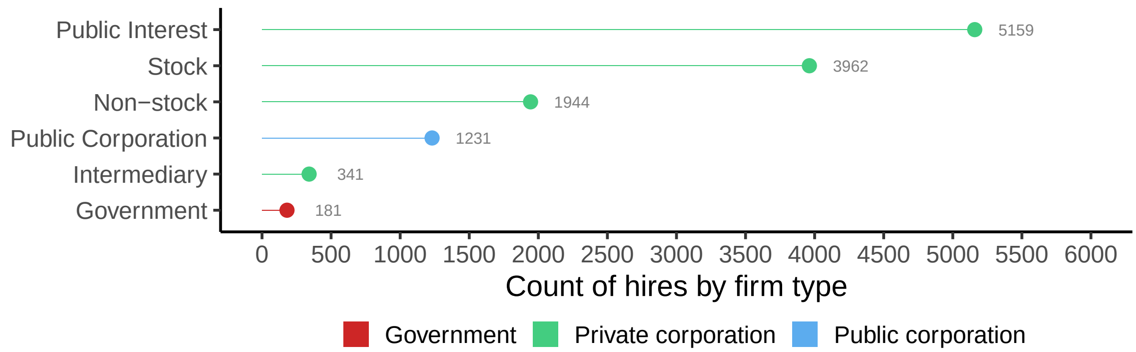


Figure 3.2: *Amakudari* destinations by firm type

Theories of amakudari in Japan

Amakudari has played a prominent role in Japanese politics and policy research over the past half century (e.g., Asai, Kawai and Nakabayashi 2021; Calder 1989; Carlson and Reed 2018; Horiuchi and Shimizu 2001; Johnson 1982; Mizoguchi and Van Quyen 2012; Rosenbluth 1989; Rosenbluth and Thies 2010; Schaede 1995; Vogel 2006). Our data allows us to precisely answer questions about this phenomenon that previously required speculation.

First, we can see which ministries place the most officials in post-retirement positions. In absolute terms, it comes as little surprise that the Ministry of Land, Infrastructure, and Transport (MLIT) placed by far the most officials over the past decade (see Figure 3.3). MLIT is the second largest ministry by number of employees (after the Ministry of Justice), and controls many lucrative government infrastructure contracts. More surprising is that the number of hires per year from MLIT increased from fewer than 100 per year in 2009 and 2010, to over 300 in 2018 (see Figure 3.7). Adjusted for number of employees, the

[5] "Public interest corporations" are a type of non-profit under the Japanese Law to Promote Specified Nonprofit Activities.

prestigious Ministry of Economy, Trade, and Industry (METI) places the most bureaucrats into *amakudari* positions (see Figure 3.8), and its officials appear to be the most valued to investors based on stock market reactions to high level METI hires (see Figure 3.9).

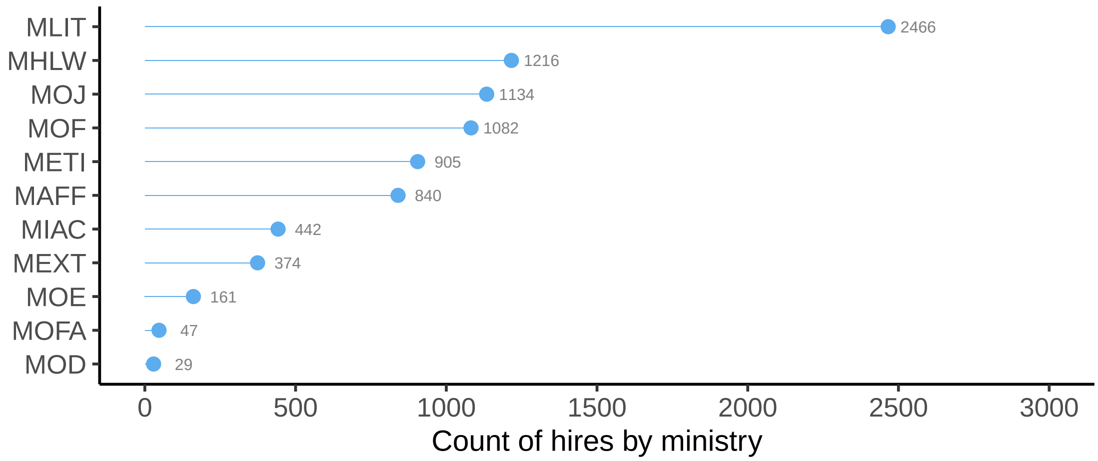


Figure 3.3: *Amakudari* ministry of origin

Second, we can see exactly where former bureaucrats go after they leave the civil service. Figure 3.2 shows that over one third of bureaucrats join public interest corporations. The top public interest corporations in terms of number of hires draw retirees primarily from a single ministry (Figure 3.4). By contrast, stock corporations and non-stock corporations tend to draw from multiple ministries, and hiring is dominated by highly regulated industries such as insurance. While large publicly traded firms tend to draw from multiple ministries, some industries do draw overwhelmingly from particular ministries. For example, the construction and transportation sectors drew the majority of their hires from the Ministry of Land, Infrastructure, and Transport (MLIT), finance drew the majority of its hires from the Ministry of Finance (MOF), and information and communications drew the majority of their hires from the Ministry of Internal Affairs and Communications (MIAC) (see Figure 3.6). Publicly traded firms that hire former bureaucrats also differ from those that do not across numerous financial metrics. For example, they tend to be larger in terms of employees, assets, and liabilities, but also have lower return on investment (see Figure 3.10).

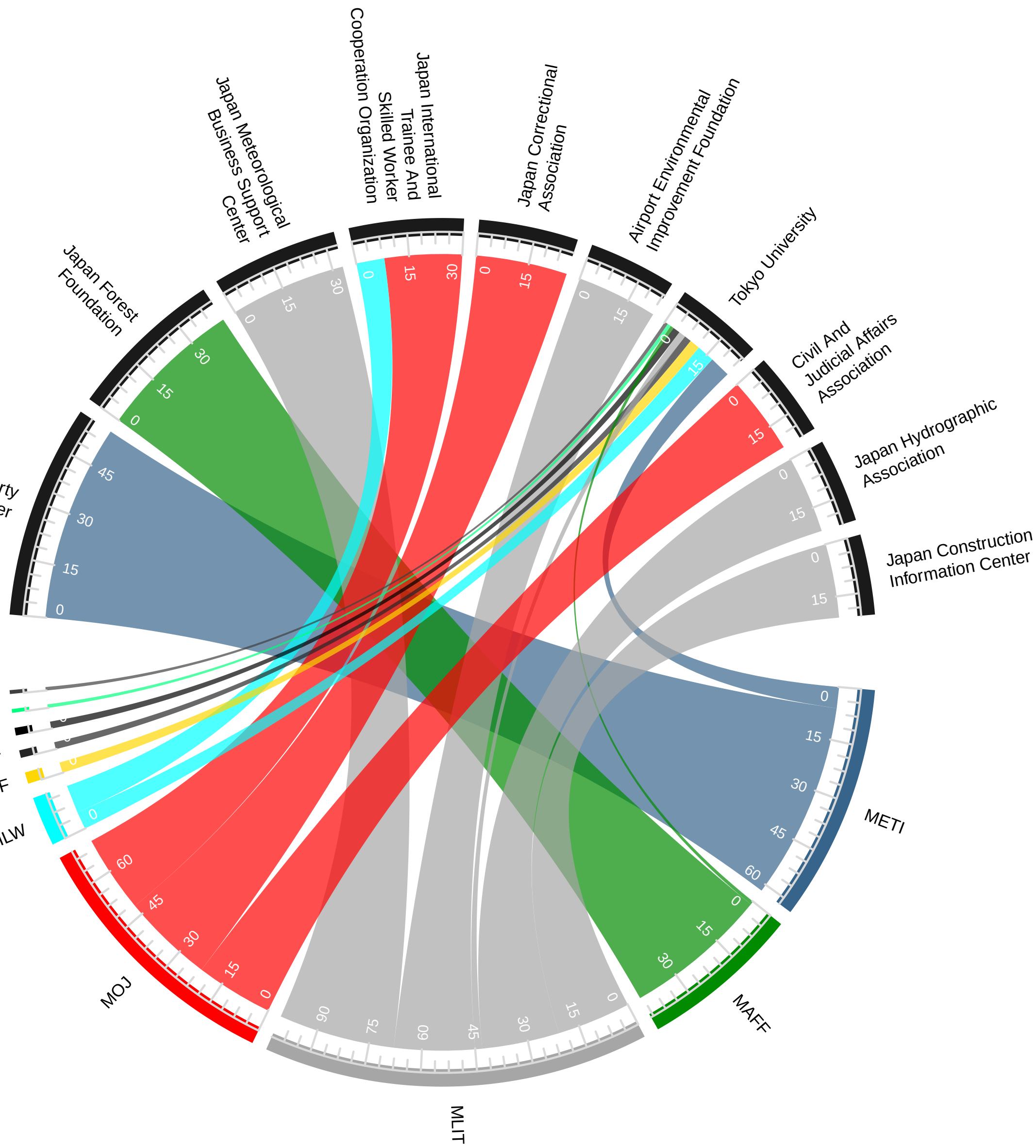


Figure 3.4: Flows of bureaucrats from ministries to top 10 public interest corporations (by number of hires)

Next, we can examine who is being hired, when, and where. Our data confirms that *amakudari* is more of an exit door than a revolving door. Japan has a mandatory "retirement" age of 60 for civil servants, and most civil servants leave at this age. However, the more prestigious ministries most heavily involved in regulation (e.g., MOF and METI) tend to have younger and wider retirement age distributions then their counterparts (see

Figure 3.11). The most common posts bureaucrats take in for-profit companies are tax advisors, consultants, auditors, lawyers, board members (internal and external), and executives. The data suggests that there is a bifurcated market for bureaucrats in which high-ranking officials are more likely to retire into large, publicly traded corporations than their subordinates. 54% of Vice-Ministers and 47% of Assistant Vice Ministers retire into publicly traded corporations, while 30% of Vice-Ministers and 38% of Assistant Vice Ministers retire into public interest corporations. By contrast, only 30% of lower ranked officials retire into publicly traded firms, while 40% retire into public interest corporations. Virtually no high ranking officials join positions in government or privately owned firms, while these organizations represent 10% and 17% of destinations for lower-ranked officials, respectively. Top officials are also more likely to retire into larger and more profitable publicly traded firms than their counterparts, as measured by number of employees, assets, liabilities, revenue, profits, return-on-investment, and EBITDA (see Table 3.2). Investors appear to monitor and reward firms for these top hires, as internal hires of vice ministers cause stock price boosts in the days following hiring announcements (see Figure 3.12).

Finally, our data allows us to speak to both recent scandals and the more pernicious aspects of *amakudari* with greater clarity than previously possible. For example, our data provides evidence that civil servants may be able to skirt rehiring regulations prohibiting former bureaucrats from immediate employment into sectors they previously regulated. 28 MOF officials retired into private sector banks since regulations were passed. 116 retired into regional credit unions known as *shinkin* banks, including 90 from regional finance bureaus. A further four officials retired into *shinkin* banks from their direct regulator—the Financial Services Agency.[6] Consistent with a scandal in which the education ministry was accused of helping its employees find jobs at universities, we see 100 appointments from the education ministry to universities between 2009-2017. This practice did not end after the scandal came to light in early 2017, with a further 12 officials entering universities in 2018 (Figure 3.13).

[6]Note that the Financial Services Agency is not located within the Ministry of Finance.

Empirical applications and future research

Corruption, capture, and favoritism

While *amakudari* is not typically envisioned as explicit corruption,[7] it is often described as such (Mizoguchi and Van Quyen 2012). However, cases in which *amakudari* does represent explicit corruption—e.g., through explicit *quid pro quos* or bribery—remain understudied.

Amakudari is more typically conceptualized as a conflict of interest that manifests through mechanisms such as regulatory capture and favoritism. Bureaucratic links have been hypothesized to impact regulation in two primary fashions: (1) explicit capture whereby regulations are modified in favor of firms over public good, and (2) as a method of easing costs for companies in financial trouble (Horiuchi and Shimizu 2001; Mizoguchi and Van Quyen 2012). Linking rehiring data with regulatory enforcement data could shed light on this issue.

A growing literature shows that connections to political office often provide tangible benefits to firms (Blanes i Vidal, Draca and Fons-Rosen 2012; Boas, Hidalgo and Richardson 2014; Faccio 2006; Faccio, Masulis and McConnell 2006; Khwaja and Mian 2005; Truex 2014). Fewer works have examined the impact of bureaucratic connections on firm performance. Exceptions include Lee and You (2020), who find that firms with bureaucratic connections allow for decreased lobbying spending; Barbosa and Straub (2020), who show that medical supply firms that hire former civil servants offer lower prices to the government; and Asai, Kawai and Nakabayashi (2021), who find increases in government contract awards for firms tied to 281 former civil servants from Japan's land ministry. Our data should allow for similar and more systematic investigations into the benefits of bureaucratic ties, such as contracts, government loans, and stock price boosts.

[7] Rather than grand corruption, petty corruption, or electoral fraud, *amakudari* fits within Rose-Ackerman (2018)'s conception of more ambiguous cases such as conflicts of interest that require "reorganization of government institutions and their relationship to the private sector" to address.

International political economy

Implications of bureaucratic ties travel beyond the sphere of domestic politics. Research on international political economy may also benefit from closer attention to firms' political connections, including the ones through bureaucratic channels. Bureaucrats play a central role in the foreign economic policymaking process, as they negotiate and implement a large swath of trade and investment policies. Given these policies' potential impact on firms' exporting, importing, and foreign production activities, we can expect firms will try to influence bureaucrats in this arena as well, in an attempt to bring policies closer to their preferences.

A growing literature in international political economy regards firms as primary actors behind politics of international economic transactions, such as trade, foreign investments and capital flows, and immigration (Kim and Osgood 2019; Kim, Liao and Miyano 2021; Peters 2017). In contrast to the earlier studies that considered cleavages at the factor of production level (e.g. capital versus labor) or at the sector/industry level as main sources of political contention, recent firm-level theories argue that heterogeneity among firms, which exists even within each industry, is a major driver of politics behind international political economy. For example, an overwhelming amount of trade lobbying in the United States are conducted at the individual firm level instead of collectively at the industry level, indicating that firm preferences are affecting the pattern of U.S. tariff reductions (Kim 2017).

Our data identifying firm-level connections with the home government offers a unique and promising opportunity to empirically test these firm-level theories, which often suffer from lack of fine-grained data. In particular, this data allows us to push the frontier in understanding how nationality and home government matters even for international business activities and even for multinational firms in this age of globally fragmented productions.

Government waste

Beyond connections leading to inefficiencies in procurement and loan decisions, our data show that over 40% of bureaucrats ultimately retire into the non-profit companies (see Figure 3.2), and that non-profit hires are more likely to be lower level bureaucrats. A generous reading of this phenomenon would suggest that bureaucrats are continuing their long-standing mission of public service even after they leave government. A less generous interpretation would suggest that the bureaucracy has established government subsidized employment centers for its former workers, and/or that bureaucrats use these positions to wait out the two year post-retirement "cooling off period" during which they cannot join a company they directly regulated. Freedom of Information Act requests could be used to illuminate links between *amakudari* and government funding of public-interest corporations, while tracing civil servant flows beyond their initial appointments could illuminate if bureaucrats move into firms they used to regulate after their "cooling off period."

Bureaucratic representation

Political scientists often analyze the degree to which elected officials reflect the demographic makeup of eligible voters. The degree to which bureaucrats reflect the general public is less clear due to lack of data availability. Our data sheds light on this question, showing that the Japanese bureaucracy—or at least elite bureaucrats entering post-retirement positions—is highly unrepresentative of the general populace.

The Japanese bureaucracy is, for example, often criticized for the lack of women amongst its ranks. Our data show that women are more likely to leave the bureaucracy at early stages of their careers than men. In fact, the youngest person to leave the civil service was a 29 year old woman who left the "Gender Equality Bureau" of the Cabinet Office for the private sector. This provides suggestive evidence that women may be more likely to find bureaucratic careers unpalatable. Note that these findings are unlikely to reflect women leaving the workforce after having children, as an individual does not appear in the data unless they join a corporation after leaving the bureaucracy.

Conclusion

We present a new dataset, *Amakudata*, which contains individual-level data of virtually all Japanese bureaucrats retiring into positions outside of the bureaucracy from 2009 to 2019.

Amakudata contains detailed information on bureaucratic retirements, including: name; age; bureaucratic position title; ministry of employment; agency of employment; date of retirement; date of re-employment; place of re-employment; industry of re-employment; identifiers for private, public, and government entities; identifiers for-profit or non-profit firms; identifiers for specific firm type (e.g., stock company, LLC, foundation, etc.); stock tickers; and identifier codes to facilitate connections to commonly used financial databases.

We show that *Amakudata* can be used by researchers in a variety of fields ranging from bureaucratic representation to international political economy, and provide an interactive online dashboard which can be used to explore the data. We hope that this resource will be widely used by scholars of both Japan and economics, management, and political science more generally.

Appendix

番号	氏名	離職時の年齢	離職時の官職	離職日	再就職日（注２）	再就職先の名称	再就職先の業務内容	再就職先における地位	国家公務員法第106条の3第2項第4号の規定に基づく承認（以下「求職の承認」という。）の有無（注３）	官民人材交流センターの援助の有無（注４）
1	三木　宏	60	内閣官房内閣参事官（内閣情報調査室内閣情報集約センター主幹）	H24.3.31	H24.4.1	財団法人世界政経調査会	内外の政治、経済、社会事情等の調査研究、資料の収集	事務局長	無	無
2	村田　啓子	49	内閣府経済社会総合研究所上席主任研究官	H23.9.30	H23.10.1	公立大学法人首都大学東京	教育・研究	教授	無	無
3	井上　侑子	29	内閣府男女共同参画局推進課課長補佐	H23.12.31	H24.1.1	高砂熱学工業株式会社	設備工事事業、設備機器の製造・販売事業	法務部社員	無	無
4	齋藤　潤	60	内閣府政策統括官（経済財政分析担当）	H24.1.10	H24.1.11	学校法人青山学院	教育・研究	青山学院大学非常勤講師	無	無
5	乾　友彦	49	内閣府経済社会総合研究所上席主任研究官併任大臣官房統計委員会担当室長	H24.3.31	H24.4.1	日本大学	教育・研究	経済学部教授	無	無
6	岩城　秀裕	55	内閣府経済社会総合研究所上席主任研究官	H24.3.31	H24.4.1	野村證券株式会社	証券業	金融市場調査部次長	無	無
7	建井　順子	42	内閣府男女共同参画局調査課調査分析専門官	H24.3.31	H24.4.1	国立大学法人東京大学	教育・研究	社会科学研究所学術支援専門職員	無	無
8	大滝　宏二	60	宮内庁長官官房用度課長	H24.3.31	H24.4.1	日本赤十字社	災害救護業務、病院経営等	常勤嘱託	無	無
9	萩原　一彦	60	宮内庁管理部車馬課長	H24.3.31	H24.4.1	財団法人菊葉文化協会	皇室に関係する伝承文化、文化財等の調査研究等	経理課長	無	無

Figure 3.5: Example original PDF document

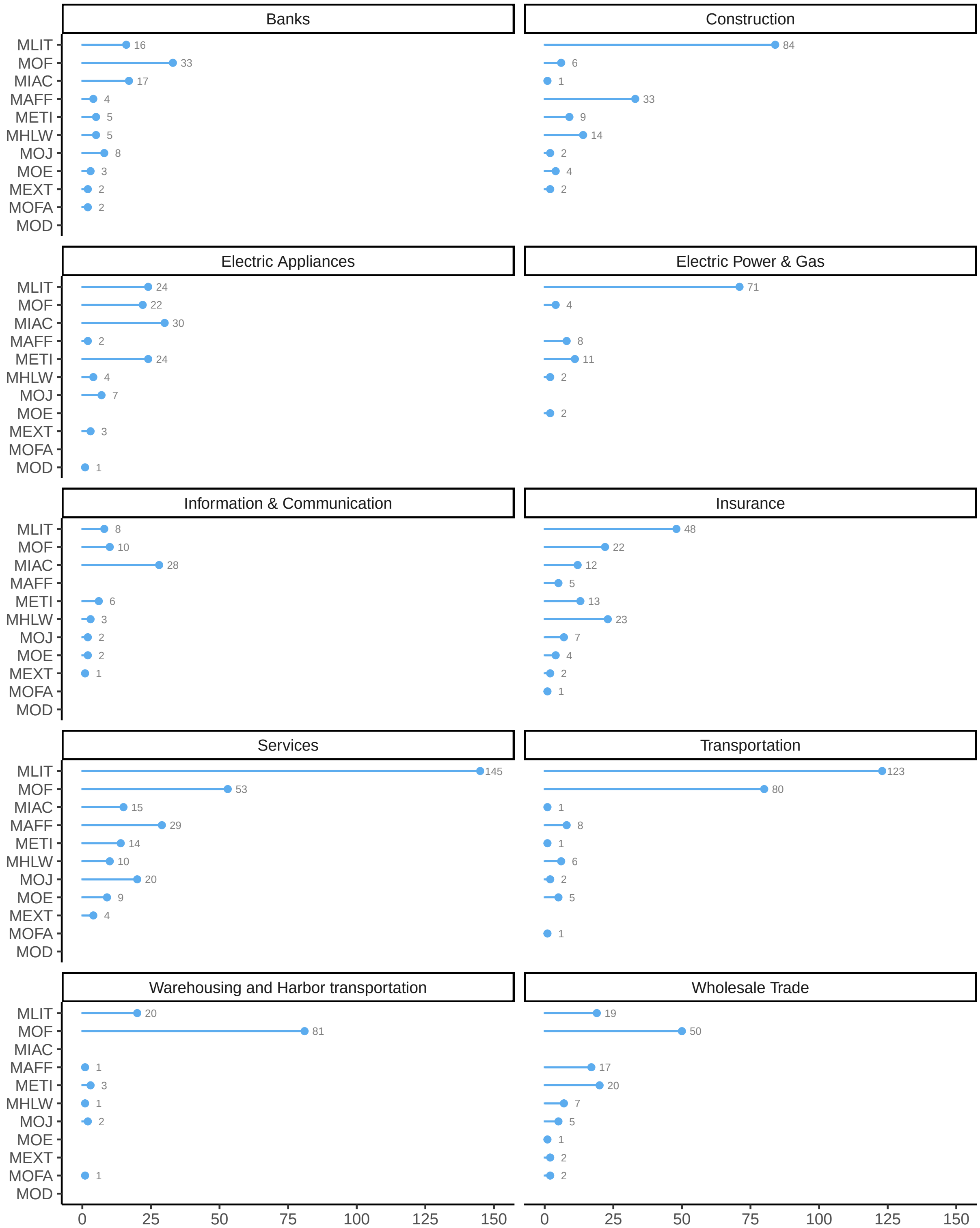


Figure 3.6: ***Amakudari*** **hires by industry and ministry (all years)**

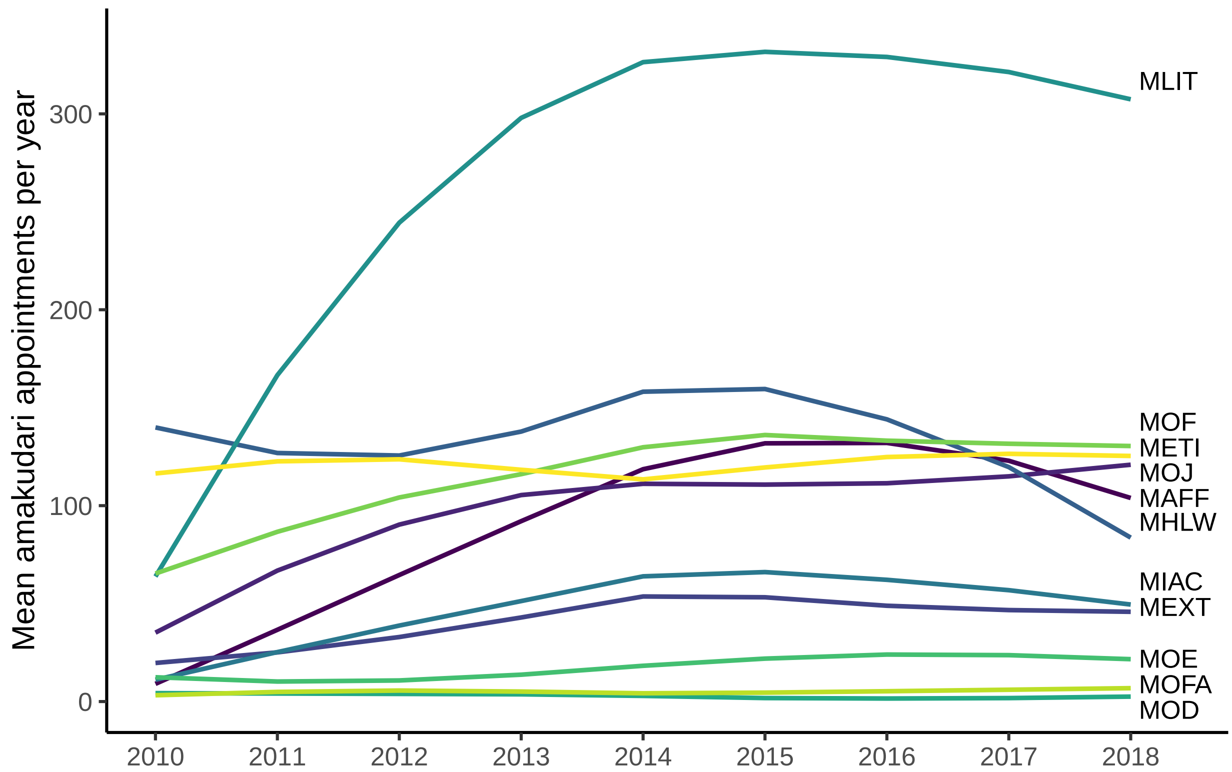


Figure 3.7: ***Amakudari*** **appointments by ministry over time**

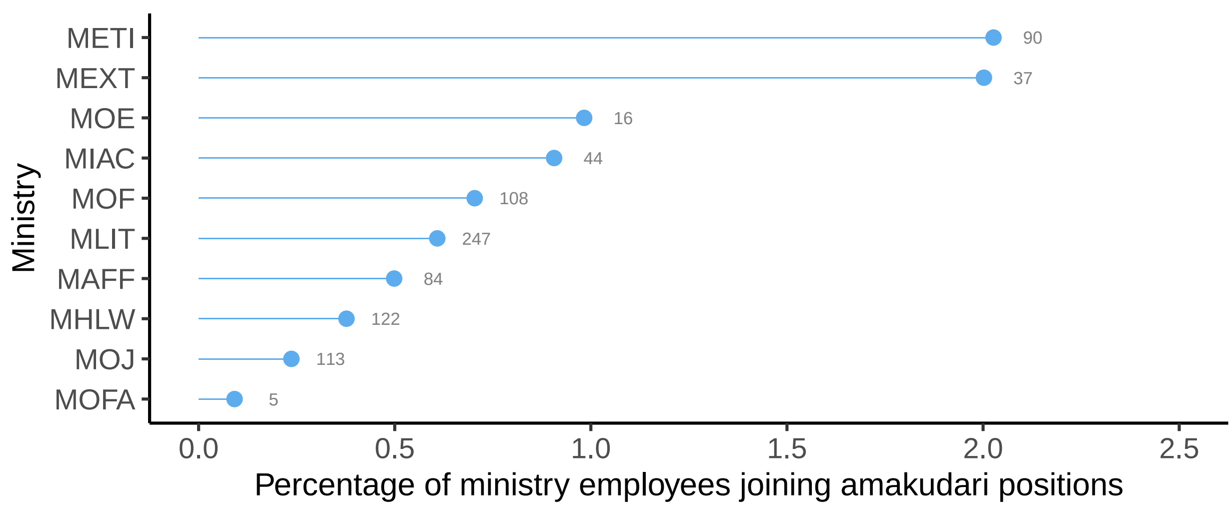


Figure 3.8: ***Amakudari*** **ministry of origin adjusted for ministry size (all years)**

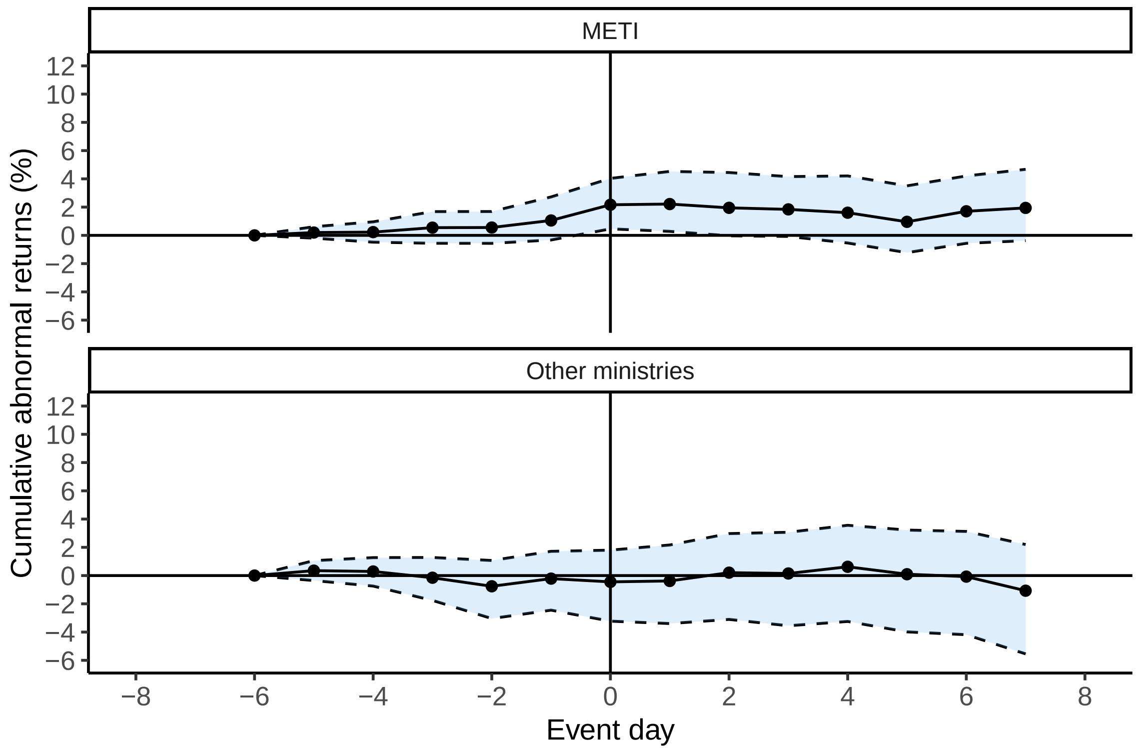


Figure 3.9: Cumulative abnormal returns from assistant vice-minister and vice-minister appointments

Note: Cumulative abnormal returns $R_{it} = \alpha_i + \beta_i R_{Mt} + \epsilon_{it}$ where R_{it} is the returns to firm i at time t, R_{Mt} is the return on the market portfolio at time t, and ϵ_{it} captures returns to firm i at time t that can be considered "abnormal" (above and beyond changes in the market porfolio R_{Mt}). 95% confidence intervals calculated using the bootstrap. The event day is the day hires are announced in Japan's largest business newspaper, the *Nikkei Shimbun*. N = 39 events.

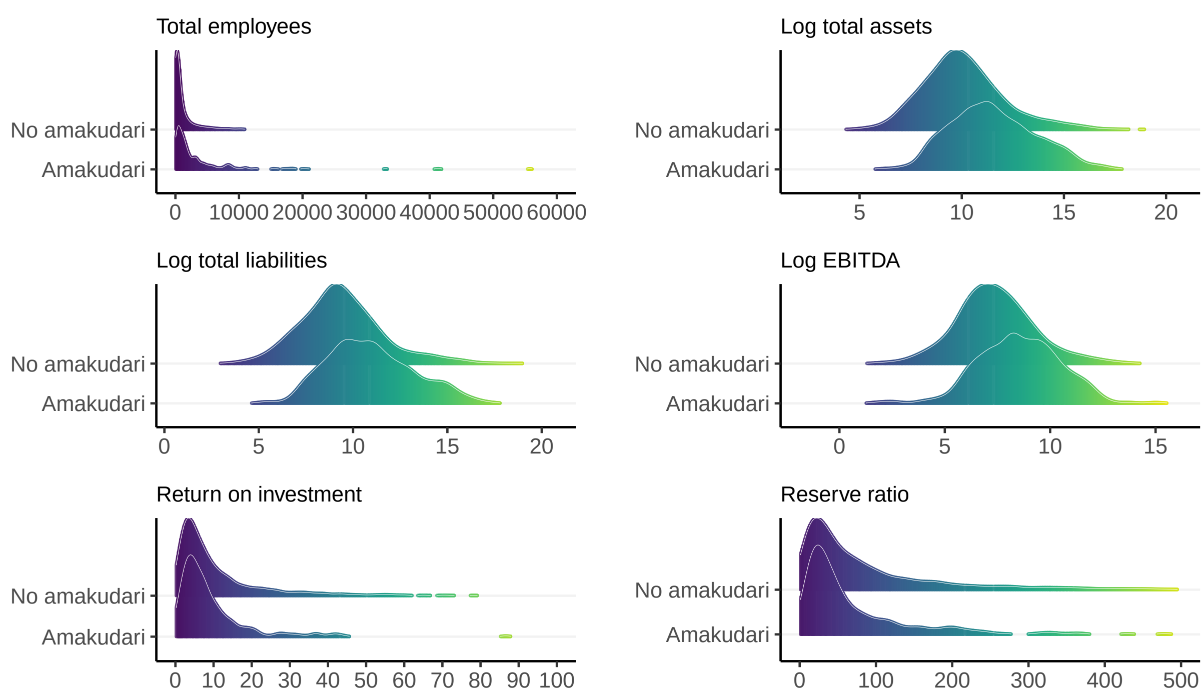


Figure 3.10: *Amakudari* vs. non-*amakudari* firm financials

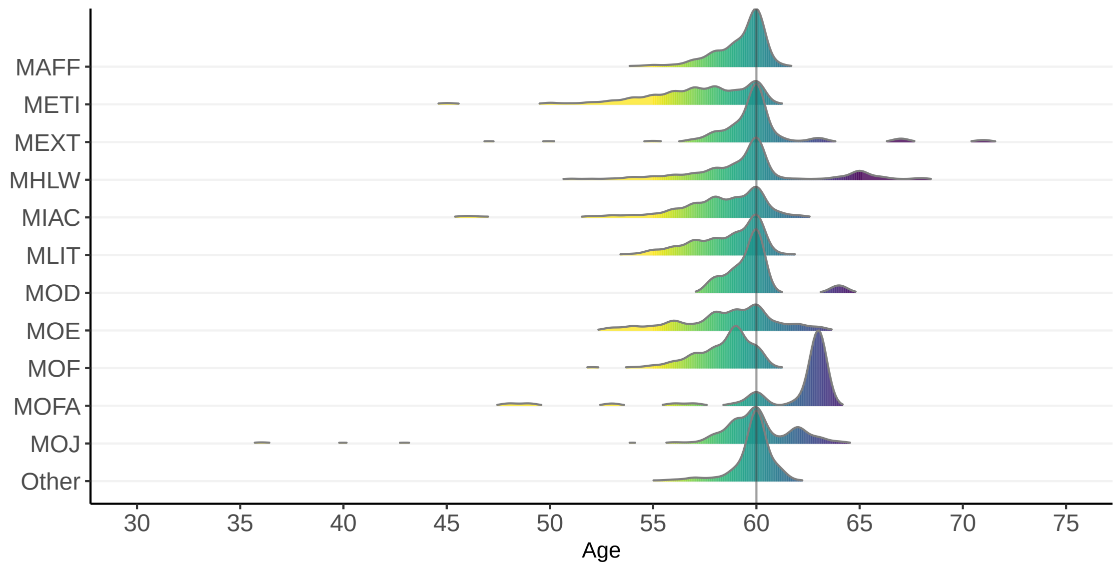


Figure 3.11: Age of exit from ministry, by ministry
Note: vertical line at "mandatory" retirement age of 60.

	Other official (N=2767)		Vice Minister or Assistant Vice Minister (N=246)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p
Total Assets	13.68	2.04	14.46	1.84	0.79	0.00
Total Liabilities	13.19	2.26	14.02	2.06	0.83	0.00
Revenue	13.08	1.82	13.50	1.56	0.42	0.01
Profit	11.47	1.89	12.24	1.56	0.77	0.00
ROI	1.61	0.60	1.87	0.64	0.25	0.00
EBITDA	10.81	2.08	11.46	1.64	0.65	0.00
Reserve Ratio	3.36	0.81	3.29	0.79	-0.07	0.39
Employees	9.06	1.81	9.58	1.65	0.53	0.00

Table 3.2: (Log of) financials of firms hiring top ranking officials vs. firms hiring lower ranked officials

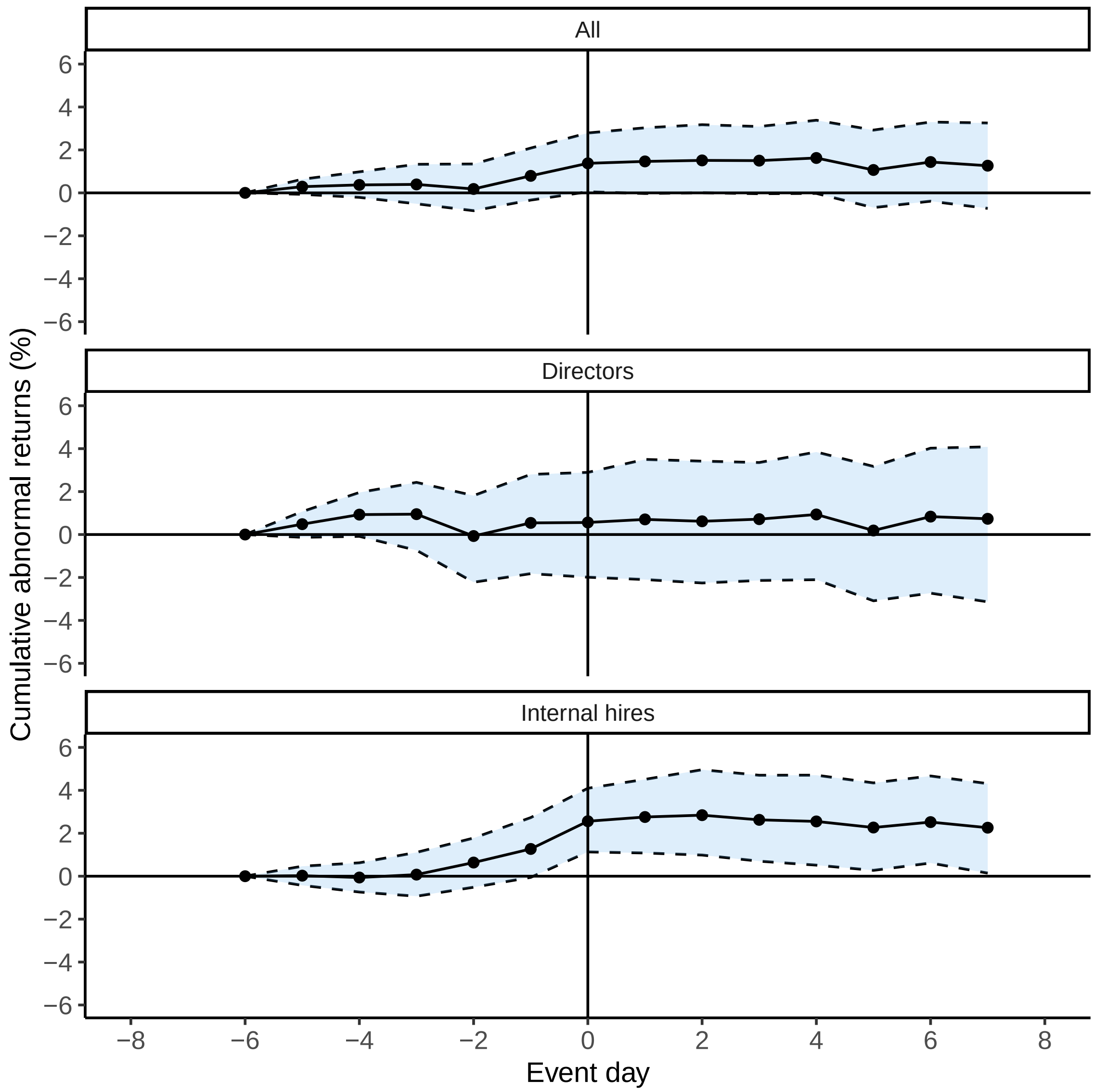


Figure 3.12: Cumulative abnormal returns from assistant vice-minister and vice-minister appointments

Note: Cumulative abnormal returns $R_{it} = \alpha_i + \beta_i R_{Mt} + \epsilon_{it}$ where R_{it} is the returns to firm i at time t, R_{Mt} is the return on the market portfolio at time t, and ϵ_{it} captures returns to firm i at time t that can be considered "abnormal" (above and beyond changes in the market porfolio R_{Mt}). 95% confidence intervals calculated using the bootstrap. The event day is the day hires are announced in Japan's largest business newspaper, the *Nikkei Shimbun*. N = 39 events.

Analyses conducted separately for internal (advisor, manager, and executive) and corporate governance (i.e. directors) related appointments as interviewees suggested that outside directors were "probably negatively correlated with the profitability of a company" and that "government outside directors have no meaning" as they lack business experience (Author Interview November 25, 2019; Author Interview December 2, 2019).

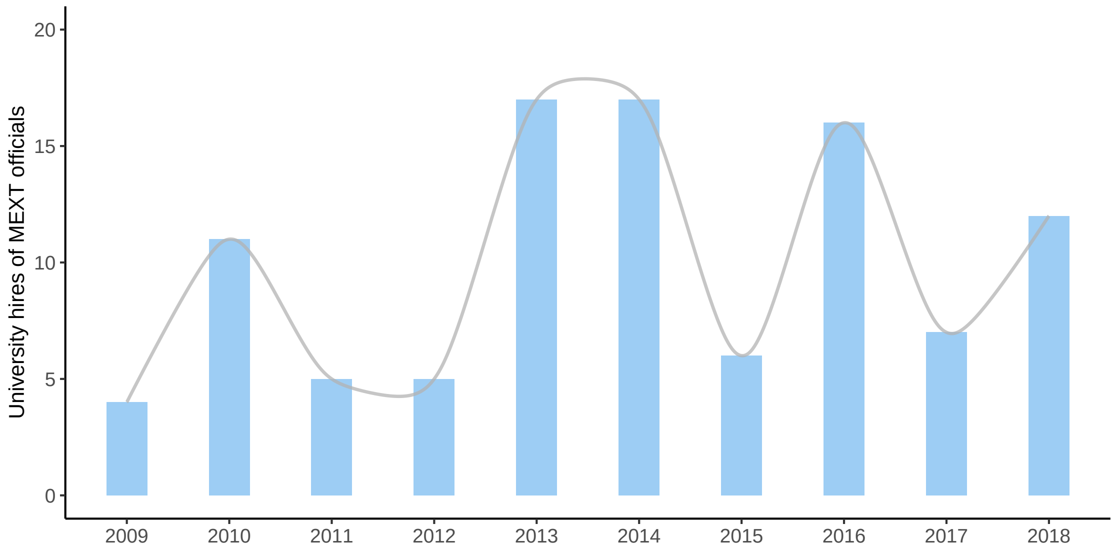


Figure 3.13: University hires of MEXT officials

chapter 4

Are regime changes always bad economics? Evidence from daily financial data

With Devin Incerti

Abstract: Political instability is commonly thought to discourage investment and reduce economic growth. We challenge this consensus by showing that instability does not systematically depress access to financial capital. Using an event study approach, we examine daily returns of national financial indices in every country that experienced an irregular regime change subject to data availability. Returns following resignations are large and positive (+4%), while those following assassinations are negative and smaller in magnitude (-2%). The impact of coups tends to be negative (-2%), but we show that a pro-capitalist coup results in large positive returns (+10%). We also find evidence that authoritarian or anti-capitalist regime changes are more likely to lead to capital flight than democratic or pro-business changes. The immediate impact of political instability on capital access is therefore dependent on the type of regime change and its expected impact on future growth.

Economies are now global, financialized, and integrated, with domestic economies intrinsically linked to international investment. However, political instability appears to jeopardize access to capital—instability is negatively correlated with investment, financial development, and GDP growth (Aisen and Veiga 2013; Alesina and Perotti 1996; Alesina, Özler, Roubini and Swagel 1996; Baker, Bloom and Davis 2016; Fosu 1992; Jong-A-Pin 2009; Roe and Siegel 2011), as well as associated with increases in stock market variance (Jensen and Schmith 2005; Leblang and Mukherjee 2005; Liu and Zhang 2015).

However, how different *types* of political instability affect domestic capital access remains unclear. In contrast to the aforementioned cross country studies, we therefore separately examine the effect of different types of instability on domestic firms' access to financial capital. Specifically, we examine whether changes in financial flows differ for coups, resignations, assassinations, and protests, as well as for authoritarian vs. democratic and pro vs. anti business shifts.

To test whether there are meaningful differences in capital flows following different types of political instability, we examine changes in stock market returns surrounding politically unstable domestic events. Specifically, we conduct event studies of daily financial data, which estimate a local average treatment effect of an unexpected event on stock prices *at the time of the event*. This interrupted time-series approach mitigates the endogeneity problems in previous cross country regressions—confounding events would need to occur on the same day as instability, and do so for a large portion of all of our independently tested events in order to influence our estimates.

We analyze the full sample of politically unstable events for which national-level daily financial data is available.[1] Like previous research, we find that all types of political instability in our sample cause large increases in financial volatility. However, we find that market returns are large and positive (+4%) following resignations and negative and smaller in magnitude following assassinations and coups (-2%). We also find that the failed pro-business

[1]13 coups, 8 assassinations, 15 forced resignations, and 11 public protests.

2002 Venezuelan coup caused large positive market returns (+10%), followed by immediate capital flight (-8%) after the reinstatement of a left-wing populist.

Our primary contributions are empirical and methodological. We provide the first estimates of the effects of different types of political instability on domestic capital access. We show that instability does not systematically depress investment, and find evidence that authoritarian regime changes are more likely to lead to negative returns, but that leaders who are clearly pro-business can be rewarded by financial markets even if they use extra-judicial methods to take power. The capital flows we document are not insubstantial—they are in some cases larger shocks than the 2008 stock market crash on their respective domestic economies. Methodologically, we (1) employ a method less susceptible to endogeneity concerns than previous studies, and (2) integrate synthetic control and event study methods to allow for control portfolios when a control candidate is not present.

Reexamining capital response to political instability

Stable capital flows are highly important to financial stability in emerging markets, which are particularly exposed to shifts in the availability of foreign capital (Cohen, Domanski, Fender and Shin 2017; Koepke 2019; Obstfeld 2012). High country risk reduces capital flows, which in turn has been shown to reduce domestic output growth (Koepke 2019).

Conventional theory and cross-country empirical evidence suggests that political instability increases country risk and therefore depresses financial returns and economic growth (Boutchkova, Doshi, Durnev and Molchanov 2012; Irshad 2017; Le and Zak 2006; Lehkonen and Heimonen 2015; Lensink, Hermes and Murinde 2000), and increases volatility (Bialkowski, Gottschalk and Wisniewski 2008; Irshad 2017; Jensen and Schmith 2005; Leblang and Mukherjee 2005; Liu and Zhang 2015). But if capital flows reflect expectations of future economic growth and returns, such hypotheses may be too simplistic. For example, capital access may increase in response to a coup if the current regime is anti-business or anti-global.

Not all irregular regime changes are equivalent. Some—e.g., the resignation of an ineffective leader—may foreshadow better policy and increased stability. Assassinations, by contrast, are not always related to the effectiveness of a leader, can occur seemingly at random, and the successor to the assassinated leader may be unclear. Coups can be democratic or autocratic. These differences are often overlooked in past studies, which have proxied for instability using the number of coups (Alesina et al. 1996; Londregan and Poole 1990), assassinations or revolutions (Barro 1991), or combined events into single indices (Alesina and Perotti 1996; Gupta 1990; Jong-A-Pin 2009; Venieris and Gupta 1986). By contrast—recognizing the different degrees of instability that these events may reflect—we estimate effects separately for coups, assassinations, resignations, and protests. Our event study approach affords us this freedom without encountering the statistical power issues that would arise in a regression setting.

A change in regime may be seen as a positive event from the perspective of foreign investors if the current regime is regarded as anti-business or anti-global. Regime changes that attempt to overthrow anti-business autocrats may be expected to result in positive abnormal returns as investors see little risk of a replacement government "worse" than the status quo. For example, an active debate exists on the subject of "good coups"—i.e. coups that lead to democratization or economic liberalization—and their effect on economic growth. Alesina et al. (1996) take the possibility of a new government adopting better economic policies seriously, but argue that the negative effects of uncertainty dominate the positive effects of coups staged by pro growth factions. Alesina and Perotti (1996) argue that "unconstitutional" regime changes such as coups are worse for economic growth than other types of political instability, while Londregan and Poole (1990) find that coups have no impact on economic growth and hypothesize that this may be because there is a bimodal distribution of coups, some of which enhance growth and others which restrict it. However, most research suggests that "good coups" are not the norm (Derpanopoulos, Frantz, Geddes and Wright 2016; Powell and Thyne 2011; Thyne and Powell 2016; Varol 2011). But while "good coups" may be the minority, some evidence suggests that they have become more frequent (Marinov and Goemans 2014) and could have positive economic effects. Meyersson (2016) finds that coup attempts that occur in democratic regimes are associated with negative GDP growth, while those that occur in autocratic regimes are associated with positive GDP growth. In another study examining the same outcome variable as our own—financial returns—Girardi and Bowles (2018) show that Salvador Allende's socialist government was associated with declines in stock prices while the coup that replaced him boosted them.

Previous research therefore suggests coups may on average cause capital flight, but lead to positive flows if the coup's instigators are more democratic or pro-market than the regime they replace. Assassinations should have a neutral or negative impact as they increase uncertainty, but institutional responses to assassinations vary by country. By contrast, resignations may be viewed positively on average, as they often signal the departure of an ineffective leader.

Our empirical findings validate these predictions. We find that the type of political event and its expected impact on economic policy determines the direction of abnormal returns. Events expected to lead to more stable governance, economic liberalization, or democratization (such as willful resignations and coups that overthrow protectionist or leftist autocrats) are associated with positive returns, while those that consolidate authoritarian rule (e.g. military coups), exacerbate poor economic policies, or merely increase policy uncertainty (such as assassinations) have the opposite effect.

Data

Financial data are from the Global Financial Data database, which includes the longest available daily time series of stock prices. We collect data on national equity indices and two global equity indices, the S&P/IFC Emerging Market Investable Composite and the Morgan Stanley Capital International (MSCI) World Index. The S&P/IFC index includes securities from emerging markets while the MSCI index includes securities from developed markets only. We collect national stock index data on every country in which there was a coup or coup attempt, an assassination or failed assassination, or a forced resignation and for which daily financial data is available.[2] The longest available daily time series for these stock indices are listed in Table 4.1.

Political data are primarily drawn from the Center for Systemic Peace's (CSP) Polity IV Coup d'etat dataset and Coup d'etat Events handbook. The Coup d'etat dataset includes the date of 1) successful coups, 2) attempted coups, 3) plotted coups and 4) alleged coup plots. We focus on successful coups because it is difficult to classify failed coups, and the choice of dataset and classification methodology would directly alter our findings. Needler (1966, p. 617) has gone so far as to say that "the categories of coups that were aborted, suppressed, or abandoned melt into each other and into a host of other non-coup phenomena so as to defy accounting," the Center for Systemic Peace states that it is "confident that [its] list of successful coups is comprehensive" but does not extend this confidence to attempted or failed coups, and Powell and Thyne (2011) state that it is "difficult to identify more ambiguous forms of coup activity, such as coup failures, plots, and rumors."

The Coup d'etat Events handbook also provides a list of 1) auto-coups[3], 2) the ouster of leadership by foreign forces, 3) the ouster of leadership by rebel forces, 4) assassinations of the executive and 5) resignations of the executive due to poor performance and/or loss of authority. Daily financial data is available for countries in categories 4 and 5, so we supplement the coups with assassinations and resignations to form a list of "irregular" regime changes. The resignations are those in which the ruling executive was coerced to resign due to poor performance, public discontent and popular demonstrations. Note that the Polity IV definition of "poor performance" is not synonymous with poor *economic* performance, and in practice the reasons cited for resignation across events are: loss in conflict/war, anti-authoritarian protest, corruption scandals, Supreme Court ruling against unconstitutional

[2]The list of failed assassinations are from Jones and Olken (2009). Coup attempts are those in category 2 in the CSP Coup d'etat dataset.

[3]Defined by Center for Systemic Peace as the "occurrence of subversion of the constitutional order by a ruling (usually elected) executive and the imposition of an autocratic regime."

Table 4.1: List of stock indices

Date	Country	Begin Date	Start Date
Argentina	Beunos Aires SE General Index	Dec-66	Jan-17
Australia	Australia ASX All Ordinaries	Jan-58	Jan-17
Bangladesh	Dhaka SE Index	Jan-90	Jan-17
Canada	Canada S&P/TSX 300 Composite	Jan-76	Jan-17
Chile	Santiago IGBC General Index	Jan-75	Jan-17
Colombia	Colombia IGBC General Index	Jan-92	Jan-17
Ecuador	Ecuador Bolsa de Valores de Guayaquil (BVG)	Jan-94	Jan-17
Egypt	Cairo SE Index	Dec-92	Jan-17
Emerging Market	S&P/IFC Emerging Markets Investable Composite	Jul-95	Jan-17
Greece	Athens SE General Index	Oct-88	Jan-17
India	Bombday SE SENSEX	Apr-79	Jan-17
Indonesia	Jakarta SE Composite Index	Apr-83	Jan-17
Iran	Tehran SE Price Index	Jan-95	Jan-17
Israel	Tel Aviv 100 Index	May-87	Jan-17
Japan	Tokyo SE Price Index (TOPIX)	Jan-53	Jan-17
Latin America	Dow Jones Latin America Index	Jan-92	Jan-17
Lithuania	Lithuania Lit-10 Index	Jan-99	May-05
Malaysia	Malaysia KLSE Composite	Jan-80	Jan-17
Nepal	Nepal NEPSE Stock Index	Jan-01	Jan-17
Netherlands	Netherlands All-Share Price Index	Jan-80	Jan-17
Pakistan	Karachi SE 100 Index	Jan-89	Jan-17
Paraguay	Asuncion SE PDV General Index	Oct-93	Sep-08
Peru	Lima SE General Index	Jan-82	Jan-17
Philippines	Manila SE Composite Index	Jan-86	Jan-17
Portugal	Oporto PSI-20 Index	Jan-86	Jan-17
Singapore	Singapore FTSE ST Index	Jul-65	Jan-17
South Korea	Korea SE Stock Price Index	Jan-62	Jan-17
Southeast Asia	Dow Jones Southeast Asia Index	Jan-92	Jan-17
Spain	Madrid SE General Index	Aug-71	Jan-17
Sri Lanka	Colombo SE All-Share Index	Dec-84	Jan-17
Sweden	Sweden OMX Affarsvarlden General Index	Jan-80	Jan-17
Taiwan	Taiwan SE Capitalization Weighted Index	Jan-67	Jan-17
Thailand	Thailand SET General Index	Apr-75	Jan-17
Tunisia	Tunisia SE Index	Dec-97	Jan-17
Turkey	Instanbul IMKB 100 Price Index	Oct-87	Jan-17
Ukraine	Ukraine PFTS OTC Index	Jan-98	Jan-17
United Kingdom	UK FTSE All-Share Index	Nov-62	Jan-17
United States	Dow Jones Industrial Average	Feb-1885	Jan-17
Uruguay	Bolsa de Valores de Montevideo Index	Jan-08	Jul-16
Venezuela	Dow Jones Venezuela Stock Index	Jan-92	Jul-07
Venezuela	Caracas SE General Index	Jan-94	Jan-17
World	MSCI World Price Index	Jan-76	Jan-17
Zambia	Lusaka SE Index	Jan-02	Apr-06
Zambia	Lusaka SE Index	Jul-11	Jan-17

Table 4.2: Regime changes

Date	Country	Political Outcome
	Coups	
06/30/1970	Argentina	Autocracy to autocracy
03/22/1971	Argentina	Autocracy to autocracy
03/24/1976	Argentina	Democracy to autocracy
10/06/1976	Thailand	Anocracy to autocracy
10/20/1977	Thailand	Autocracy to anocracy
12/12/1979	South Korea	Autocracy to autocracy
02/23/1991	Thailand	Anocracy to anocracy
04/05/1992	Peru	Democracy to anocracy
10/12/1999	Pakistan	Democracy to autocracy
10/04/2002	Nepal	Democracy to autocracy
09/19/2006	Thailand	Democracy to anocracy
01/11/2007	Bangladesh	Democracy to autocracy
07/03/2013	Egypt	Anocracy to anocracy
05/22/2014	Thailand	Democracy to anocracy
	Failed coup	
04/11/2002	Venezuela	Democracy to democracy
	Assassinations	
09/06/1901	United States	Democracy to democracy
11/22/1963	United States	Democracy to democracy
10/26/1979	South Korea	Autocracy to autocracy
10/31/1984	India	Democracy to democracy
02/28/1986	Sweden	Democracy to democracy
05/01/1993	Sri Lanka	Anocracy to anocracy
11/04/1995	Israel	Democracy to democracy
06/01/2001	Nepal	Democracy to democracy
	Resignations	
06/17/1982	Argentina	Autocracy to autocracy
02/25/1986	Philippines	Autocracy to democracy
12/06/1990	Bangladesh	Anocracy to anocracy
05/24/1992	Thailand	Anocracy to anocracy
04/18/1993	Pakistan	Democracy to democracy
11/05/1996	Pakistan	Democracy to democracy
06/30/1997	Turkey	Democracy to democracy
05/21/1998	Indonesia	Autocracy to anocracy
01/20/2001	Philippines	Democracy to democracy
12/20/2001	Argentina	Democracy to democracy
04/06/2004	Lithuania	Democracy to democracy
12/26/2004	Ukraine	Democracy to democracy
04/20/2005	Ecuador	Democracy to democracy
04/24/2006	Nepal	Autocracy to democracy
01/14/2011	Tunisia	Anocracy to democracy

Notes: The Polity score is used to classify political outcomes as follows: autocracy = $-10 \leq$ score ≤ -6, anocracy = $-5 \leq$ score ≤ 5, and democracy = $6 \leq$ score ≤ 10.

actions, contested elections, and abuse of power.

We further supplement the CSP data with leadership data from Archigos Version 4.1, which allows us to identify additional cases in which a "leader lost power through irregular means." Irregular transfers of power are those in which leaders do not leave office "in a manner prescribed by either explicit rules or established conventions." Nearly all removals by irregular means result from the threat or use of force (e.g. coups, revolts and assassinations).

A list of the political events in our dataset is shown in Table 4.2. Coups tend to have the largest impact on the level of democratization as a number of countries have subsequently transitioned from democracies to anocracies or autocracies. On the other hand, neither assassinations or resignations have much of an impact on the level of democratization.

The discussion above shows that there is considerable debate about classification of regime changes. We recognize that some readers may feel that certain events are missing based on their own substantive knowledge. However, we choose to rely on commonly used third-party classifications in order to minimize the possibility that our results are driven by our own subjective classifications. The one exception is that we separately study the failed coup attempt in Venezuela in April 2002 in which the President of Venezuela, Hugo Chavez, was removed from office for two days because it provides a natural test of the impact of the seemingly successful removal of a left-wing populist with a pro-business regime, and the ensuing reinstatement of a left-wing populist.

To gain perspective on the relationship between irregular regime changes and the stock market, Figure 4.1 plots the absolute value of daily stock returns averaged across all events. Returns are for 200 trading days before and after regime changes. The absolute value of returns on the event day (trading day 0) are significantly larger than on any other day. In addition, the magnitude of returns begins increasing just before the event day and remains high for a short period after. This suggests that financial volatility increases during the days surrounding regime changes, although it does not provide any evidence on mean returns.

Impact of Political Instability on Stock Returns tests these results more formally. It will first formally estimate the amount of volatility surrounding all irregular regime changes, then analyze coups, assassinations and resignations separately and determine what impact they have on mean returns.

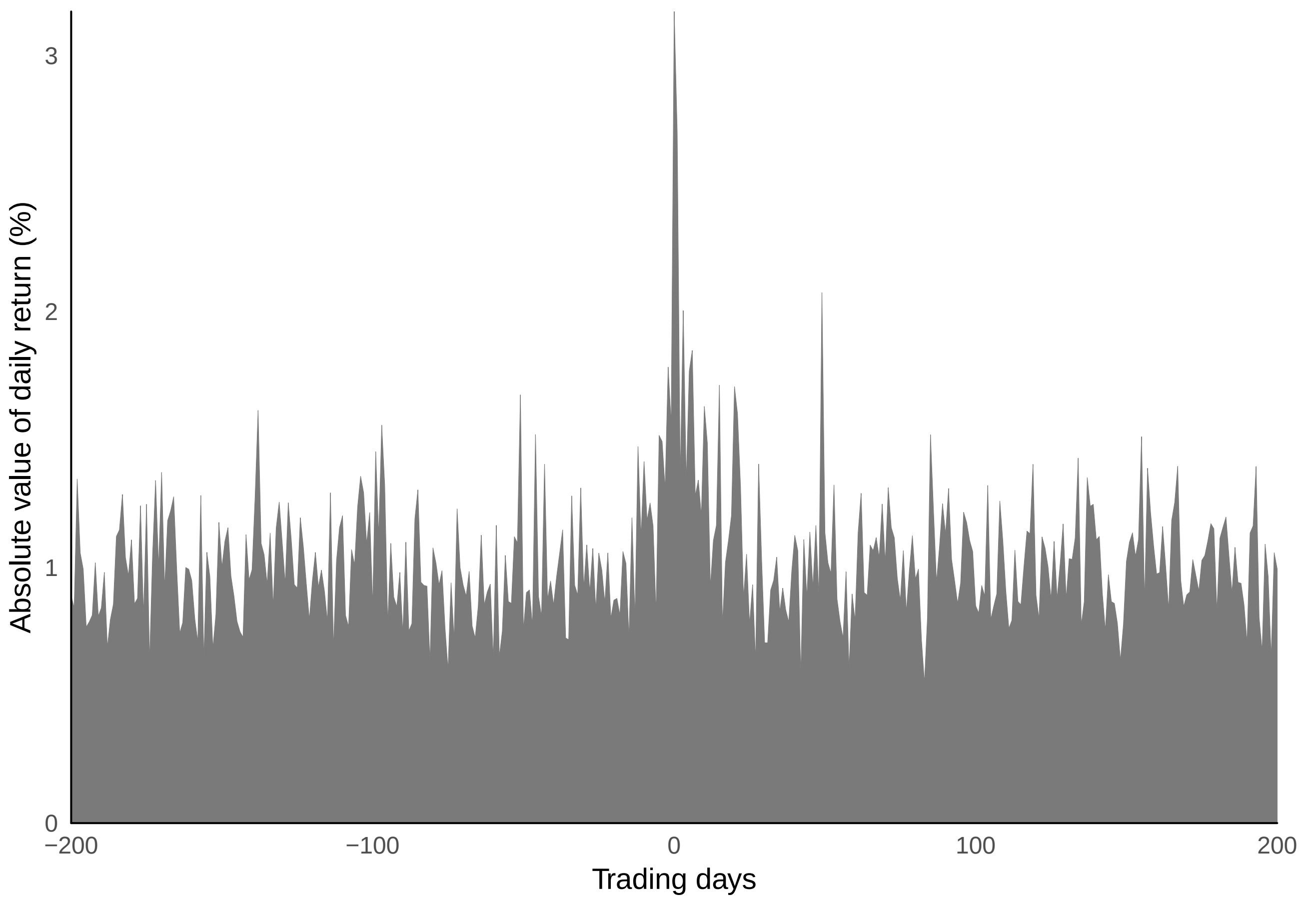


Figure 4.1: Absolute value of daily returns

Estimation

Volatility

To estimate the effect of irregular regime changes on financial volatility, we use a generalized autoregressive conditional heteroskedasticity (GARCH) model estimated using 1000 pre-event days, the event day and 1000 post-event days. As in Jensen and Schmith (2005) and

Leblang and Mukherjee (2005), we use the GARCH (1,1) specification. In particular, for national stock index i,

$$R_{it} = \mu_i + \epsilon_{it}, \qquad \epsilon_{it} \sim \mathcal{N}\left(0, \sigma_{it}^2\right),$$

where μ_i is a constant and,

$$\sigma_{it}^2 = \gamma_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i \sigma_{i,t-1}^2.$$

The key parameter of interest is the conditional variance, σ_{it}^2. The one-period-ahead volatility forecasts, σ_{it}, are larger when $\epsilon_{i,t-1}^2$ and $\sigma_{i,t-1}^2$ are larger. In other words, the model predicts that large shocks will be followed by other large shocks.

Abnormal returns

To estimate the effect of irregular regime changes on financial flows, we follow standard event study methodology (Campbell, Lo and MacKinlay 1997; MacKinlay 1997). Normal performance is measured with a constant mean return model,

$$R_{it} = \mu_i + \epsilon_{it}, \tag{4.1}$$

where R_{it} is the logged return of national stock index i on trading day t and ϵ_{it} is the error term. We calculate abnormal returns (ARs) in an "event window" around the date of each event, $AR_{i\tau} = R_{i\tau} - \widehat{\mu}_i$, where τ is a date in the event window, and $\widehat{\mu}_i$ is estimated in an "estimation window" preceding the event window with Equation 4.1. We use a 41 trading day event window (20 pre-event days, event day, and 20 post-event days) and 250 trading day estimation window. The ARs are then used to calculate cumulative abnormal returns (CARs) between event day τ_1 and event day τ_2: $CAR(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} AR_{i\tau}$. Standard errors and p-values are calculated using asymptotic t-statistics as in MacKinlay (1997).[4] The event

[4]This is appropriate because the length of the estimation window is sufficiently long (250 trading days).

date is the first trading day a market could react to news of the event.

We do not use a market model (where a market index is included as a control) as our unit of analysis *is* the country-wide market index (i.e., not a firm). To address concerns regarding use of a constant mean return model, we combine synthetic control methods (Abadie, Diamond and Hainmueller 2010) with event study estimation. We create a "synthetic" control portfolio for each event, where each country is given a weight representing its influence in the portfolio. The weight is chosen so that the daily returns and the variance of the daily returns of the control portfolio and the event country are most similar in the estimation window. The possible countries in the control portfolio are all countries listed in Table 4.1. A more formal exposition of the methodology can be found in Synthetic Control Portfolio.

We define the event date as the first trading day in which the market could have reacted to news of the event. For example, during the October 12, 1999 coup d'etat in Pakistan led by General Pervez Musharraf, the army announced that Prime Minister Nawaz Sharif had been dismissed after market hours at 10:15 pm. We code October 14th, the day in which the market re-opened, as the event day. When events occurred on weekends, we change the event date to the following Monday.

$(0, \tau - 1)$ is used to denote the τ-day period beginning with the event day and $(-1, \tau)$ to denote the negative τ-day period beginning with the day prior to the event day. In other words, for cumulative abnormal returns prior to the event date, we aggregate backwards starting at the day of the event. For example, $CAR(-1, -2)$ is the sum of the abnormal returns on event date -1 and event day -2. We present results for the sum of abnormal returns over the post-event windows of the event date only $(0, 0)$, the event date plus 6 days $(0, 6)$, and the event date plus 19 days $(0, 19)$. In addition, we present results for the pre-event windows $(-1, -7)$ and $(-1, -20)$.

As we hypothesize that different types of regime changes will have disparate effects on markets, we report abnormal returns separately for coups, assassinations and resignations. Standard errors and p-values are calculated using asymptotic t-statistics as in MacKinlay

(1997).[5]

This interrupted time series approach implies that the "control group" being compared against is not, for example, regular regime[6] changes or failed regime changes, but the expected returns in the same country on the same day in a but-for world in which no regime change occurred. We report abnormal returns separately for coups, assassinations and resignations. However, each event CAR is by itself also a valid causal estimate of the effect of a specific regime change.

Impact of Political Instability on Stock Returns

Volatility

We first confirm that our sample of events exhibits the increases in volatility suggested by previous literature. Figure 4.2 shows the mean volatility ($\overline{\sigma_t}$) estimates across all irregular regime changes for 250 trading days prior to and 250 days after each event. Volatility stays between a narrow range at nearly all dates except those surrounding the regime change. There is an enormous volatility jump on the day of the regime change, with levels then decreasing to normal within a month of the event.

Abnormal Returns

Coups

Table 4.3 shows abnormal returns for national stock indices both preceding and following coup d'etat. Table 4.3 contains all coups presented in Table 4.2 with the exception of the Argentinian coup of March 24, 1976. The March 24, 1976 Argentinian coup is excluded from our analysis because the stock market remained closed from March 24 to April 5, 1976, or a

[5]It is appropriate to use the standard normal distribution to calculate test statistics because the length of the estimation window is sufficiently long (250 trading days).

[6]Past studies have examined market reactions to regular regime changes. This work finds that volatility increases around elections—particularly close elections and those that lead to a change in government political orientation (Bialkowski, Gottschalk and Wisniewski 2008) —and that positive returns tend to precede elections (Pantzalis, Stangeland and Turtle 2000).

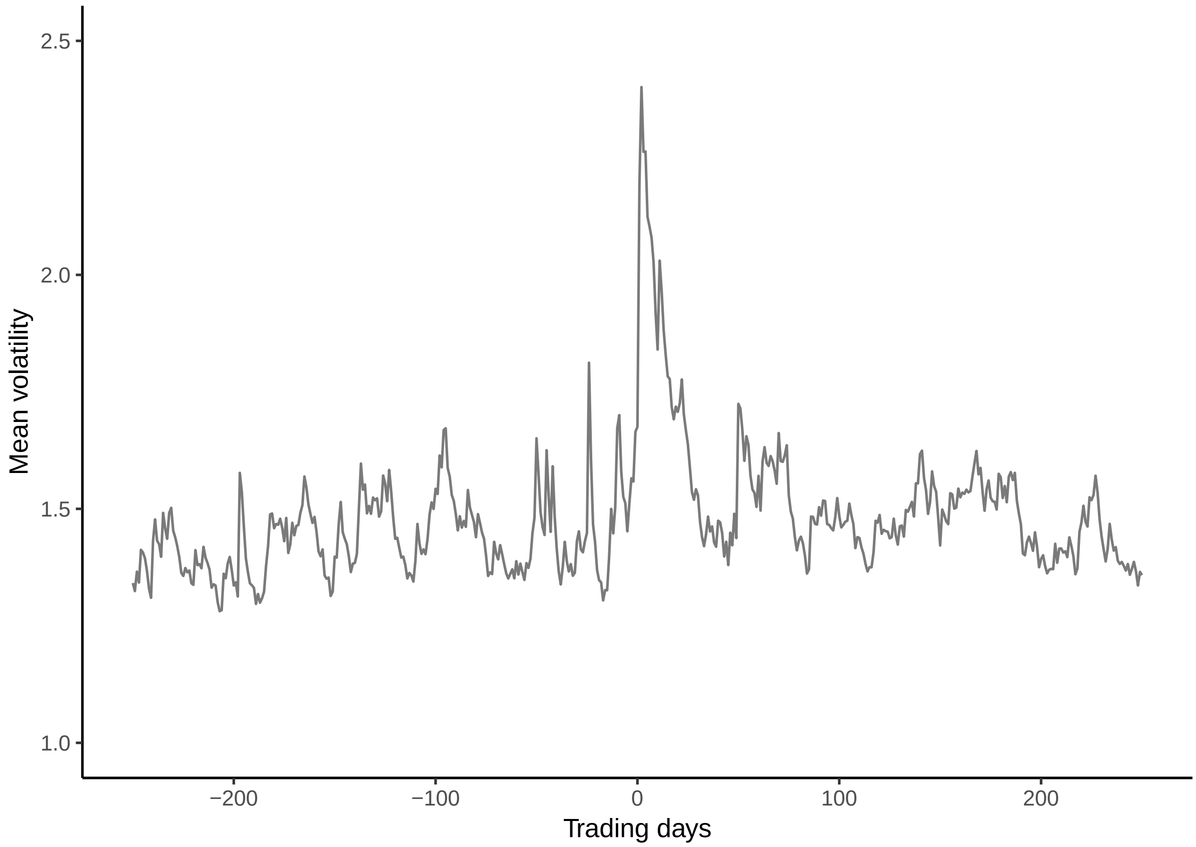


Figure 4.2: Mean of volatility estimates from GARCH(1,1) models

period of twelve days.[7]

The average coup has a −2.1% event day AR. Event day ARs for the 1970 coup in Argentina, the 1991 coup in Thailand, the 1992 coup in Peru, and the 1999 coup in Pakistan are all negative and statistically different than zero. Moreover, all of these cases except Thailand have negative post-event CARs and pre-event CARs that are statistically indistinguishable from zero. In all of these cases, the coup either overthrew a democratically elected government or changed governance from one military ruler to another. The initial negative reaction followed by additional post-event negativity is consistent with the expected market reaction from a successful authoritarian coup followed by post-event consolidation of power.

The only events with positive ARs are the 1971 coup in Argentina and the 2002 coup in

[7]Treating this twelve day period as a single day CARs results in a positive abnormal return of 58%, a fluctuation that seems qualitatively unreasonable.

Table 4.3: Abnormal returns following coups

Country	Event Date	Post-Event CAR (0,0)	(0,6)	(0,19)	Pre-Event CAR (-1,-7)	(-1,-20)	Days to rebound
Argentina	06/08/1970	−1.919	−0.530	−2.011	0.247	4.728	204
		(0.949)	(2.510)	(4.243)	(2.510)	(4.243)	
Argentina	03/22/1971	0.925	14.294	24.218	0.131	0.274	
		(1.216)	(3.218)	(5.439)	(3.218)	(5.439)	
Bangladesh	01/11/2007	−0.320	10.351	14.883	−0.896	2.250	2
		(1.166)	(3.086)	(5.217)	(3.086)	(5.217)	
Egypt	07/03/2013	−0.346	5.169	7.144	6.776	−4.869	2
		(1.515)	(4.009)	(6.776)	(4.009)	(6.776)	
Nepal	10/04/2002	0.090	1.563	5.567	−1.014	−0.493	2
		(1.206)	(3.190)	(5.392)	(3.190)	(5.392)	
Pakistan	10/14/1999	−7.737	−9.431	−7.130	4.151	4.900	36
		(1.943)	(5.141)	(8.690)	(5.141)	(8.690)	
Peru	04/06/1992	−6.819	−5.814	−25.027	−2.075	−10.519	5
		(2.210)	(5.848)	(9.885)	(5.848)	(9.885)	
South Korea	12/12/1979	−1.784	−3.474	−24.465	−1.678	−6.187	418
		(1.152)	(3.047)	(5.150)	(3.047)	(5.150)	
Thailand	10/06/1976	−0.541	0.837	0.731	0.001	0.713	3
		(0.639)	(1.691)	(2.859)	(1.691)	(2.859)	
Thailand	10/20/1977	−0.951	4.096	7.290	9.961	10.198	2
		(1.232)	(3.260)	(5.510)	(3.260)	(5.510)	
Thailand	02/25/1991	−7.326	2.860	14.162	6.326	26.262	7
		(2.884)	(7.631)	(12.899)	(7.631)	(12.899)	
Thailand	09/19/2006	−0.481	−2.640	0.111	1.848	0.131	17
		(1.094)	(2.894)	(4.892)	(2.894)	(4.892)	
Thailand	05/23/2014	−0.571	2.800	4.591	2.350	−0.424	5
		(1.201)	(3.177)	(5.370)	(3.177)	(5.370)	
Mean		−2.137	1.545	1.543	2.010	2.074	58
		(0.424)	(1.121)	(1.896)	(1.121)	(1.896)	

Notes: Standard errors are in parentheses. "Days to rebound" is the number of trading days following a negative stock return for the national stock index to return to pre-event level (it is calculated if the price decreases on the event day, not if the event day abnormal return is negative). Returns are inflation adjusted.

Nepal. These results provide evidence that coups do not necessarily lead to negative abnormal returns. While the 1971 Argentinian coup did result in another military leader, it did so while calling for free and democratic elections and replaced a government that had adopted extreme protectionist economic policies. In fact, by 1973 Argentina had transitioned to a democracy.[8] The 2002 coup in Nepal resulted in a monarchical restoration, but occurred after the country's prime minister postponed general elections, itself a democratically subversive action.

[8]Based on Center for System Peace Polity IV polity score of 6. Values of 6-10 are defined as democracies.

Assassinations

The results in Table 4.4 are produced from analyses identical to those in Table 4.3 but for successful assassinations rather than coups. Like the majority of coups, there is evidence that assassinations decrease stock prices. The mean event day abnormal return is negative and statistically different than zero. However, the result is driven by five events: the shooting of U.S. President William McKinley on September 6, 1901; the assassination of U.S. President John F. Kennedy on November 22, 1963; the assassination of Indian Prime Minister Indira Gandhi on October 31, 1984; the suicide bombing that killed Sri Lankan president Ranasinghe Premadasa on May 1, 1993; and the assassination of Israeli Prime Minister Yitzhak Rabin on the evening of November 4, 1995.

Table 4.4: Abnormal returns following assassinations

		Post-Event CAR			Pre-Event CAR		Days to
Country	Event Date	(0,0)	(0,6)	(0,19)	(-1,-7)	(-1,-20)	rebound
India	11/05/1984	−2.416	−1.259	−2.416	−3.916	1.344	5
		(0.668)	(1.767)	(2.987)	(1.767)	(2.987)	
Israel	11/05/1995	−3.460	−3.177	0.743	−0.857	−10.316	12
		(1.473)	(3.897)	(6.587)	(3.897)	(6.587)	
Nepal	06/12/2001	−0.513	2.965	15.516	5.956	1.791	20
		(3.513)	(9.295)	(15.711)	(9.295)	(15.711)	
South Korea	10/26/1979	−0.364	−9.376	1.186	0.690	−0.368	14
		(1.058)	(2.800)	(4.734)	(2.800)	(4.734)	
Sri Lanka	05/03/1993	−3.231	−0.983	3.515	−0.541	−1.360	7
		(0.767)	(2.030)	(3.432)	(2.030)	(3.432)	
Sweden	03/03/1986	0.698	5.038	10.908	−3.754	0.955	
		(0.927)	(2.452)	(4.145)	(2.452)	(4.145)	
United States	09/07/1901	−4.522	−3.055	−8.920	−0.733	3.456	963
		(1.283)	(3.394)	(5.738)	(3.394)	(5.738)	
United States	11/22/1963	−2.973	2.451	2.267	−2.666	−2.720	2
		(0.470)	(1.242)	(2.100)	(1.242)	(2.100)	
Mean		−2.098	−0.924	2.850	−0.728	−0.902	146
		(0.550)	(1.456)	(2.462)	(1.456)	(2.462)	

Notes: Standard errors are in parentheses. “Days to rebound” is the number of trading days following a negative stock return for the national stock index to return to pre-event level (it is calculated if the price decreases on the event day, not if the event day abnormal return is negative). Returns are inflation adjusted.

These results are consistent with our hypothesis that the nature of the political event and its expected impact on policy matters, and that assassinations should have a negative effect as they occur seemingly at random and increase uncertainty. While the mean effect of assassinations is negative, it is smaller in magnitude than for coups. Unlike a coup, an assassination may not necessarily be expected to cause immediate change in economic policy, particularly in the presence of an institutionalized line of succession. As such, we would expect CARs to be negative due to increased instability and uncertainty, but smaller in magnitude to a coup or resignation due to greater expectations of policy inertia.

There is no evidence of post or pre-event CARs in almost any of the assassinations. This is consistent with expectations as assassinations are typically not predictable. As with coups, the number of days that it took the stock market to rebound to pre-event levels is fairly low.[9]

Resignations

In contrast to coups and assassinations, abnormal returns following resignations are large and positive (see Table 4.5). The mean event day abnormal return is over 4% and the positive returns are persistent and grow larger over time (mean 20-day CAR $\approx$ 12%). Furthermore, event day ARs are only negative and statistically significant at even the ten percent level in two out of the fifteen resignations (Pakistan on April 19, 1993 and Tunisia on June 31, 2011).

These results are again consistent with our hypothesis that different events will have disparate effects, and that resignations may lead to capital inflows. The positive event day abnormal return following resignations is not surprising as resignations typically occur because of poor performance and/or loss of authority. Among our sample of events, leaders were ousted following loss in conflict/war, anti-authoritarian protest, corruption scandals, Supreme Court ruling against unconstitutional actions, contested elections, and abuse of power.

For example, consider Ferdinand Marcos' resignation from office as President of the Philip-

[9] One exception is the assassination of William Mckinley in which the stock market didn't fully recover for 963 days, or almost 4 calendar years. However, this was likely caused by the Panic of 1901, which began when the stock market crashed on May 17th, 1901, and not by McKinley's death (although the assassination may have exacerbated the panic).

Table 4.5: Abnormal returns following resignations

Country	Event Date	Post-Event CAR (0,0)	Post-Event CAR (0,6)	Post-Event CAR (0,19)	Pre-Event CAR (-1,-7)	Pre-Event CAR (-1,-20)	Days to rebound
Argentina	06/18/1982	18.892	24.904	65.863	−2.819	28.234	
		(3.334)	(8.822)	(14.912)	(8.822)	(14.912)	
Argentina	12/20/2001	14.015	48.103	62.191	14.656	36.165	
		(1.976)	(5.227)	(8.836)	(5.227)	(8.836)	
Bangladesh	12/07/1990	0.323	1.002	2.171	1.880	3.654	
		(0.871)	(2.305)	(3.896)	(2.305)	(3.896)	
Ecuador	04/20/2005	−0.084	−0.249	−0.595	−1.305	0.710	
		(0.945)	(2.499)	(4.225)	(2.499)	(4.225)	
Indonesia	05/20/1998	2.817	4.296	4.543	−2.695	−17.868	
		(3.392)	(8.974)	(15.168)	(8.974)	(15.168)	
Lithuania	04/06/2004	−0.575	−3.319	−11.704	2.182	5.426	159
		(1.137)	(3.007)	(5.083)	(3.007)	(5.083)	
Nepal	04/25/2006	1.915	8.132	9.937	−1.951	−4.205	
		(0.665)	(1.760)	(2.975)	(1.760)	(2.975)	
Pakistan	04/19/1993	−3.265	−0.432	2.771	−0.312	−0.485	15
		(1.108)	(2.930)	(4.953)	(2.930)	(4.953)	
Pakistan	11/06/1996	5.084	1.229	−0.441	4.182	7.597	
		(1.416)	(3.746)	(6.331)	(3.746)	(6.331)	
Philippines	02/26/1986	12.938	21.473	23.086	−1.847	−6.884	
		(0.477)	(1.263)	(2.134)	(1.263)	(2.134)	
Philippines	01/19/2001	1.150	16.837	18.469	−5.382	3.581	
		(1.591)	(4.209)	(7.115)	(4.209)	(7.115)	
Thailand	05/25/1992	3.248	−6.574	3.789	−5.085	−10.841	
		(1.433)	(3.793)	(6.411)	(3.793)	(6.411)	
Tunisia	01/31/2011	−2.705	2.982	−11.787	−13.610	−13.445	5
		(0.671)	(1.776)	(3.002)	(1.776)	(3.002)	
Turkey	06/30/1997	2.010	−2.861	−7.629	12.876	4.532	
		(3.015)	(7.976)	(13.481)	(7.976)	(13.481)	
Ukraine	12/28/2004	5.118	12.837	18.445	4.170	32.085	
		(2.797)	(7.401)	(12.511)	(7.401)	(12.511)	
Mean		4.059	8.557	11.941	0.329	4.550	59
		(0.496)	(1.312)	(2.217)	(1.312)	(2.217)	

Notes: Standard errors are in parentheses. "Days to rebound" is the number of trading days following a negative stock return for the national stock index to return to pre-event level (it is calculated if the price decreases on the event day, not if the event day abnormal return is negative). Returns are inflation adjusted.

pines in February 1986. Prior to his resignation, the Philippine regime was known for rampant corruption, crony capitalism, extreme inequality, high unemployment, failed import substitution industrialization policy, and oligarchic control of the economy (Overholt 1986; Traywick 2014). In fact, the Philippines was the least preferred site for foreign investment amongst East Asian capitalist economies and possessed one of the worst capital investment to economic output ratios in Asia (Overholt 1986). Marcos held a snap presidential election on February 7, 1986, in which he declared victory despite overwhelming evidence of electoral fraud. Public protests ensued, and two weeks later the military withdrew its support of the Marcos regime (Lee 2009). Marcos was replaced by his electoral opponent, Corazon Aquino, who had run on a platform of economic liberalization and elimination of crony capitalism (Villegas 1987). This event was associated with an approximately 13% positive event day AR.

By contrast, the largest negative event in our sample (-3%) is the 1993 resignation of President Ghulam Ishaq Khan and Prime Minister Nawaz Sharif in Pakistan. The resignations occurred after months of political infighting when the army demanded the President and Prime Minister resign and call for new elections. An interim prime minister was installed, but uncertainty about Pakistan's political and economic future remained high prior to the next round of elections.

The resignations studied in this paper are those in which leaders left office because of poor performance, public discontent and popular protests. It is therefore not unreasonable to expect the political actions preceding the resignations to have similarly large effects on financial markets.[10] To examine this, we explore all resignations that were driven by significant popular demonstrations, riots, non-violent civil resistance and other forms of public discontent in Public Protests in the appendix.[11] When taking directionality into account, it appears that public protests have no effect on stock returns. However, this occurs because some political movements increase stock prices while others decrease them, and the absolute value of stock returns are approximately 1.5% higher during public protests (see Table 4.10).[12]

[10] Indeed, corporate investors in the 2013 MIGA *World Investment and Political Risk* ranked civil disturbances as the fourth most concerning type of political risk.

[11] The set of resignations includes all those listed in either the Coup d'etat Events Handbook or the Archigos Version 4.1 data set with available financial data. In practice, this is the 2011 Egyptian Revolution and the list of resignations in Table 4.5.

[12] These results hold when controlling for emerging market index fluctuations.

Exploring possible mechanisms

While markets may generally dislike political instability, the immediate effect of regime changes on capital access may not always be unpredictable. For example, investors may generally value democracy if it is perceived to provide stronger property rights and lower susceptibility to capital appropriation (North and Weingast 1989; Przeworski and Wallerstein 1982; Svensson 1998). In addition, when the successor in an irregular regime change is clear, investors may have strong priors about the effect of the new leader on economic and/or market performance. Two possible mechanisms that could be driving the differences may therefore be: (1) whether the regime change is associated with an authoritarian or democratic shift in governance, and (2) whether a new leader is clearly more pro or anti business than their predecessor.

We first attempt to explore these mechanisms by aggregating all of the events in our sample by whether they resulted in a shift in an authoritarian or democratic direction.[13] We find suggestive evidence that regime changes associated with authoritarian shifts are on average perceived negatively by investors (see Table 4.6), while those that move governance in a democratic direction are on average perceived positively (see Table 4.7). We refer to this evidence as suggestive despite statistical significance due to the small sample of cases which fit these criteria, particularly with regard to democratic shifts. We observe ten cases associated with an authoritarian shift in governance. Only two of these events result in positive CARs, and neither are significantly different from zero. Five cases resulted in a democratic shift in governance. Three of these five cases result in positive CARs. Of the two negative CARs, one is not significantly different from zero, and the other is associated with a forced market closure that lasted 17 days. However, the majority of the positive returns from democratic regime changes come from the 12% positive CAR associated with the resignation of Ferdinand Marcos in the Philippines in 1986.

[13]As defined by the Polity project.

Table 4.6: Abnormal returns following authoritarian regime changes

		Post-Event CAR			Pre-Event CAR		Days to
Country	Event Date	(0,0)	(0,6)	(0,19)	(-1,-7)	(-1,-20)	rebound
Bangladesh	01/11/2007	−0.320	10.351	14.883	−0.896	2.250	2
		(1.166)	(3.086)	(5.217)	(3.086)	(5.217)	
Egypt	07/03/2013	−0.346	5.169	7.144	6.776	−4.869	2
		(1.515)	(4.009)	(6.776)	(4.009)	(6.776)	
Nepal	10/04/2002	0.090	1.563	5.567	−1.014	−0.493	2
		(1.206)	(3.190)	(5.392)	(3.190)	(5.392)	
Pakistan	10/14/1999	−7.737	−9.431	−7.130	4.151	4.900	36
		(1.943)	(5.141)	(8.690)	(5.141)	(8.690)	
Peru	04/06/1992	−6.819	−5.814	−25.027	−2.075	−10.519	5
		(2.210)	(5.848)	(9.885)	(5.848)	(9.885)	
Thailand	10/06/1976	−0.541	0.837	0.731	0.001	0.713	3
		(0.639)	(1.691)	(2.859)	(1.691)	(2.859)	
Thailand	02/25/1991	−7.326	2.860	14.162	6.326	26.262	7
		(2.884)	(7.631)	(12.899)	(7.631)	(12.899)	
Thailand	09/19/2006	−0.481	−2.640	0.111	1.848	0.131	17
		(1.094)	(2.894)	(4.892)	(2.894)	(4.892)	
Thailand	05/23/2014	−0.571	2.800	4.591	2.350	−0.424	5
		(1.201)	(3.177)	(5.370)	(3.177)	(5.370)	
Turkey	06/30/1997	2.010	−2.861	−7.629	12.876	4.532	
		(3.015)	(7.976)	(13.481)	(7.976)	(13.481)	
Mean		−2.204	0.283	0.740	3.034	2.248	8
		(0.585)	(1.548)	(2.616)	(1.548)	(2.616)	

Notes: Standard errors are in parentheses. "Days to rebound" is the number of trading days following a negative stock return for the national stock index to return to pre-event level (it is calculated if the price decreases on the event day, not if the event day abnormal return is negative). Returns are inflation adjusted.

A similar analysis is not possible for shifts from pro or anti business leaders, as examples of such clear and plausibly exogenous shifts in leader economic ideology do not exist in our sample.[14] We therefore look outside of the sample above and conduct an in-depth case study of a seemingly pro-business and anti-socialist coup followed by the reinstatement of a socialist government: the 2002 failed coup against Hugo Chavez in Venezuela.

The ultimately failed Venezuelan coup against Hugo Chavez replaced a left-wing populist government with a new pro-business president, and therefore provides a natural test of the effects of both pro-business and anti-business regime changes separately from simple uncertainty because investors reacted to an expected regime change twice: first, when Chavez

[14]Based on matching our cases with codings from the The Ideology of Heads of Government (HOG) database, as well as surveys of news reports on the day of each event. In all cases, no clear economic ideological shift can be identified.

Table 4.7: Abnormal returns following democratic regime changes

Country	Event Date	Post-Event CAR (0,0)	Post-Event CAR (0,6)	Post-Event CAR (0,19)	Pre-Event CAR (-1,-7)	Pre-Event CAR (-1,-20)	Days to rebound
Indonesia	05/20/1998	2.817	4.296	4.543	−2.695	−17.868	
		(3.392)	(8.974)	(15.168)	(8.974)	(15.168)	
Nepal	04/25/2006	1.915	8.132	9.937	−1.951	−4.205	
		(0.665)	(1.760)	(2.975)	(1.760)	(2.975)	
Philippines	02/26/1986	12.938	21.473	23.086	−1.847	−6.884	
		(0.477)	(1.263)	(2.134)	(1.263)	(2.134)	
Thailand	10/20/1977	−0.951	4.096	7.290	9.961	10.198	2
		(1.232)	(3.260)	(5.510)	(3.260)	(5.510)	
Tunisia	01/31/2011	−2.705	2.982	−11.787	−13.610	−13.445	5
		(0.671)	(1.776)	(3.002)	(1.776)	(3.002)	
Mean		2.803	8.196	6.614	−2.028	−6.441	3
		(0.752)	(1.990)	(3.364)	(1.990)	(3.364)	

Notes: Standard errors are in parentheses. “Days to rebound” is the number of trading days following a negative stock return for the national stock index to return to pre-event level (it is calculated if the price decreases on the event day, not if the event day abnormal return is negative). Returns are inflation adjusted.

was ousted, and second, when he was reinstated. On the evening of April 11, 2002, coup plotters removed Chavez from office and later detained him. Pedro Carmona, a Venezuelan economist and business leader, was named the transitional President of Venezuela. Two days later, on April 13, 2002, a popular uprising led to Chavez’s reinstatement as president. This provides an estimate of the market’s valuation of a transition from the Chavez regime to the Carmona regime and its valuation of a transition from the Carmona regime back to the Chavez regime. By extension, it provides an estimate of the impact of a shift from a left-wing populist government to a pro business regime in an emerging market.

Figure 4.3 provides graphical evidence on the effect of the coup attempt. The top panel shows CARs for the 10 days prior to and following the event, along with 95% confidence intervals. The daily ARs and corresponding confidence intervals are displayed in the bottom panel. The abnormal return on April 12, the first trading day in which investors could react to the coup, was +10%. The market reacted in the opposite direction to Chavez’s reinstatement as president: the abnormal return on the next trading day, April 15, 2002, was -8%.

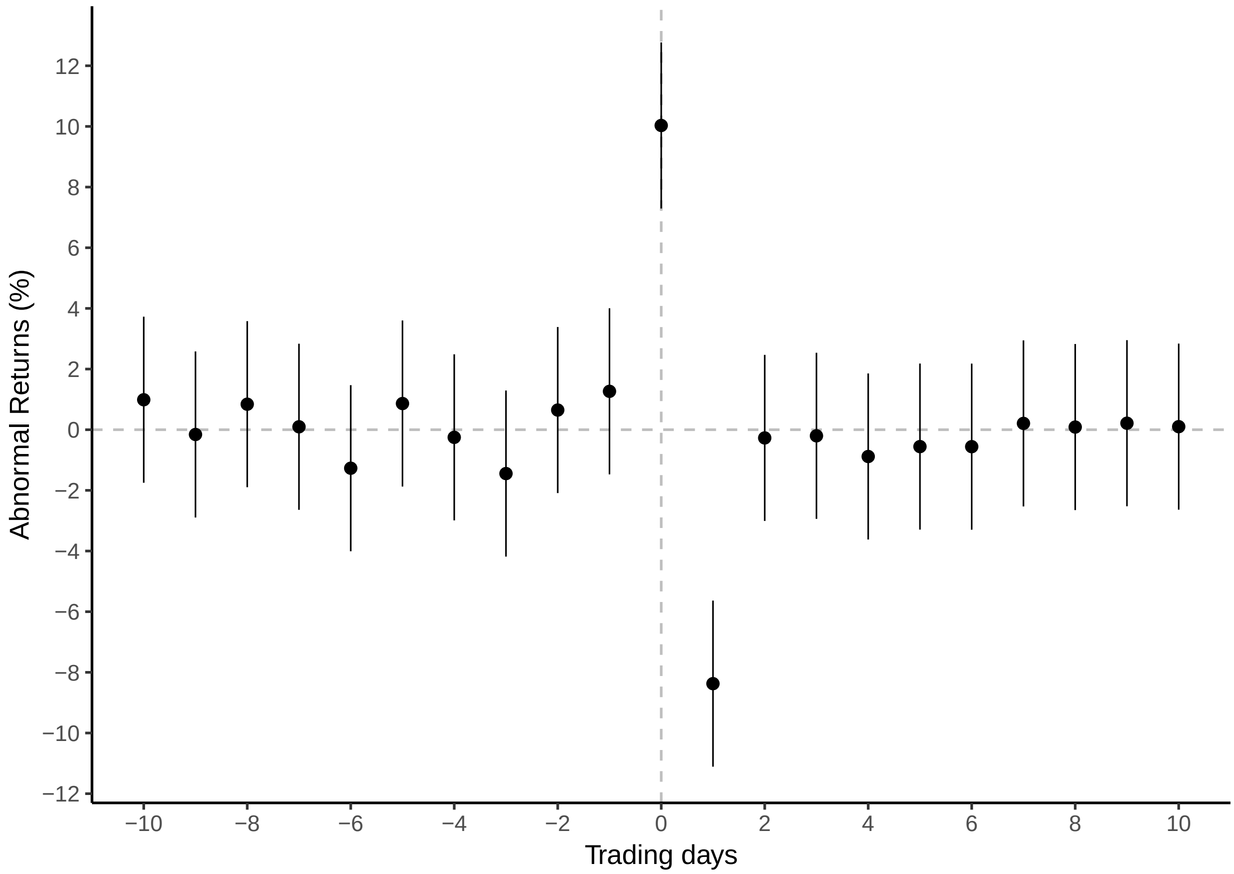


Figure 4.3: Abnormal returns surrounding the 2002 Venezuelan coup attempt

The results in Figure 4.3 are particularly striking given the discrepancy between the ARs on event days 0 and 1 and all other days. Consistent with our earlier findings that coups tend to have pre and post-event CARs that are statistically indistinguishable from zero, the only days on which the 2002 Venezuelan coup d'etat attempt ARs are statistically different from zero is on event days 0 and 1 after the coup attempt. The almost 0% 10-day CAR preceding the coup makes this an ideal case as it implies that investors were completely unaware of the coup plot, increasing our confidence that the abnormal returns capture the true effect of the Chavez to Carmona and Carmona to Chavez regime changes on capital flows. This failed coup therefore demonstrates a large positive market reaction to the attempted overthrow of a socialist leader, and an equally large negative reaction to his reinstatement. More generally, the large magnitudes and precision of these effects suggest that investors primarily value transition to a pro-business government regime, regardless of how the regime change is achieved.

Robustness

There are some potential concerns with the results in the Coups, Assassinations, and Resignations sections. First, the abnormal returns could have been driven by factors unrelated to the regime changes. Second, the true effects of regime changes on firm value may be underestimated if investors had apriori information. Third, the reported means are based on small sample sizes so confidence intervals based on normally distributed abnormal returns may be inappropriate.

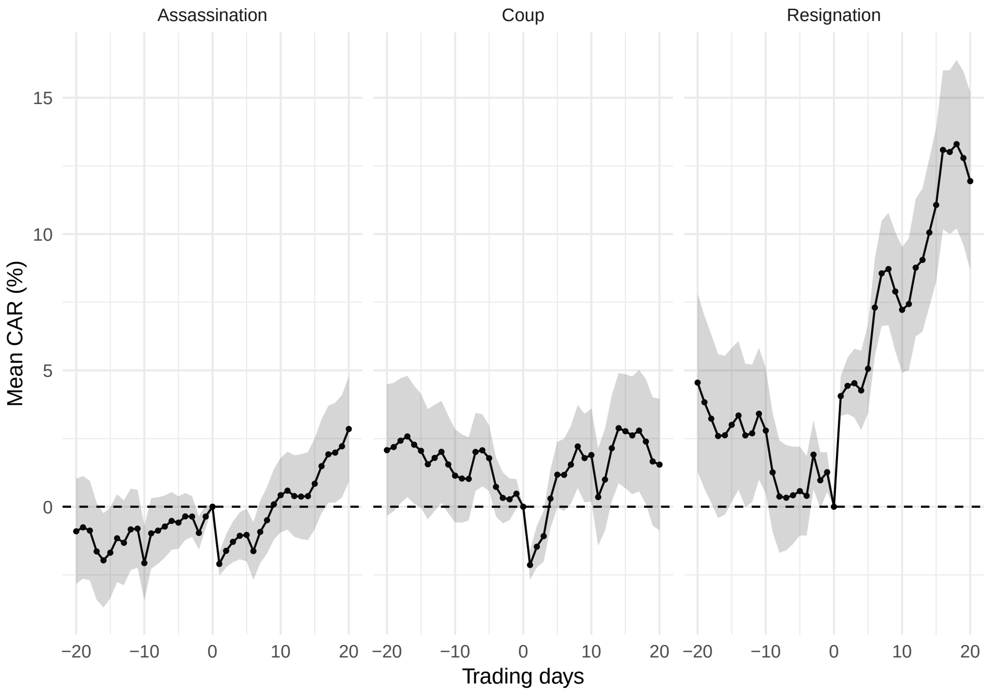
We explore these concerns in two ways. First, we reestimate mean CARs on a set of time-shifted placebo dates, with means computed across all events for each type of regime change. We shift event dates surrounding the actual event date backwards and forwards in increments of five days (-20, 15, 10, 5, 0, 5, 10, 15, and 20 days). In addition, we extend the forward shifted event dates to one year (110, 195, 285, and 365 days) to capture dates that are likely to be completely unaffected by the regime change. The general intuition is that we should not observe significant abnormal returns when performing an identical test on dates where no intervention (i.e. a politically unstable event) occurred. Observing such effects would call the research design and modeling assumptions into question, and raise concerns that the abnormal returns were caused by factors other than the regime changes.

Figure 4.4 compares mean CARs estimated using the actual event date (Figure 4.4a) to CARs estimated with the event date shifted 1-year (365 days) into the future (Figure 4.4b). Figure 4.4a—which reproduces the tabular results presented in the Coups, Assassinations, and Resignations sections—shows that assassinations and coups are associated with negative event day ARs while resignations are associated with positive event day ARs. In contrast, there are no discernible abnormal returns in Figure 4.4b. The event day ARs are considerably smaller in magnitude and are not statistically different from zero for either coups or resignations. Moreover, while CARs estimated using the actual event date for resignations trend upwards following the event day, there is no consistent trend in the placebo analysis. Figure 4.4 therefore suggests that the main results are not merely an artifact of the data.

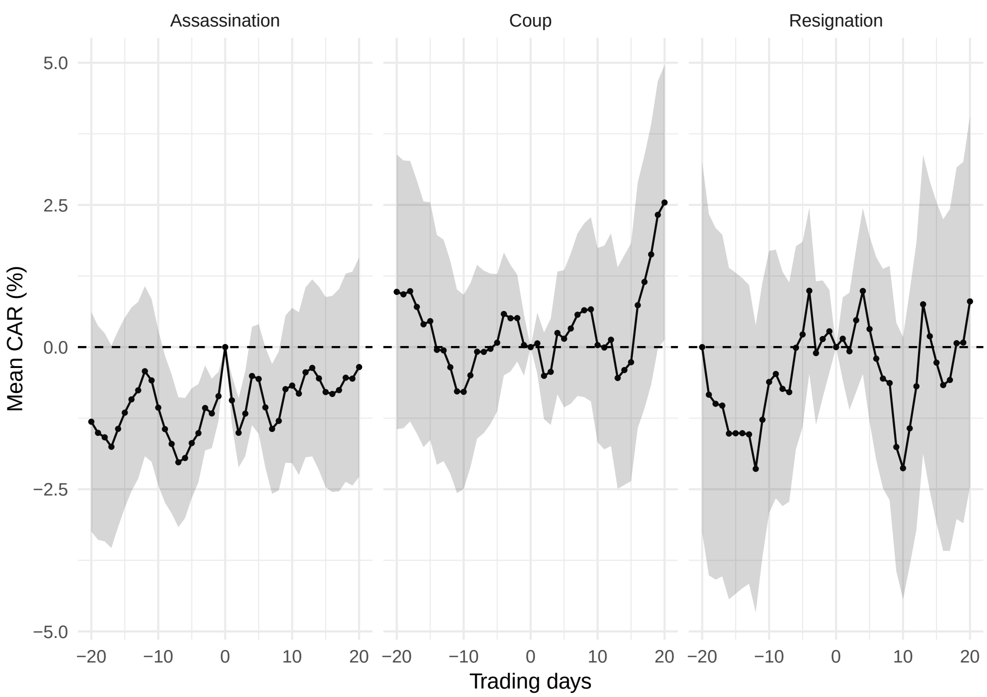
To ensure that results of the placebo test based on 1-year did not occur by chance, Figure 4.6 plots event-day abnormal against the number of days shifted. There are a few instances in which there are statistically significant (at the 5% level) ARs in the same direction as those on the actual event day, but they are always smaller in magnitude than the ARs estimated using the actual event date. Most of the statistically significant placebo estimates also occur when dates are shifted within the post-event window (days 0-20), a period during which stock returns remain volatile and, in the case of resignations, there is a consistent upward trend in the CARs. When dates are shifted forward further ($\geq$ 110 days), the event day AR is only statistically significant in one case (day 365 for assassinations). Overall, these results reinforce the main results: the ARs on the actual event date capture most of effect of the regime change, although effects can sometimes persist in the short event window following the event date.

These figures also provide evidence on the extent to which regime changes appear to be unexpected. For instance, the CARs prior to assassinations and coups presented in Figure 4.4a tend to be close to zero, suggesting that investors were unaware that a negative event was likely to occur in the coming days. For resignations, CARs trend downward in the 10 days prior to the event; however, if investors were aware that a resignation were about to occur one would expect the pre-event CARs to be positive given the positive CARs observed in the post-event window. There is thus little evidence that the CARs in the post-event window are not capturing most of the effect caused by the regime changes.

Second, we create a synthetic control portfolio for each event based on the techniques introduced in Abadie and Gardeazabal (2003) and Abadie, Diamond and Hainmueller (2010). Each country is given a weight which represents its influence in the synthetic control portfolio. The weight is chosen so that the daily returns and the variance of the daily returns of the



(a) Actual event date



(b) Event date shifted forward by one year

Figure 4.4: Mean cumulative abnormal returns by type of regime change

control portfolio and the event country are most similar in the estimation window. The set of possible countries in the control portfolio consists of all countries listed in Table 4.1.[15]

Non-parametric statistical techniques that are free from distributional assumptions are used to address concerns about inferences from small sample sizes. We employ the sign and the rank tests, which are based on the sign and the rank of the event day ARs respectively.[16] Both tests are less influenced by departures from normality than statistics based on traditional t-tests such as those reported earlier in this paper. Table 4.8 compares event day ARs as well as "abnormal absolute returns" between the event country and the synthetic control portfolio using the non-parametric methods discussed above. The "abnormal absolute returns" are abnormal returns for the absolute value of stock returns. This is done to combine events since resignations tend to increase returns while assassinations and coups tend to decrease them. The idea that the absolute value of returns might increase during irregular regime changes is similar to the finding that volatility increases and is consistent with Figure 4.1.

Table 4.8: Non-parametric tests of the impact of regime changes

	Regime Change Country			Synthetic Control Portfolio			
Event Type	Mean CAR (0,0)	Rank p-value	Sign p-value	Mean CAR (0,0)	Rank p-value	Sign p-value	Wilcoxon Rank Test p-Value
Coups	−2.137	0.006	0.022	0.024	0.702	1.000	0.002
Assassinations	−2.098	0.001	0.070	−0.125	0.255	0.453	0.078
Resignations	4.059	0.010	0.118	0.366	0.778	0.607	0.048
All (Absolute Value)	2.410	0.002	0.033	−0.079	0.879	0.955	0.000

Notes: Estimates for assassinations do not include the assassination of U.S. president William McKinley in 1901 because no control portfolios are available.

As shown in Table 4.8, the mean event day abnormal returns for coups, assassinations and resignations are all statistically different from zero at the 1% level using the rank test statistic and the abnormal returns for coups and assassinations are significant at at least the 10% level using the sign test. In addition, abnormal absolute returns for all events are

[15]See Appendix A for a more formal explanation.

[16]See section 8 in MacKinlay (1997) for more details.

statistically significant at at least the 5% level using both the rank and sign tests. On the other hand, the event day abnormal returns for the control portfolio are never statistically different from zero at even the 10% level using the rank or sign tests. Finally, the difference in means between the regime change country and the control portfolio are statistically different from zero for coups (1% level), assassinations (10% level), resignations (5% level), and all events combined (1% level) when using two-sided p-values from the Wilcoxon rank test.[17] In sum, the synthetic control and small sample tests suggest that the main results are not a result of deviations from normality or confounding world events.

Conclusion

Conventional theory suggests political instability causes capital flight. We show that this is not always the case. Unexpected changes in ruler virtually always increase volatility, but when political instability is broken down into types, evidence emerges that some types of instability increase capital access while other types lead to capital flight.

Coups and other types of regime changes remain common, highlighting their relevance for economic and political development. Our sample consists of 5 coups, 1 assassination, 7 resignations, and 7 instances of public protect since 2000. The Arab spring is perhaps the most notable, with protests in response to oppressive regimes and low living standards spreading throughout the Middle East in late 2010. There have also been a number of failed coups such as the 2002 coup attempt against Hugo Chavez and the 2016 failed coup in Turkey.[18] Yet despite their frequency there is little evidence on their economic consequences.

We examine changes in stock market returns surrounding politically unstable domestic events to show that resignations increase capital access on average, but assassinations and coups usually cause capital flight. We also find evidence that capital markets tend to prefer democratic regime changes to authoritarian shifts, but that even democratically subversive

[17]The Wilcoxon rank test is a non-parametric statistical technique that can be used to compare differences between matched samples.

[18]The Turkish coup attempt led to negative event-day CARs of approximately -7%. See Figure A2 for a visual depiction.

coups can boost capital access if the instigators are clearly pro-business.

The results are consistent with the idea that perceptions of government competence and changes in government have large impacts on investor confidence. But although the effect of regime changes and protests on stock volatility is substantial, the effect on the direction of capital flows is dependent on the expected impact of the regime change on economic growth.

There are a number of avenues for future research. First, more research is needed to identify the pro and anti market characteristics of regime changes. Second, more work is needed to determine the extent to which financial market flows translate to broader economic development outcomes. It remains unclear whether the direction of the effects of different types of regime changes on outcomes such as economic growth, investment, debt, inflation, infant mortality, and years of schooling are consistent with their impact on financial flows, or if investor perceptions are at odds with certain development goals.

Appendix

Public Protests

The resignations studied in this paper are those in which leaders left office because of poor performance, public discontent and popular protests. It is therefore not unreasonable to expect the political actions preceding the resignations to have similarly large effects on financial markets.[19] To examine this, we explore all resignations that were driven by significant popular demonstrations, riots, non-violent civil resistance and other forms of public discontent (see Table 4.9).[20]

Table 4.9: List of public protests preceding resignations

Country	Name	Start Date	End Date
Philippines	EDSA 1/Yellow Revolution	2/22/1986	2/25/1986
Bangladesh	Bangladeshi Spring of 1990	11/27/1990	12/7/1990
Thailand	Black May	5/17/1992	5/20/1992
Indonesia	Indonesian Riots	5/12/1998	5/21/1998
Philippines	EDSA II	1/17/2001	1/20/2001
Argentina	Argentina Riots	12/16/2001	12/20/2001
Ukraine	Orange Revolution	11/22/2004	1/23/2005
Ecuador	Ecuadorian Protests	4/13/2005	4/20/2005
Nepal	Nepalese People's Revolution	4/6/2006	4/24/2006
Tunisia	Tunisian Revolution	12/18/2010	1/14/2011
Egypt	Egyptian Revolution	1/25/2011	2/11/2001

A recent example of a popular uprising preceding a resignation is the 2011 Egyptian Revolution that resulted in the overthrow of President Hosni Mubarak's regime.[21] Clashes between security forces and protestors led to the deaths of hundreds of citizens and injuries to thousands more. The uprising began on January 25, 2011 when millions of protestors demanded the overthrow of the Egyptian leadership. Examples of public discontent included

[19]Indeed, corporate investors in the 2013 MIGA *World Investment and Political Risk* ranked civil disturbances as the fourth most concerning type of political risk.

[20]The set of resignations includes all those listed in either the Coup d'etat Events Handbook or the Archigos Version 4.1 data set with available financial data. In practice, this is the 2011 Egyptian Revolution and the list of resignations in Table 4.5.

[21]Abnormal returns for this event are not shown in Table 4.5 because the stock market was closed on the day of Mubarak's resignation.

demonstrations, marches, riots, non-violent civil disobedience, and labor strikes.

The short-term impact of the Egyptian Revolution on the economy was disastrous. As shown in Figure 4.5, abnormal returns on the Egyptian Stock Exchange Index (EGX 30) were around -7% on January 26th and -10% the day after. To prevent further decline during the uprising, the Egyptian Stock Exchange closed at the end of trading on January 27th. President Mubarak resigned on February 11, but the market remained closed until March 23, when CARs declined by another 9%, before rebounding slightly thereafter.

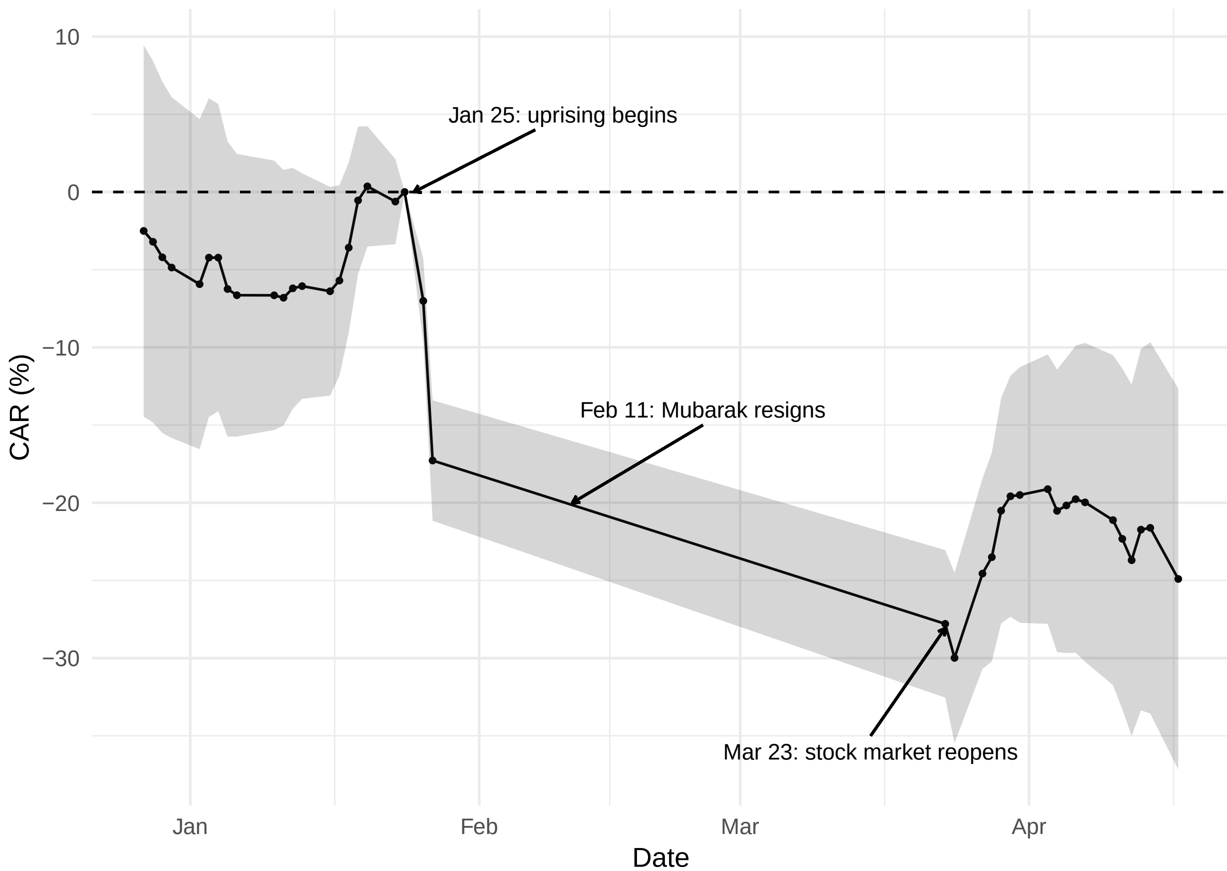


Figure 4.5: Cumulative abnormal returns during the Egyptian revolution

An important question is whether other popular uprisings have had similar adverse economic consequences. To examine this, we explore all resignations that were driven by significant public protests.[22] Public protests include popular demonstrations, riots, non-violent

[22]The set of resignations includes all those listed in either the Coup d'etat Events Handbook or the

civil resistance and other forms of public discontent. We find that both volatility and the absolute value of returns increase during times of protest. Similarly to coups, however, the direction of returns is dependent upon the nature of the protest in question.

The start and end dates in Table 4.10 are the dates that protests began and leader's resigned respectively. Resignations caused by popular uprisings were identified by examining the descriptions in the Coup d'etat Events Handbook and Archigos Version 4.1. Additional Lexis Nexis searches were used to verify these descriptions.

In Table 4.10, we examine whether public protests influence stock prices. The variable *Protest* is equal to 1 during the dates in which citizens participate in political activities demanding the resignation of the executive and 0 otherwise. Non-protest dates are the 250 days prior to the start dates and after the end dates listed in Table A.1.[23]

Column (1) suggests that public protests have no effect on stock returns. However, this occurs because some political movements increase stock prices while others decrease them. As shown in column (2), the absolute value of stock returns are approximately 1.5% higher during public protests. These estimates would be biased if protest dates are correlated with higher world or regional stock market indices. To address this potential confounder, column (3) controls for returns on the S&P/IFC Emerging Markets Investable Composite Stock Index. The coefficient on *Protest* barely changes and the absolute value of returns are still about 1.5% higher during public protests. Finally, column (4) shows that stock volatility is approximately 1 percentage point higher during political movements.[24]

We therefore find that both volatility and the absolute value of returns increase during times of protest. Similarly to coups, however, the direction of returns is dependent upon the nature of the protest in question.

Archigos Version 4.1 data set with available financial data. In practice, this is the 2011 Egyptian Revolution and the list or resignations in Table 4.5.

[23]The volatility estimates used as the dependent variable in column (4) are estimated on the 250 days prior to the start date, the protest dates, and the 250 days following the end date.

[24]Volatility estimation methodology is described in detail in Volatility.

Table 4.10: Effect of public protests on stock prices

	Returns	Absolute Value of Returns		Volatility
	(1)	(2)	(3)	(4)
Protest	0.261	1.485	1.313	0.891
	(0.700)	(0.412)	(0.387)	(0.346)
Emerging market index			0.075	
			(0.058)	
Event fixed effect?	Yes	Yes	Yes	Yes
Observations	3,537	3,537	2,676	3,537
Events	11	11	8	11

Notes: Standard errors clustered by event are in parentheses.

Time-shifted placebo test

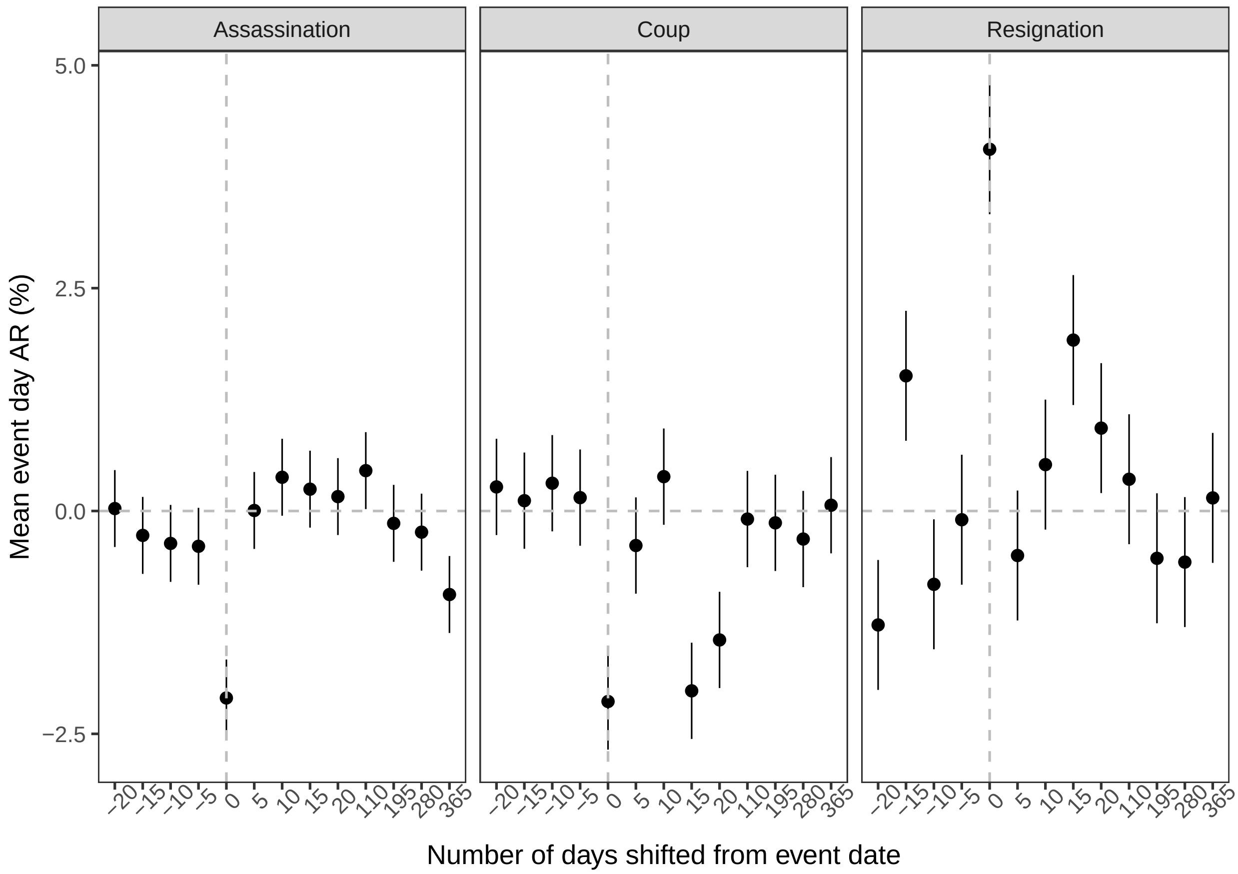


Figure 4.6: Time-shifted placebo sensitivity analysis of mean event day abnormal return by type of regime change

Graphical depictions of additional events

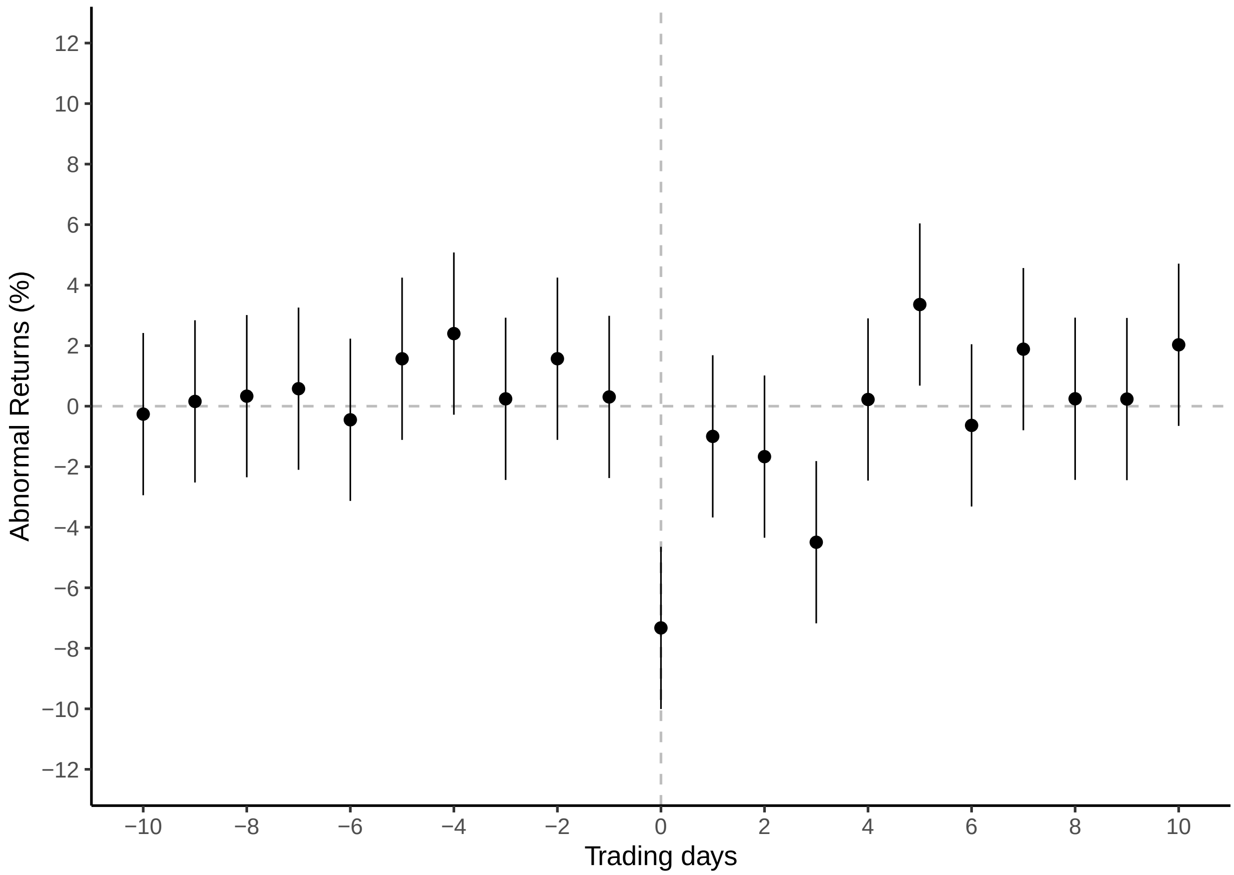


Figure 4.7: Abnormal returns surrounding the 2016 Turkish coup attempt

Synthetic Control Portfolio

Let $\boldsymbol{R}_k$ be the vector of returns for the event country in the estimation window, $\boldsymbol{R}_{-k}$ be the vector of returns for all other countries in the estimation window, $\boldsymbol{X}_1 = (\boldsymbol{R}_k, \mathrm{Var}(\boldsymbol{R}_k))$, $\boldsymbol{X}_0 = (\boldsymbol{R}_{-k}, \mathrm{Var}(\boldsymbol{R}_{-k}))$, and $\boldsymbol{W}_{-k}$ be a $((N-1)\times 1)$ vector of weights where N is the number of countries listed in Table 4.1. Then $\boldsymbol{W}^*$ is chosen to minimize $(\boldsymbol{X}_1 - \boldsymbol{X}_0\boldsymbol{W})'\boldsymbol{V}(\boldsymbol{X}_1 - \boldsymbol{X}_0\boldsymbol{W})$ subject to $w_i \geq 0$ $(i = 1, 2, \ldots, N-1)$ and $\sum_i^{N-1} w_i = 1$, and the vector $\boldsymbol{V}$ is chosen so that stock returns for the control portfolio during the estimation window are are close as possible to the event country.[25]

[25] See Abadie and Gardeazabal (2003) for further details.

chapter 5

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program." *Journal of the American Statistical Association* 105(490).

Abadie, Alberto and Javier Gardeazabal. 2003. "The economic costs of conflict: A case study of the Basque Country." *American economic review* pp. 113–132.

Abramson, Scott F, Korhan Koçak and Asya Magazinnik. 2019. "What Do We Learn About Voter Preferences From Conjoint Experiments?" *Working paper* .

URL: https://scholar.princeton.edu/sites/default/files/kkocak/files/conjoint_draft.pdf

Adida, Claire, Jessica Gottlieb, Eric Kramon and Gwyneth McClendon. 2019. Under what conditions does performance information influence voting behavior? Lessons from Benin. In *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I*, ed. Thad Dunning, Guy Grossman, Macartan Humphreys, Susan D Hyde, Craig McIntosh and Gareth Nellis. Cambridge: Cambridge University Press chapter 4, pp. 81–116.

Agerberg, Mattias. 2020. "The lesser evil? Corruption voting and the importance of clean alternatives." *Comparative Political Studies* 53(2):253–287.

Aisen, Ari and Francisco José Veiga. 2013. "How does political instability affect economic growth?" *European Journal of Political Economy* 29:151–167.

Alesina, A. and R. Perotti. 1996. "Income distribution, political instability, and investment." *European Economic Review* 40(6):1203–1228.

Alesina, A., S. Özler, N. Roubini and P. Swagel. 1996. "Political instability and economic growth." *Journal of Economic Growth* 1(2):189–211.

Allcott, Hunt. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95(9-10):1082–1095.

Anduiza, Eva, Aina Gallego and Jordi Muñoz. 2013. "Turning a blind eye: Experimental evidence of partisan bias in attitudes toward corruption." *Comparative Political Studies* 46(12):1664–1692.

Ansolabehere, Stephen Daniel. 2012. "Movers, stayers, and registration: Why age is correlated with registration in the US." *Quarterly Journal of Political Science* .

Arias, Eric, Horacio Larreguy, John Marshall and Pablo Querubin. 2018. Priors Rule: When do Malfeasance Revelations Help or Hurt Incumbent Parties? Technical report National Bureau of Economic Research.

URL: *https://www.nber.org/papers/w24888*

Arias, Eric, Pablo Balán, Horacio Larreguy, John Marshall and Pablo Querubín. 2019. "Information provision, voter coordination, and electoral accountability: Evidence from Mexican social networks." *American Political Science Review* 113(2):475–498.

Arvate, Paulo and Sergio Mittlaender. 2017. "Condemning corruption while condoning inefficiency: an experimental investigation into voting behavior." *Public Choice* 172(3-4):399–419.

Asai, Kentaro, Kei Kawai and Jun Nakabayashi. 2021. "Regulatory capture in public procurement: Evidence from revolving door bureaucrats in Japan." *Journal of Economic Behavior & Organization* 186:328–343.

Avenburg, Alejandro. 2019. "Public Costs versus Private Gain: Assessing the Effect of Different Types of Information about Corruption Incidents on Electoral Accountability." *Journal of Politics in Latin America* 11(1):71–108.

Aytaç, S Erdem and Susan C Stokes. 2019. *Why Bother?: Rethinking Participation in Elections and Protests*. Cambridge University Press.

Azfar, Omar and William Robert Nelson. 2007. "Transparency, wages, and the separation of powers: An experimental analysis of corruption." *Public Choice* 130(3-4):471–493.

Baker, Scott R, Nicholas Bloom and Steven J Davis. 2016. "Measuring economic policy uncertainty." *The Quarterly Journal of Economics* 131(4):1593–1636.

Banerjee, Abhijit, Donald Green, Jennifer Green and Rohini Pande. 2010. Can voters be primed to choose better legislators? Experimental evidence from rural India. In *Presented and the Political Economics Seminar, Stanford University*. Citeseer.

URL: *https://pdfs.semanticscholar.org/a204/eb3e92d382dd312790f47df9aefde657fd13.pdf*

Banerjee, Abhijit, Donald P Green, Jeffery McManus and Rohini Pande. 2014. "Are poor voters indifferent to whether elected leaders are criminal or corrupt? A vignette experiment in rural India." *Political Communication* 31(3):391–407.

Banerjee, Abhijit, Selvan Kumar, Rohini Pande and Felix Su. 2011. "Do informed voters make better choices? Experimental evidence from urban India." *Unpublished manuscript* .

URL: *https://pdfs.semanticscholar.org/45aa/1e275e770103f7a7d7b02ba86fb46afa89c0.pdf*

Barbosa, Klenio and Stephane Straub. 2020. "The Value of Revolving Doors in Public Procurement." *TSE Working Paper* .

Barro, Robert J. 1991. "Economic growth in a cross section of countries." *The quarterly journal of economics* 106(2):407–443.

Bialkowski, Jedrzej, Katrin Gottschalk and Tomasz Piotr Wisniewski. 2008. "Stock market volatility around national elections." *Journal of Banking & Finance* 32(9):1941–1953.

Blair, Graeme, Alexander Coppock and Margaret Moor. 2018. "When to Worry About Sensitivity Bias: Evidence from 30 Years of List Experiments." *Working Paper* .
URL: *https://alexandercoppock.com/papers/BCM_list.pdf*

Blanes i Vidal, Jordi, Mirko Draca and Christian Fons-Rosen. 2012. "Revolving door lobbyists." *American Economic Review* 102(7):3731–48.

Boas, Taylor C, F Daniel Hidalgo and Marcus André Melo. 2019. "Norms versus action: Why voters fail to sanction malfeasance in Brazil." *American Journal of Political Science* 63(2):385–400.

Boas, Taylor C, F Daniel Hidalgo and Neal P Richardson. 2014. "The spoils of victory: campaign donations and government contracts in Brazil." *The Journal of Politics* 76(2):415–429.

Botero, Sandra, Rodrigo Castro Cornejo, Laura Gamboa, Nara Pavao and David W Nickerson. 2015. "Says who? An experiment on allegations of corruption and credibility of sources." *Political Research Quarterly* 68(3):493–504.

Boutchkova, Maria, Hitesh Doshi, Art Durnev and Alexander Molchanov. 2012. "Precarious politics and return volatility." *The Review of Financial Studies* 25(4):1111–1154.

Bovens, Luc. 2009. The ethics of nudge. In *Preference change*. Springer pp. 207–219.

Breitenstein, Sofia. 2019. "Choosing the crook: A conjoint experiment on voting for corrupt politicians." *Research & Politics* 6(1):1–8.

Broockman, David E, Joshua L Kalla and Jasjeet S Sekhon. 2017. "The design of field experiments with survey outcomes: A framework for selecting more efficient, robust, and ethical designs." *Political Analysis* 25(4):435–464.

Brueckner, Jan K. 2009. Government land use interventions: An economic analysis. In *Urban* *land markets*. Springer pp. 3–23.

Buntaine, Mark T, Ryan Jablonski, Daniel L Nielson and Paula M Pickering. 2018. "SMS texts on corruption help Ugandan voters hold elected councillors accountable at the polls." *Proceedings of the National Academy of Sciences* 115(26):6668–6673.

Calder, Kent E. 1989. "Elites in an equalizing role: ex-bureaucrats as coordinators and intermediaries in the Japanese government-business relationship." *Comparative Politics* 21(4):379–403.

Camerer, Colin. 2011. "The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List." *SSRN 1977749* .

URL: *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1977749*

Campbell, J.Y., A.W. Lo and A.C. MacKinlay. 1997. *The econometrics of financial markets*. Vol. 1 princeton University press Princeton, NJ.

Carlson, Matthew M and Steven R Reed. 2018. *Political corruption and scandals in Japan*. Cornell University Press.

Carter, Evan C, Felix D Schönbrodt, Will M Gervais and Joseph Hilgard. 2019. "Correcting for bias in psychology: A comparison of meta-analytic methods." *Advances in Methods and Practices in Psychological Science* 2(2):115–144.

Cervero, Robert and Jin Murakami. 2010. "Effects of built environments on vehicle miles traveled: evidence from 370 US urbanized areas." *Environment and planning A* 42(2):400–418.

Chang, Eric CC, Miriam A Golden and Seth J Hill. 2010. "Legislative malfeasance and political accountability." *World Politics* 62(2):177–220.

Charette, Allison, Chris Herbert, Andrew Jakabovics, Ellen Tracy Marya and Daniel T McCue. 2015. "Projecting trends in severely cost-burdened renters: 2015–2025." *Harvard University's Joint Center for Housing Studies and Enterprise Community Partners Inc* .

Chauchard, Simon, Marko Klašnja and SP Harish. 2019. "Getting Rich Too Fast? Voters'

Reactions to Politicians' Wealth Accumulation." *The Journal of Politics* 81(4):1197–1209.

Chester, Mikhail, Andrew Fraser, Juan Matute, Carolyn Flower and Ram Pendyala. 2015. "Parking infrastructure: A constraint on or opportunity for urban redevelopment? A study of Los Angeles County parking supply and growth." *Journal of the American Planning Association* 81(4):268–286.

Chong, Alberto, Ana L De La O, Dean Karlan and Leonard Wantchekon. 2014. "Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification." *The Journal of Politics* 77(1):55–71.

Citrin, Jack, Beth Reingold and Donald P Green. 1990. "American identity and the politics of ethnic change." *The Journal of Politics* 52(4):1124–1154.

Citrin, Jack, Donald P Green, Christopher Muste and Cara Wong. 1997. "Public opinion toward immigration reform: The role of economic motivations." *The Journal of Politics* 59(3):858–881.

Cohen, Benjamin H, Dietrich Domanski, Ingo Fender and Hyun Song Shin. 2017. "Global liquidity: a selective review." *Annual Review of Economics* 9:587–612.

Cook, Scott J, Jude C Hays and Robert J Franzese. 2020. "Fixed effects in rare events data: a penalized maximum likelihood solution." *Political Science Research and Methods* 8(1):92–105.

Coppock, Alexander and Donald P Green. 2015. "Assessing the correspondence between experimental results obtained in the lab and field: A review of recent social science research." *Political Science Research and Methods* 3(1):113–131.

Dal Bó, Ernesto. 2006. "Regulatory capture: A review." *Oxford Review of Economic Policy* 22(2):203–225.

De Figueiredo, Miguel FP, F Daniel Hidalgo and Yuri Kasahara. 2011. "When do voters punish corrupt politicians? Experimental evidence from Brazil." *Working Paper* .

URL: *https://law.utexas.edu/wp-content/uploads/sites/25/figueiredo_when_do_voters_punish.pdf*

De la Cuesta, Brandon, Naoki Egami and Kosuke Imai. 2019. "Improving the External Validity of Conjoint Analysis: The Essential Role of Profile Distribution." *Working Paper* .

URL: *https://imai.fas.harvard.edu/research/files/conjoint.pdf*

De Rooij, Eline A, Donald P Green and Alan S Gerber. 2009. "Field experiments on political behavior and collective action." *Annual Review of Political Science* 12:389–395.

De Vries, Catherine E and Hector Solaz. 2017. "The electoral consequences of corruption." *Annual Review of Political Science* 20:391–408.

Derpanopoulos, G., E. Frantz, B. Geddes and J. Wright. 2016. "Are coups good for democracy." *Research and Politics* pp. 1–7.

Diet of Japan. 2012. "The Fukushima Nuclear Accident Independent Investigation Commission.".

Doberstein, Carey, Ross Hickey and Eric Li. 2016. "Nudging NIMBY: Do positive messages regarding the benefits of increased housing density influence resident stated housing development preferences?" *Land Use Policy* 54:276–289.

Dunning, Thad, Guy Grossman, Macartan Humphreys, Susan D Hyde, Craig McIntosh and Gareth Nellis. 2019. *Information, Accountability, and Cumulative Learning: Lessons from Metaketa I.* Cambridge: Cambridge University Press.

Duval, Sue and Richard Tweedie. 2000. "A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis." *Journal of the American Statistical Association* 95(449):89–98.

Eggers, Andrew C, Nick Vivyan and Markus Wagner. 2018. "Corruption, accountability, and gender: do female politicians face higher standards in public life?" *The Journal of Politics* 80(1):321–326.

Einstein, Katherine Levine, David Glick, Luisa Godinez Puig and Maxwell Palmer. 2021.

"Zoom Does Not Reduce Unequal Participation: Evidence from Public Meeting Minutes." *Working Paper* .

Einstein, Katherine Levine, Joseph T Ornstein and Maxwell Palmer. 2019. "Who Represents the Renters?" *Working Paper* .

Einstein, Katherine Levine, Maxwell Palmer and David M Glick. 2019. "Who participates in local government? Evidence from meeting minutes." *Perspectives on Politics* 17(1):28–46.

Enamorado, Ted, Benjamin Fifield and Kosuke Imai. 2019. "Using a probabilistic model to assist merging of large-scale administrative records." *American Political Science Review* 113(2):353–371.

Faccio, Mara. 2006. "Politically connected firms." *American Economic Review* 96(1):369–386.

Faccio, Mara, Ronald W Masulis and John J McConnell. 2006. "Political connections and corporate bailouts." *The Journal of Finance* 61(6):2597–2635.

Ferraz, Claudio and Frederico Finan. 2008. "Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes." *The Quarterly Journal of Economics* 123(2):703–745.

Fischel, William A. 2005. *The homevoter hypothesis: How home values influence local government taxation, school finance, and land-use policies.* Harvard University Press.

Fosu, Augustin Kwasi. 1992. "Political instability and economic growth: evidence from Sub-Saharan Africa." *Economic Development and Cultural Change* 40(4):829–841.

Franchino, Fabio and Francesco Zucchini. 2015. "Voting in a multi-dimensional space: a conjoint analysis employing valence and ideology attributes of candidates." *Political Science Research and Methods* 3(2):221–241.

Ganong, Peter and Daniel Shoag. 2017. "Why has regional income convergence in the US declined?" *Journal of Urban Economics* 102:76–90.

Gelman, Andrew and Jennifer Hill. 2006. *Data analysis using regression and multi-*

level/hierarchical models. Cambridge university press.

Gerber, Alan S, Donald P Green and Christopher W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American political Science review* pp. 33–48.

Girardi, Daniele and Samuel Bowles. 2018. "Institution shocks and economic outcomes: Allende's election, Pinochet's coup and the Santiago stock market." *Journal of Development Economics* 134:16–27.

Glaeser, Edward and Joseph Gyourko. 2018. "The economic implications of housing supply." *Journal of Economic Perspectives* 32(1):3–30.

Glaeser, Edward L and Joseph Gyourko. 2002. "The impact of zoning on housing affordability.".

Glaeser, Edward L, Joseph Gyourko and Raven E Saks. 2005*a*. "Why have housing prices gone up?" *American Economic Review* 95(2):329–333.

Glaeser, Edward L, Joseph Gyourko and Raven Saks. 2005*b*. "Why is Manhattan so expensive? Regulation and the rise in housing prices." *The Journal of Law and Economics* 48(2):331–369.

Green, Donald P, Adam Zelizer, David Kirby et al. 2018. "Publicizing Scandal: Results from Five Field Experiments." *Quarterly Journal of Political Science* 13(3):237–261.

Green, Donald P and Alan S Gerber. 2019. *Get out the vote: How to increase voter turnout*. Washington, D.C.: Brookings Institution Press.

Gupta, Dipak K. 1990. *The economics of political violence: The effect of political instability on economic growth*. Praeger.

Gyourko, Joseph and Raven Molloy. 2015. Regulation and housing supply. In *Handbook of regional and urban economics*. Vol. 5 Elsevier pp. 1289–1337.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political Analysis* 22(1):1–30.

Hainmueller, Jens, Dominik Hangartner and Teppei Yamamoto. 2015. "Validating vignette and conjoint survey experiments against real-world behavior." *Proceedings of the National Academy of Sciences* 112(8):2395–2400.

Hall, Andrew B and Jesse Yoder. Forthcoming. "Does homeownership influence political behavior? Evidence from administrative data." *The Journal of Politics* .

Hankinson, Michael. 2018. "When do renters behave like homeowners? High rent, price anxiety, and NIMBYism." *American Political Science Review* 112(3):473–493.

Hernández, Antonia. 2021. "Our Common Purpose: Reinventing American Democracy for the 21st Century." *National Civic Review* 110(1):29–37.

Horiuchi, Akiyoshi and Katsutoshi Shimizu. 2001. "Did amakudari undermine the effectiveness of regulator monitoring in Japan?" *Journal of Banking & Finance* 25(3):573–596.

Horiuchi, Yusaku, Zachary D Markovich and Teppei Yamamoto. 2018. "Can Conjoint Analysis Mitigate Social Desirability Bias?" *MIT Political Science Department Research Paper* .

URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3219323

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing constraints and spatial misallocation." *American Economic Journal: Macroeconomics* 11(2):1–39.

Irshad, Hira. 2017. "Relationship among political instability, stock market returns and stock market volatility." *Studies in business and economics* 12(2):70–99.

Jensen, Nathan M and Scott Schmith. 2005. "Market Responses to Politics The Rise of Lula and the Decline of the Brazilian Stock Market." *Comparative Political Studies* 38(10):1245–1270.

Johnson, Chalmers. 1982. *MITI and the Japanese miracle: the growth of industrial policy, 1925-1975*. Stanford University Press.

Jones, B.F. and B.A. Olken. 2009. "Hit or miss? The effect of assassinations on institutions and war." *American Economic Journal: Macroeconomics* pp. 55–87.

Jong-A-Pin, Richard. 2009. "On the measurement of political instability and its impact on economic growth." *European Journal of Political Economy* 25(1):15–29.

Judge-Lord, Devin. 2022. "The Environmental Justice Movement's Impact on Technocratic Policymaking." *Working paper* .

Kalla, Joshua L and David E Broockman. 2018. "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." *American Political Science Review* 112(1):148–166.

Kato, Sota. 2017. "Getting to the Root of Amakudari: Sweeping Reform Needed to Close the Revolving Door." *The Tokyo Foundation for Policy Research* .

Khwaja, Asim Ijaz and Atif Mian. 2005. "Do lenders favor politically connected firms? Rent provision in an emerging financial market." *The Quarterly Journal of Economics* 120(4):1371–1411.

Kim, In Song. 2017. "Political Cleavages within Industry: Firm-Level Lobbying for Trade Liberalization." *American Political Science Review* 111(01):1–20.

Kim, In Song. 2018. "Lobbyview: Firm-level lobbying & congressional bills database." *Unpublished manuscript, MIT, Cambridge, MA. http://web. mit. edu/insong/www/pdf/lobbyview. pdf Google Scholar Article Location* .

Kim, In Song and Iain Osgood. 2019. "Firms in Trade and Trade Politics." *Annual Review of Political Science* 22(1):399–417.

Kim, In Song, Steven Liao and Sayumi Miyano. 2021. Why Trade and FDI Should Be Studied Together. Technical report.

King, Gary and Langche Zeng. 2001. "Logistic regression in rare events data." *Political analysis* 9(2):137–163.

Klašnja, Marko and Joshua A Tucker. 2013. "The economy, corruption, and the vote: Evidence from experiments in Sweden and Moldova." *Electoral Studies* 32(3):536–543.

Klašnja, Marko, Noam Lupu and Joshua A Tucker. 2017. "When Do Voters Sanction Corrupt Politicians?" *Working paper* .

URL: http://noamlupu.com/corruption_sanction.pdf

Koepke, Robin. 2019. "What drives capital flows to emerging markets? A survey of the empirical literature." *Journal of Economic Surveys* 33(2):516–540.

Kolstad, Ivar and Arne Wiig. 2009. "Is transparency the key to reducing corruption in resource-rich countries?" *World development* 37(3):521–532.

Konstantinidis, Iannis and Georgios Xezonakis. 2013. "Sources of tolerance towards corrupted politicians in Greece: The role of trade offs and individual benefits." *Crime, Law and Social Change* 60(5):549–563.

Le, Quan Vu and Paul J Zak. 2006. "Political risk and capital flight." *Journal of International Money and Finance* 25(2):308–329.

Leblang, David and Bumba Mukherjee. 2005. "Government partisanship, elections, and the stock market: examining American and British stock returns, 1930–2000." *American Journal of Political Science* 49(4):780–802.

Lee, Kyuwon and Hye Young You. 2020. "Bureaucratic Revolving Doors and Interest Group Participation in Policymaking." *NYU Department of Political Science Working Paper* .

Lee, Terence. 2009. "The armed forces and transitions from authoritarian rule: Explaining the role of the military in 1986 Philippines and 1998 Indonesia." *Comparative Political Studies* 42(5):640–669.

Leeper, Thomas J, Sara B Hobolt and James Tilley. 2019. "Measuring subgroup preferences in conjoint experiments." *Political Analysis* .

Lehkonen, Heikki and Kari Heimonen. 2015. "Democracy, political risks and stock market performance." *Journal of International Money and Finance* 59:77–99.

Lens, Michael C and Paavo Monkkonen. 2016. "Do strict land use regulations make metropolitan areas more segregated by income?" *Journal of the American Planning Association* 82(1):6–21.

Lensink, Robert, Niels Hermes and Victor Murinde. 2000. "Capital flight and political risk." *Journal of international Money and Finance* 19(1):73–92.

Li, Zeren. 2021. "Chinese Revolving-Door Officials Database." *Unpublished manuscript, Yale University* .

Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *Annals of Applied Statistics* 7(1):295–318.

Liu, Li and Tao Zhang. 2015. "Economic policy uncertainty and stock market volatility." *Finance Research Letters* 15:99–105.

Londregan, J.B. and K.T. Poole. 1990. "Poverty, the coup trap, and the seizure of executive power." *World Politics* 42(02):151–183.

Loomis, John. 2011. "What's to know about hypothetical bias in stated preference valuation studies?" *Journal of Economic Surveys* 25(2):363–370.

Los Angeles Times. 2021. "Editorial: Some good from the pandemic era: Online access to government meetings.".

URL: *https://www.latimes.com/opinion/story/2021-04-28/require-livestreamed-public meetings*

MacKinlay, A.C. 1997. "Event studies in economics and finance." *Journal of economic literature* 35(1):13–39.

Malhotra, Neil, Melissa R Michelson, Ali Adam Valenzuela et al. 2012. "Emails from official sources can increase turnout." *Quarterly Journal of Political Science* 7(3):321–332.

Marble, William and Clayton Nall. 2021. "Where self-interest trumps ideology: liberal homeowners and local opposition to housing development." *The Journal of Politics* 83(4):1747–1763.

Mares, Isabela and Giancarlo Visconti. 2019. "Voting for the lesser evil: Evidence from a conjoint experiment in Romania." *Political Science Research and Methods* pp. 1–14.

Marinov, Nikolay and Hein Goemans. 2014. "Coups and democracy." *British Journal of Political Science* 44(4):799–825.

McCabe, Brian J. 2016. *No place like home: Wealth, community, and the politics of homeownership*. Oxford University Press.

McDermott, Mark. 2021. “BLACK LIVES MATTER: City Council spars over Bruce’s Beach plaque.”.

URL: *https://easyreadernews.com/black-lives-matter-city-council-spars-over-bruces-beach-plaque/*

McDermott, Rose and Peter K Hatemi. 2020. “Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms.” *Proceedings of the National Academy of Sciences* 117(48):30014–30021.

McDonald, Jared. 2019. “Avoiding the Hypothetical: Why “Mirror Experiments” are an Essential Part of Survey Research.” *International Journal of Public Opinion Research* .

Meyersson, Erik. 2016. “Political man on horseback: coups and development.” *Stockholm Institute for Transition Economics (SITE)* 5.

Milkman, Katherine L, John Beshears, James J Choi, David Laibson and Brigitte C Madrian. 2011. “Using implementation intentions prompts to enhance influenza vaccination rates.” *Proceedings of the National Academy of Sciences* 108(26):10415–10420.

Mishima, Ko. 2013. “A Missing Piece in Japan’s Political Reform: The Stalemate of Reform of the Bureaucratic Personnel System.” *Asian Survey* 53(4):703–727.

Mizoguchi, Tetsuro and Nguyen Van Quyen. 2012. “Amakudari: The Post-Retirement Employment of Elite Bureaucrats in Japan.” *Journal of Public Economic Theory* 14(5):813–847.

Molloy, Raven et al. 2020. “The effect of housing supply regulation on housing affordability: A review.” *Regional Science and Urban Economics* 80(C).

Monkkonen, Paavo and Michael Manville. 2019. “Opposition to development or opposition to developers? Experimental evidence on attitudes toward new housing.” *Journal of Urban Affairs* 41(8):1123–1141.

Muñoz, Jordi, Eva Anduiza and Aina Gallego. 2012. Why do voters forgive corrupt politicians? Cynicism, noise and implicit exchange. In *International Political Science*

Association Conference, Madrid, Spain.

URL: https://www.researchgate.net/profile/Eva_Anduiza/publication/268056601_Why_do_voters_forgive_corrupt_politicians_Cynicism_noise_and_implicit_exchange/links/54677eb30cf2f5eb18036b4d.pdf

National Public Service Act. 1947.

URL: http://www.japaneselawtranslation.go.jp/law/detail_main?re=&vm=02&id=2713

Needler, M.C. 1966. "Political development and military intervention in Latin America." *The American Political Science Review* 60(3):616–626.

Nickerson, David W. 2008. "Is voting contagious? Evidence from two field experiments." *American political Science review* 102(1):49–57.

Nickerson, David W and Todd Rogers. 2010. "Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making." *Psychological Science* 21(2):194–199.

Nickerson, David W et al. 2007. "Does email boost turnout." *Quarterly Journal of Political Science* 2(4):369–379.

North, Douglass C and Barry R Weingast. 1989. "Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England." *The journal of economic history* 49(4):803–832.

Obstfeld, Maurice. 2012. "Financial flows, financial crises, and global imbalances." *Journal of International Money and Finance* 31(3):469–480.

Olson, Mancur. 1965. The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard University Press.

Ortalo-Magné, François and Andrea Prat. 2014. "On the political economy of urban growth: Homeownership versus affordability." *American Economic Journal: Microeconomics* 6(1):154–81.

Ostrom, Elinor. 2000. "Collective action and the evolution of social norms." *Journal of*

Economic Perspectives 14(3):137–158.

Overholt, William H. 1986. "The rise and fall of Ferdinand Marcos." *Asian Survey* 26(11):1137–1163.

Pantzalis, Christos, David A Stangeland and Harry J Turtle. 2000. "Political elections and the resolution of uncertainty: the international evidence." *Journal of banking & finance* 24(10):1575–1604.

Peters, Jaime L, Alex J Sutton, David R Jones, Keith R Abrams and Lesley Rushton. 2007. "Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity." *Statistics in Medicine* 26(25):4544–4562.

Peters, Margaret. 2017. *Trading Barriers: Immigration and the Remaking of Globalization.* Princeton University Press.

Philp, Mark and Elizabeth David-Barrett. 2015. "Realism about political corruption." *Annual Review of Political Science* 18:387–402.

Powell, Jonathan M and Clayton L Thyne. 2011. "Global instances of coups from 1950 to 2010: A new dataset." *Journal of Peace Research* 48(2):249–259.

Przeworski, Adam and Michael Wallerstein. 1982. "The structure of class conflict in democratic capitalist societies." *American Political Science Review* 76(2):215–238.

Quigley, John M and Larry A Rosenthal. 2005. "The effects of land use regulation on the price of housing: What do we know? What can we learn?" *Cityscape* pp. 69–137.

Reed, W Robert, Raymond JGM Florax and Jacques Poot. 2015. "A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias." *Economics Discussion Papers* .

Reeves, Richard V. 2018. *Dream hoarders: How the American upper middle class is leaving everyone else in the dust, why that is a problem, and what to do about it.* Brookings Institution Press.

Riker, William H and Peter C Ordeshook. 1973. *An introduction to positive political theory.* Vol. 387 Prentice-Hall Englewood Cliffs, NJ.

Roe, M.J. and J.I. Siegel. 2011. "Political instability: Effects on financial development, roots in the severity of economic inequality." *Journal of Comparative Economics* .

Rogers, Todd and David Nickerson. 2013. "Can Inaccurate Beliefs About Incumbents be Changed? And Can Reframing Change Votes?" *HKS Working Paper No. RWP13-018* .

URL: https://research.hks.harvard.edu/publications/getFile.aspx?Id=941

Rose-Ackerman, Susan. 2018. "Corruption & purity." *Daedalus* 147(3):98–110.

Rose-Ackerman, Susan and Bonnie J Palifka. 2016. *Corruption and government: Causes, consequences, and reform*. Cambridge: Cambridge university press.

Rosenbluth, Frances M. 1989. *Financial politics in contemporary Japan*. Cornell University Press.

Rosenbluth, Frances McCall and Michael F Thies. 2010. *Japan transformed*. Princeton University Press.

Rundquist, Barry S, Gerald S Strom and John G Peters. 1977. "Corrupt politicians and their electoral support: some experimental observations." *American Political Science Review* 71(3):954–963.

Schaede, Ulrike. 1995. "The" Old Boy" network and government-business relationships in Japan." *Journal of Japanese Studies* 21(2):293–317.

Schubert, Christian. 2017. "Green nudges: Do they work? Are they ethical?" *Ecological Economics* 132:329–342.

Sears, David O and Carolyn L Funk. 1991. "The role of self-interest in social and political attitudes." *Advances in experimental social psychology* 24:1–91.

Sears, David O and Jack Citrin. 1982. *Tax revolt: Something for nothing in California*. Harvard University Press.

Simonsohn, U, JP Simmons and LD Nelson. 2015. "Better P-curves: Making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a Reply to Ulrich and Miller

(2015).” *Journal of Experimental Psychology: General* 144(6):1146.

Simonsohn, Uri, Leif D Nelson and Joseph P Simmons. 2014*a*. “P-curve: a key to the file-drawer.” *Journal of Experimental Psychology: General* 143(2):534.

Simonsohn, Uri, Leif D Nelson and Joseph P Simmons. 2014*b*. “p-curve and effect size: Correcting for publication bias using only significant results.” *Perspectives on Psychological Science* 9(6):666–681.

Slough, Tara. 2019. “The Ethics of Electoral Experimentation: Design-Based Recommendations.” *Working Paper* .

URL: *http://taraslough.com/assets/pdf/eee.pdf*

Slough, Tara and Scott A Tyson. 2021. “External Validity and Meta-Analysis.” *Working Paper* .

Solaz, Hector, Catherine E De Vries and Roosmarijn A de Geus. 2019. “In-group loyalty and the punishment of corruption.” *Comparative Political Studies* 52(6):896–926.

Stanley, Tom D and Hristos Doucouliagos. 2014. “Meta-regression approximations to reduce publication selection bias.” *Research Synthesis Methods* 5(1):60–78.

Stanley, Tom D and Hristos Doucouliagos. 2017. “Neither fixed nor random: Weighted least squares meta-regression.” *Research Synthesis Methods* 8(1):19–42.

Sterne, Jonathan AC, Matthias Egger and George Davey Smith. 2001. “Investigating and dealing with publication and other biases in meta-analysis.” *BMJ* 323(7304):101–105.

Stokes, Susan C, Thad Dunning, Marcelo Nazareno and Valeria Brusco. 2013. *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge: Cambridge University Press.

Sulitzeanu-Kenan, Raanan, Yoav Dotan and Omer Yair. Forthcoming. “Can Institutions Make Voters Care about Corruption?” *The Journal of Politics* .

Svensson, J. 1998. “Investment, property rights and political instability: Theory and evidence.” *European Economic Review* 42(7):1317–1341.

Teele, Dawn Langan, Joshua Kalla and Frances Rosenbluth. 2018. “The Ties That Double

Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112(3):525–541.

Terada, Mayu. 2019. "The Changing Nature of Bureaucracy and Governing Structure in Japan." *Pacific Rim Law & Policy Journal* 28:431.

Terrin, Norma, Christopher H Schmid, Joseph Lau and Ingram Olkin. 2003. "Adjusting for publication bias in the presence of heterogeneity." *Statistics in Medicine* 22(13):2113–2126.

The Boston Globe. 2021. "The pandemic taught us a better way to do public business.".
URL: *https://www.bostonglobe.com/2021/05/27/opinion/pandemic-taught-us-better-way-do-public-business/*

Thyne, Clayton L and Jonathan M Powell. 2016. "Coup d?état or Coup d'Autocracy? How Coups Impact Democratization, 1950–2008." *Foreign Policy Analysis* 12(2):192–213.

Traywick, Catherine. 2014. "Shoes, Jewels, and Monets: The Immense Ill-Gotten Wealth of Imelda Marcos." *Foreign Policy* .

Truex, Rory. 2014. "The returns to office in a 'rubber stamp' parliament." *American Political Science Review* 108(2):235–251.

United States Census Bureau. 2020. "New Privately-Owned Housing Units Authorized by Building Permits in Permit-Issuing Places.".
URL: *https://www.census.gov/construction/bps/pdf/annualhistorybystate.pdf*

Valentine, Jeffrey C, Therese D Pigott and Hannah R Rothstein. 2010. "How many studies do you need? A primer on statistical power for meta-analysis." *Journal of Educational and Behavioral Statistics* 35(2):215–247.

van Aert, Robbie CM, Jelte M Wicherts and Marcel ALM van Assen. 2016. "Conducting meta-analyses based on p values: Reservations and recommendations for applying p-uniform and p-curve." *Perspectives on Psychological Science* 11(5):713–729.

Van Aert, Robbie CM, Jelte M Wicherts and Marcel ALM Van Assen. 2019. "Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis."

PloS one 14(4):e0215052.

Varol, Ozan O. 2011. "The democratic coup d?état." *Harvard Journal of International Law* .

Venieris, Yiannis P and Dipak K Gupta. 1986. "Income distribution and sociopolitical instability as determinants of savings: a cross-sectional model." *Journal of Political Economy* 94(4):873–883.

Vera, Sofia B. 2019. "Accepting or Resisting? Citizen Responses to Corruption Across Varying Levels of Competence and Corruption Prevalence." *Political Studies* .

Villegas, Bernardo M. 1987. "The Philippines in 1986: Democratic Reconstruction in the Post-Marcos Era." *Asian Survey* 27(2):194–205.

Vogel, Steven Kent. 2006. *Japan remodeled: How government and industry are reforming Japanese capitalism*. Cornell University Press.

Weitz-Shapiro, Rebecca and Matthew S Winters. 2017. "Can citizens discern? Information credibility, political sophistication, and the punishment of corruption in Brazil." *The Journal of Politics* 79(1):60–74.

Weschle, Simon. 2016. "Punishing personal and electoral corruption: Experimental evidence from India." *Research & Politics* 3(2):1–6.

Wheeler, Stephen M, Christopher M Jones and Daniel M Kammen. 2018. "Carbon footprint planning: quantifying local and state mitigation opportunities for 700 California cities." *Urban Planning* 3(2):35–51.

Winters, Matthew S and Rebecca Weitz-Shapiro. 2013. "Lacking information or condoning corruption: When do voters support corrupt politicians?" *Comparative Politics* 45(4):418–436.

Winters, Matthew S and Rebecca Weitz-Shapiro. 2015. "Political corruption and partisan engagement: evidence from Brazil." *Journal of Politics in Latin America* 7(1):45–81.

Winters, Matthew S and Rebecca Weitz-Shapiro. 2016. "Who's in charge here? Direct and indirect accusations and voter punishment of corruption." *Political Research Quarterly*

69(2):207–219.

Winters, Matthew S and Rebecca Weitz-Shapiro. 2020. “Information credibility and responses to corruption: a replication and extension in Argentina.” *Political Science Research and Methods* 8(1):169–177.

Yoder, Jesse. 2020. “Does Property Ownership Lead to Participation in Local Politics? Evidence from Property Records and Meeting Minutes.” *American Political Science Review* 114(4):1213–1229.