

Expansions of Executive Authority:
Government Leaders' Near-Term Pressures and Long-Term Fates

By

Natalie Ahn

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Public Policy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Sarah Anzia, Chair

Assistant Professor Jennifer Bussell

Assistant Professor Aila Matanock

Assistant Professor David Bamman

Spring 2018

Abstract

Expansions of Executive Authority:
Government Leaders' Near-Term Pressures and Long-Term Fates

by

Natalie Ahn

Doctor of Philosophy in Public Policy

University of California, Berkeley

Associate Professor Sarah Anzia, Chair

In many developing countries, even those that now have democratically elected governments, there remain widespread concerns about strong leaders exercising unchecked power and imposing important decisions through executive orders or decrees. Concentrated executive authority seems most popular during times of crisis or political gridlock, but can backfire in the long run. Many developing democracies have also had presidents removed early and/or prosecuted for abuses of power. Yet it remains unclear whether these forms of accountability actually work, when prosecutions are often politicized and may drive incumbents to cling to power to protect themselves.

In this dissertation, I study how the fates of former Latin American presidents influence successors' efforts to consolidate power, through their use of executive decrees. I use computational methods to parse decrees and identify those that not only enact external policies, but that make changes to internal government institutions. I also introduce measures of the consistency or predictability of former leaders' punishments and rewards, and test whether that predictability influences subsequent leaders' decisions about how to use their own power while in office. This dissertation contributes to our understanding of how strong political leaders make important decisions – even restructuring powerful components of the governing apparatus – and what signals about their future punishment or reward encourage them to seek more power for themselves or to exercise restraint.

Table of Contents

Table of Contents	i
List of Figures	iv
List of Tables	v
Acknowledgements	vi
Chapter 1. Introduction	1
1.1 Motivating examples	2
1.1.1 Peru’s Alberto Fujimori and the 1992 <i>autogolpe</i>	2
1.1.2 Ecuador’s weak presidents and the rise of Rafael Correa	5
1.2 The puzzle: power consolidation in developing democracies	6
1.3 The argument: objectivity and predictability of observed fates	8
1.4 Theoretical challenges	9
1.5 Methodological challenges	10
1.6 Case selection	12
1.6.1 Introducing the remaining cases: Bolivia	13
1.6.2 Introducing the remaining cases: Colombia	14
1.6.3 Introducing the remaining cases: Venezuela	16
1.6.4 Summary of country-years covered	18
1.7 Overview of remaining chapters	20
Chapter 2. A Theory of the Motivations for and Costs of Consolidating Power	21
2.1 Existing literature on presidential power and executive orders	21
2.2 Existing literature on presidential removal and post-tenure fates	23
2.3 Why do leaders seek greater unilateral decision-making powers?	25
2.4 How to tell if leaders are trying to get things done or seeking power alone?	28
2.5 How to verify that these different actions align with different motivations?	30
2.6 How do leaders view the risks of future punishment or reward?	33
2.6.1 Types of sanctions, reprieves, and rewards	34
2.6.2 Types of actors initiating a sanction or reward	36
2.6.3 Different types of sanctions, stages in the legal process	39
2.6.4 Consistency or predictability of former leaders’ fates	40
2.6.5 Interactions between near-term pressures and post-tenure fates	41
2.7 Summary of theoretical discussion	42
Chapter 3. Methodology for Measuring Executive Action through Decrees	44
3.1 Overview of methodological decisions	44
3.2 Sources of information on government action	44

3.2.1 Shortcomings of news reporting and advantages of public records	45
3.2.2 Challenges to analyzing government documents	46
3.2.3 Documents chosen for this project	47
3.3 Established methods for extracting actions and entities from text	49
3.4 Project strategy for document classification	51
Step 1. Project coding scheme	53
Step 2. Text preprocessing and feature extraction	54
Step 3. Classification algorithms	57
Step 3.1: Supervised machine learning classifiers	58
Step 3.2: Rule-based information extraction	59
3.5 Evaluation of accuracy in relation to human coding	60
3.5.1 Results of accuracy tests, comparing models	61
3.5.2 Results for ensemble classification	63
3.5.3 Caveats to evaluating data measures by accuracy alone	66
 Chapter 4. Description of Decree Usage and Contextual Validation of Decree Categories	 67
4.1 Summary of decrees collected by country	67
4.2 Using established correlates to distinguish motivations for decrees	71
4.2.1 What actions do leaders take when facing legislative opposition?	73
4.2.2 What actions do leaders take during disasters and economic shocks?	76
4.2.3 What actions do leaders take as they near the end of their term?	79
4.3 Finalizing the decree categories to use in main hypothesis tests	81
 Chapter 5. The Relationship Between Power Consolidation and Leader's Fates	 83
5.1 Data sources for leaders' fates	84
5.2 Coding scheme for post-tenure events	86
5.3 Structure of the post-tenure event data	87
5.4 Descriptive statistics for presidents' fates in dataset	88
5.5 Statistical models and parameters for hypothesis tests	91
5.6 Results for hypothesis tests, baseline post-tenure event categories	95
5.7 Results for hypothesis tests, post-tenure events by instigating actor	99
5.8 Summary of analysis for general categories of post-tenure events	105
 Chapter 6. The Role of Context, Consensus, and Predictability in Post-Tenure Sanctions	 106
6.1 Data on more specific sanction events within the legal process	106
6.2 Results for post-tenure sanctions at different stages of legal process	107
6.3 Measuring the probability or predictability of post-tenure fates	111
6.4 Results for post-tenure sanctions by predicted probability	116
6.5 Robustness checks using alternative predictions of sanction probability	120
6.6 Relationship between near-term pressures and post-tenure fates	124

6.7 Results for interactions between sanctions and economic shocks	125
6.8 Summary of additional analyses of post-tenure sanctions	129
Chapter 7. Conclusion	131
7.1 Theoretical contributions and policy implications	131
7.2 Methodological contributions and broader applications	134
References	138

List of Figures

Figure 1. Coups d'état and democracies vs autocracies over time	7
Figure 2. Literary references to military coups and political conflict over time	8
Figure 3. Parse tree for a typical decree title	55
Figure 4. Bag-of-words vs. structured features for example text	56
Figure 5. Example rules to match verb-object clauses to action or target labels	59
Figure 6. Accuracy of supervised learning by size of training corpus	62
Figure 7. Executive decrees issued by month for each country in the dataset	68
Figure 8. Decrees in dataset by country, broken down by main action category	69
Figure 9. Decrees in dataset by country, broken down by target entity category	70
Figure 10. Decrees in dataset, by combination of main action and target entity	71
Figure 11. Relationship between total decrees issued and legislative control	73
Figure 12. Bivariate regression of decree categories on governing party control of congress	74
Figure 13. Bivariate regression of decree categories on party fractionalization	75
Figure 14. Bivariate regression of decree categories on earthquake disasters	77
Figure 15. Bivariate regression of total decrees on export commodity price shocks	78
Figure 16. Bivariate regression of decree categories on export commodity price shocks	79
Figure 17. Bivariate regression of decree categories on months left in term	80
Figure 18. Former leaders' post-tenure events, by country and by instigating actor	90
Figure 19. Summary of estimates, separate models for sanctions, reprieves, and rewards	98
Figure 20. Summary of estimates, combined model for sanctions, reprieves, and rewards	98
Figure 21. Summary of regression models for sanctions by actor category	100
Figure 22. Summary of regression models for reprieves by actor category	101
Figure 23. Summary of regression models for rewards by actor category <i>(Recording reward events only in the month in which they began)</i>	103
Figure 24. Summary of regression models for rewards by actor category <i>(Recording reward events in all months in which the reward was in effect)</i>	104
Figure 25. Summary of regression models for sanction events at stages of legal process <i>(Recording events only in the month in which they began)</i>	109
Figure 26. Summary of regression models for sanction events at stages of legal process <i>(Recording events in all months in which the sanction was in effect)</i>	110
Figure 27. Summary of estimates for sanctions interacted with predicted probability	117
Figure 28. Summary of estimates for sanctions by predicted probability quantile	118
Figure 29. Sanctions interacted with predicted probability, predictions stratified by leader	121
Figure 30. Sanctions by predicted probability quantile, predictions stratified by leader	121
Figure 31. Sanctions interacted with probability, predicted from prior fates chronologically	122
Figure 32. Sanctions by probability quantile, predicted from prior fates chronologically	123
Figure 33. Bivariate regression of presidential approval on export price shocks	125
Figure 34. Summary of estimates for export price drop terms in interaction models	126
Figure 35. Summary of estimates for sanction terms in interaction models	127
Figure 36. Summary of estimates for sanction-price shock interaction terms	128

List of Tables

Table 1. Government leaders' motivations and strategies for use of power in office	27
Table 2. Decree actions related to leaders' different motivations and strategies	29
Table 3. Established covariates that might reveal leaders' motivations and strategies	32
Table 4. Hypothesized relationships for post-tenure fates and successors' actions	36
Table 5. Hypotheses for post-tenure fates by instigating actor	38
Table 6. Hypotheses for specific types or stages of sanctions	39
Table 7. Hypotheses for consistency or predictability of sanctions	41
Table 8. Hypothesized interactions between crises and post-tenure fates	42
Table 9. Examples of decree titles in dataset	49
Table 10. Summary of Methodological Options Selected for Comparison	52
Table 11. Coding Scheme with main action and target entity categories	54
Table 12. Accuracy against human coding, comparing classification algorithms	61
Table 13. Accuracy against human coding, ensemble classification	64
Table 14. Evaluation scores for ensemble classification by high-level category	65
Table 15. Total decrees and years covered for each country in the project dataset	67
Table 16. Summary of post-tenure fate events, actors involved, and additional factors	86
Table 17. Former leaders' post-tenure events in the dataset, observation event counts	89
Table 18. Control variables included in main regression models, by random selection	93
Table 19: Regression estimates for main hypothesis tests	96
Table 20: Average regression estimates for decrees enabling government executive offices	107
Table 21. Predicted probabilities of actual sanctions in dataset, by actor or stage	114
Table 22. Example values for new sanction variables incorporating predicted probability	115

Acknowledgements

This dissertation benefited from the help and support of many people and institutions, to whom I am deeply grateful. First and foremost, my advisor and committee chair, Sarah Anzia, has been unfailingly supportive, encouraging, pragmatic, perceptive, and wise throughout my pursuit of this degree. She has taught me more than anyone how to be an effective scholar and social scientist, seeking to produce meaningful research on societal problems in ways that are not only intellectually rigorous but are applicable to the development of practical, real-world solutions. She has been my rock throughout this process in so many ways; it would not have been possible without her continuous guidance and support.

I also owe many thanks to the rest of my committee. Jennifer Bussell has provided extensive advice and mentorship throughout this project, from the theoretical development, through the practical implementation of the research, to the analysis and presentation of meaningful results. Aila Matanock also provided thoughtful guidance on the theoretical foundations of the project, the process of conducting field research, and the formulation of the final analysis. David Bamman provided invaluable advice and mentorship on the computational methods that play a central role in this dissertation, especially the use of natural language processing and machine learning to extract information from text records, as well as in the development of my skills as a data scientist in general. And finally, Leonardo Arriola, who served on my qualifying exam committee and as an advisor since the early stages of the process, provided very helpful guidance especially during the conceptual development and practical planning for the project, and was a valuable source of mentorship and support throughout.

I am also deeply grateful to fellow graduate students at both the Goldman School of Public Policy (GSPP) and the Political Science Department at the University of California, Berkeley, who formed the community of colleagues on which I depended throughout this process. Special thanks to Gabrielle Elul, Fiona Shen-Bayh, Rachel Bernhard, Paul Thissen, and Amanda Charbonneau for many hours of formative discussions and feedback, and for their steadfast friendship and moral support. Thanks also to other members of the GSPP faculty, especially to Amy Lerman for her excellent mentorship and advice through the GSPP Ph.D. seminar and thoughtful comments on research drafts, to Michael Nacht for his warm and welcoming mentorship and support during my first years at GSPP, and to Geno Smolensky and John Ellwood for their early stewardship of the Ph.D. program, thoughtful advice and support. I am also deeply grateful to Terry Moe, for his longtime encouragement and support of my academic and intellectual pursuits, his advice that brought me to GSPP and connected me to Sarah Anzia, and for continued conversations and invaluable advice on this dissertation.

Many thanks to Atha Davis, Arman Jaffer, Marisa Lenci, and Erika Stablow for their excellent research assistance in support of this project. For supporting my field research in South America, I am grateful to the Department of Social Sciences at the Pontificia Universidad Católica del Perú (PUCP), and the Department of Political Studies at the Facultad Latinoamericana de Ciencias Sociales (FLACSO) in Ecuador. This project benefited from

thoughtful insights provided by scholars and affiliates at each of those institutions, as well as several local journalists and practitioners, many of who's work is cited in the following pages. The project was also made possible by generous financial support from the University of California, Berkeley's Graduate Division, through the Dissertation-Year Fellowship, Doctoral Completion Fellowship, and summer research grants, as well as by early support from the Nuclear Science and Security Consortium and the Institute for Humane Studies.

Finally, I am forever grateful to my family and friends – especially my sisters Mary and Rebecca, and parents Sue and John – for their constant love and support.

Chapter 1. Introduction

In many countries today, from the Philippines to Turkey to even the United States, there is considerable debate about the power of government leaders, especially those who seem to govern with unchecked personal authority, making decisions on their own that arbitrarily benefit their friends and oppress their foes. Concentrated government power is not new, but it seems to persist even as states democratize and modernize, and as populations become more mobilized and politically engaged. Sometimes heavy-handed government leaders are thrown out of office in disgrace, or face prosecution, conviction, and/or imprisonment for abuses of power. Yet their successors often continue to follow in their footsteps, consolidating their own sources of power and circumventing checks and balances until opponents manage to remove them as well.

Controversy and backlash surround powerful leaders not only in extreme cases of violent oppression, but also in debates about executive orders or decrees, issued unilaterally by presidents (or similar government heads) to make major policy decisions and institutional changes without the involvement of the legislature or other parties. Former presidents in Latin America and other recently democratized regions are increasingly being prosecuted, not only for severe human rights violations, but also for criminal charges like embezzlement, illegal search and seizure, and other forms of corruption or infringement of constitutional rights (Pérez-Liñán 2007; Burt 2009; Conaghan 2012). These violations are often attributed to unilateral executive action, either to decisions directly enacted by decree, or undertaken by agencies that take orders from the president's office, which might themselves have been created by decree.

Do all government leaders want as much power as they can get? If former leaders are more likely to face backlash and punishment when they amass and wield more power at others' expense, there seems to be good reason for forward-thinking leaders to moderate their power ambitions. Not all modern government leaders do concentrate executive authority. In fact, a more prominent body of political science research seeks to explain why many leaders in historically autocratic or politically developing states choose to *share* power and accept constraints. Those theories assert that leaders cannot maintain power and exercise it effectively in modern societies without winning over a large number of supporters, who not only seek material benefits for their support, but also often demand constraints on the leader and a credible institutionalized share of government power for themselves (Acemoglu and Robinson 2001; Bueno de Mesquita et al 2003; Gandhi and Przeworski 2007; Magaloni and Kricheli 2010; Boix and Svobik 2013).

Why, then, do many government leaders in modern states move back in the opposite direction, seeking to concentrate power and control decision making on their own, without input or consent from other branches or parties in government? Even when there are negative consequences, why does the prosecution of former leaders for abusing power often fail to deter successors from doing the same thing?

1.1 Motivating examples

There have been many controversial “strongman” leaders around the world in recent decades, from Egypt’s Hosni Mubarak and Angola’s José Eduardo de Santos beginning in the 1980s, to Venezuela’s Hugo Chavez and Pakistan’s Pervez Musharraf in the early 2000s, to contemporary leaders like Turkey’s Recep Tayyip Erdogan and the Philippines’ Rodrigo Duterte. Some have managed to remain in office until they died, but others were eventually removed and prosecuted. The popular uprisings of the 2003–2005 “Color Revolutions” in former Soviet states and the 2011 “Arab Spring” in the Middle East removed multiple long-standing, heavy-handed rulers, giving way to new constitutions and presidential elections (Tudoroiu 2007; Bellin 2012; Tabaar 2013). Yet the removal and sanction of these leaders has not marked the end of concentrated power; many of the same countries have since empowered new personalist leaders to begin the cycle over again (Tudoroiu 2007; Tabaar 2013; Teehankee 2016).

1.1.1 Peru’s Alberto Fujimori and the 1992 *autogolpe*

An archetypal case of a president who amassed power and governed unilaterally for a time, but was ultimately removed and convicted of multiple abuses of power, is Peru’s Alberto Fujimori. Fujimori rose to power as a political outsider at a time of turmoil and public disenchantment with the established political elite, winning a surprise victory in the 1990 presidential election (Gonzales de Olarte 1993; Vargas Llosa 1995; García 2001). He quickly began to implement a series of sweeping economic and security reforms, many of them through unilateral decrees, against increasing resistance from the legislature. Within two years of taking office, at loggerheads with the opposition, he disbanded Congress, purged the judiciary, and called elections for a Constituent Assembly to draft a new Constitution that would allow his reelection to two more terms (García 2001; Conaghan 2005). The power grab was widely referred to as an *autogolpe*, or “self-coup” (Lane 1992).

After a decade in office, however, his popularity waned and his final reelection in 2000 was marred by constitutional ambiguity, allegations of fraud and a scandal involving his intelligence chief. He served a few more embattled months before traveling to Japan and sending back a letter of resignation, which the legislature rejected in favor of dismissing him themselves (García 2001; Pérez-Liñán 2007). He remained in exile until he was detained in Chile in 2005 and extradited back to Peru in 2007, where he faced multiple trials and convictions for embezzlement, illegal search and wiretapping, and ordering the kidnapping and killing of civilians (Franklin 2007; Partlow and Chauvin 2009; Cespedes and Wade 2009).

There have been many accounts by historians, journalists, and academics about Fujimori’s rise and fall (Bowen 2000; García 2001; Conaghan 2005; Murakami 2007). Yet even for such a prominent, widely documented case, there is little consensus about why he chose to consolidate a form of personal rule within what was otherwise a democratic state. Local experts’

views and daily local news coverage suggest that there was ample evidence to support multiple prominent theories about why a president might choose to consolidate power – from simultaneous economic and security crises, to severe legislative opposition and gridlock, to a populist outsider’s political naiveté and the influence of a power-hungry political manipulator.

First, Fujimori may have resorted to decrees because he was simply trying to get urgently needed policies enacted, in the face of a debt crisis, hyperinflation and a violent insurgency that he inherited when he entered office (Nash 1991; Roberts 1996; Stokes 1997; Tanaka 2006). The government was extremely low on reserves and desperate to appeal to foreign creditors with stabilization measures (Nash 1991; Gonzales de Olarte 1993). Meanwhile, Fujimori’s party controlled few seats in a highly fragmented legislature, and Peru’s established parties had been disgraced, fueling the rise of many political outsiders and atomized candidate-centered movements like Fujimori’s (Stokes 1997; Bowen 2000; Levitsky and Cameron 2003; Tanaka 2006). Also evidence of the weakness of public institutions at the time, news reports of corruption scandals hit all three government branches from 1991-92, including misuse of funds in the legislature and the dismissal of judges for ties to illegal activity.¹ Congress’s inability to quickly and effectively enact urgently needed legislation, along with low public confidence in other centers of power, probably contributed to Fujimori’s decisions to go it alone.

However, the legislature did not prove incapable of taking *any* action, and became much more obstructionist over time, at least in part in reaction to Fujimori’s actions. At first, there was some solidarity around the need to rapidly address the economic crisis, and Congress initially delegated additional decree authority to Fujimori to do so (Gonzales de Olarte 1993). Yet support eroded as Fujimori pursued unanticipated and severe structural reforms – dubbed “Fujishock” (Nash 1991; Gonzales de Olarte 1993; Stokes 1997), along with heavy-handed military and paramilitary operations (Partlow and Chauvin 2009; Cespedes and Wade 2009). As Fujimori’s expansive decree-making outstripped the authority the legislature was willing to concede, they began to take action to stop him, calling a special legislative session to review his decrees and ultimately revising or repealing many of them (McClintock 1999). The campaign against his decrees culminated in the passage of a Law of Parliamentary Control, designed to curtail the president’s decree-making powers, which Fujimori vetoed (Schmidt 1998).

Some of the factors behind this confrontation relate to Fujimori’s political outsider status, but go beyond the weakness of the party system. Relevant factors also include the types of opportunists who attach themselves to a political neophyte, and the prospects for the incumbent’s and his inner circle’s interests in and outside the state. A common feature of many explanations for Fujimori’s leadership style was the strong influence of a personal advisor – Vladimiro Montesinos (McClintock and Vallas 2003; Levitsky and Cameron 2003; Burt 2009). Montesinos served as an army captain in the 1970s but was discharged and served time in prison for illegal acts (Caistor 2000; McClintock and Vallas 2003). In the 1980s, he worked as a lawyer for drug traffickers, until he regained ties to state intelligence and became connected to Fujimori’s 1990

¹ La República, “Suprema sacará a jueces corruptos,” December 9, 1991. El Comercio, “Olivera anuncia acusará

campaign (McClintock and Vallas 2003; Conaghan 2005). He may have perceived Fujimori as his own best chance to rise to power, previous attempts having failed.

Over the course of Fujimori's first few years in office, Montesinos encouraged and orchestrated the restructuring of executive offices and agencies, and collected materials to blackmail many high-ranking officials, until he held de facto personal control over the entire national security apparatus and much of the political sphere (Rotella 1996; McClintock and Vallas 2003; Burt 2009; Gorriti 2015). He maintained close ties to foreign agencies, especially U.S. intelligence, which he is widely believed to have used to enrich himself and to secure his position against mounting criticism and fear within Peru (Gorriti 1994; Conaghan 2005; Berzon et al 2005). As his machinations and abuses grew, so did the risks of losing power, which might have propelled Montesinos to advocate even more extreme measures to secure Fujimori's rule. After Fujimori's downfall, Montesinos was captured and convicted of corruption and human rights violations, for which he is currently imprisoned (Adameczyk 2016).

Finally, what is not mentioned in most scholarly accounts, but was apparent in frequent front-page news stories during the months leading up to Fujimori's power grab, was that his immediate predecessor was facing corruption charges in a heated political battle at the time (Cam 1991; Quiroz 2008). Former president Alan García first faced investigation in the legislature, with both houses voting to strip him of legal immunity and launch criminal proceedings.² Prosecutors then went back and forth with the courts,³ where judges appointed by García initially dismissed the charges, but opponents continued to push for trial (Quiroz 2008). Several members of García's cabinet also fled arrest warrants for corruption charges in the final months of 1991.⁴

Fujimori had accepted García's last-minute support against García's better known rivals during the 1990 election, and remained in tacit defense of that alliance early in his tenure (Vargas Llosa 1995; Levitsky and Cameron 2003). It seems plausible that García's fate influenced Fujimori's view of the threats he might face himself, were he to become too weak or lose power. There were already scandals unfolding about government abuses occurring on Fujimori's watch, including investigations into crimes by military and police units, and the legislative questioning and censure of several cabinet members.⁵ Then just before the *autogolpe*, Fujimori's wife (in her first move separating herself from her husband's administration) denounced his relatives for misappropriating charitable donations, initiating an investigation (Conaghan 2005). These events might also have influenced Fujimori's fear for his own fate, if he allowed rival authorities the institutional capacity to challenge him.

² La República, "Diputados aprobó acusación contra Alan por abrumadora mayoría," October 3, 1991. La República, "Caso Alan García llegó a Cámara Alta: Mañana nombran 7 senadores que dictaminarán acusación," October 8, 1991. La República, "Comisión acusa a Alan y pide que le suspendan inmunidad parlamentaria," October 15, 1991. La República, "Por 38 votos contra 17 Senado aprobó la acusación contra Alan," October 19, 1991

³ La República, "Fiscal formalizó denuncia penal contra Alan García," November 23, 1991. La República, "Gobierno nombra nuevo procurador público en el proceso contra Alan," December 12, 1991

⁴ La República, "Policía Judicial no halla rastros de 3 ex ministros," October 2, 1991.

⁵ La República, "Exigen renuncia inmediata de canciller y ministro de Defensa," October 31, 1991. La República, "Titulares del Interior y Defensa deberán explicar repudiable crimen: Senado llama a dos ministros por la horrorosa matanza de Barrios Altos," November 5, 1991. La República, "Partidos cuestionan gestión de los dos ministros: Diputados deciden esta semana censura a Carlos Boloña y Rossl Link," November 25, 1991.

In sum, the most striking moment of power consolidation in Fujimori's tenure occurred at a time when the legislature had become a determined obstacle to the enactment of his agenda, while economic and security crises threatened to overwhelm his administration. It also occurred at a time when the president's immediate predecessor – García – was being investigated for criminal acts and facing a controversial political battle over whether or not the case would reach the courts. Meanwhile, the novice outsider president – Fujimori – was listening to advice from a political manipulator – Montesinos – who had himself faced jailtime when he lost the protections of an earlier official position. Accusations and investigations targeting officials within Fujimori's administration were also unfolding, and a scandal had finally reached his own family, suggesting that he might have begun to fear for his own fate if he did not establish direct unilateral control.

Each of these factors may have contributed to the president's actions in various ways, but disentangling their effects requires more than a single case. Not all recent presidents increased their power in the ways Fujimori did, even when they faced some of the same challenges.

1.1.2 Ecuador's weak presidents and the rise of Rafael Correa

During the 1990s, neighboring Ecuador experienced its own economic and security crises, including a border conflict and a financial crisis toward the end of the decade (Martinez 2006). Established political parties were under fire from a disenchanted population, and voters elected a bombastic populist to the presidency – Abdalá Bucaram – in 1996 (Goering 1997). From the 1990s to early 2000s, three Ecuadorian presidents (Bucaram, Jamil Mahuad, and Lucio Gutiérrez) were removed early, and they and others were eventually prosecuted for misuses of power (Conaghan 2008; Basabe-Serrano et al 2010). Yet the presidency in Ecuador remained relatively weak compared to that in Peru through the 1980s and 1990s, unable to diminish or circumvent the pressures and constraints imposed by other government branches and political actors (Conaghan 2008; Basabe-Serrano et al 2010; Acosta and Polga-Hecimovich 2011).

It remains unclear whether Ecuadorian presidents tried to expand their authority, but were removed before they had gotten very far. Many of them did issue controversial executive decrees that were seen to encroach on legislative prerogatives, although Basabe-Serrano et al (2010) assess that the resulting executive-legislative tensions hindered the executives' ability to enact their agendas. Mid-1990s reforms and a new Constitution in 1998 were designed to strengthen executive powers, but presidential instability and failure to effectively enact policies continued (Basabe-Serrano et al 2010) and may have even been exacerbated by the reforms (Acosta and Polga-Hecimovich 2011). These challenges continued until the election in 2007 of Rafael Correa, who finally consolidated power in a new Constitution in 2008 and would govern for a decade as a strong personalist leader (Conaghan 2008; Basabe-Serrano et al 2010).

As with Fujimori, there were many potential factors at play during Correa's rise to power, including the fates of his predecessors. Several former Ecuadorian presidents had faced stop-and-go criminal proceedings before Correa's presidency (Lucas 2003), but the cases came to a head

in the months leading up to his election. In 2004, former president Bucaram and then-President Gutiérrez orchestrated a purge of the Supreme Court, which stalled pending criminal proceedings against Bucaram, as well as against Jamil Mahuad and Gustavo Noboa. The new court was dismissed, however, amid Gutiérrez's own removal from power in 2005 (Gumbel 2005). When the court reopened in 2006, it reinstated charges against Bucaram, as well as somewhat reduced charges against Noboa, but dismissed those against Mahuad (Conaghan 2012).

Launching his presidential campaign that year, Correa aligned himself with anti-corruption groups demanding to revive the case against Mahuad, but urged amnesty for his long-time associate Noboa, whom he depicted as the victim of politicized attacks (Conaghan 2012). Once he took office, Correa managed to petition for Noboa's exoneration through a legislative amnesty bill (Conaghan 2012). Meanwhile, government prosecutors continued to press the case against Mahuad; successive courts rejected Mahuad's petitions for dismissal and claims of politicization, moving the case to formal trial and ultimately convicting him in 2014 (Alvaro 2014). The political drama surrounding his predecessors' fates – and Correa's ability to punish his opponents and protect his friends – might have taught Correa that leaders' fates could be manipulated by whoever wielded the most power, influencing his own subsequent efforts to consolidate power.

1.2 The puzzle: power consolidation in developing democracies

Concerns about the concentration and potential abuse of government power have a long history in political studies and societal debates. Specific forms of arbitrary or unchecked power have evolved over time, especially since the last waves of democratization in the late 20th century. Yet the decline of hereditary or military dictatorships has not eliminated concerns about powerful leaders acting unilaterally or with impunity. As the examples above suggest, a trend has emerged even within democratic states, one that involves more nuance and debate about the merits and risks of concentrated power. This trend is the rise of “strongman” presidents (or similar heads of government), who may be popularly elected, yet manage to govern without effective constraints on their power.

The leaders cited in the motivating examples above were all elected to office, as were many of the leaders removed during the Color Revolutions and the Arab Spring. Some might even be considered especially democratic, since popular masses rather than party elites brought them to power (Anria 2017). These leaders are often described as charismatic populists or political outsiders, who campaign against the political establishment, win a personal mandate to govern as they see fit, then make decisions unilaterally while coopting or circumventing established parties, legislatures, and judiciaries. If scholars or observers call these leaders less than democratic (Way 2008; Bellin 2012), it is not because they suddenly became dynastic monarchs or the heads of military juntas. The concern is about something other than the way the

leaders were selected: it is about the power they wield once in office – how much they are able to decide and control on their own, and how they may be held accountable for those actions.

The following graphs highlight the shift from concerns about autocrats and military coups to greater concerns with the concentration and use of power in democratic states. Figure 1 shows several widely-cited datasets of military coups d’etat, which all show a clear global decline in coups since the 1970s, especially since the end of the Cold War (Powell and Thyne 2011, Nardulli et al 2013, Marshall and Marshall 2017). The decline is true of both successful coups (bolder lines) and coup attempts (fainter lines above). To the right is a plot of the most prominently cited index of democratic and autocratic regimes, Polity IV, which shows a sharp decline in autocracies and rise in democracies from the 1980s to today (Marshall et al 2017).

Figure 1. Coups d’etat and democracies vs autocracies over time

Coup data (left): Powell and Thyne 2011 (P&T), Nardulli et al 2013 Cline Center Coup D’etat Project (CDP), and Marshall and Marshall 2017 from the Center for Systemic Peace (CSP). Regime type data (right): Polity IV (Marshall, Gurr and Jagers 2017).

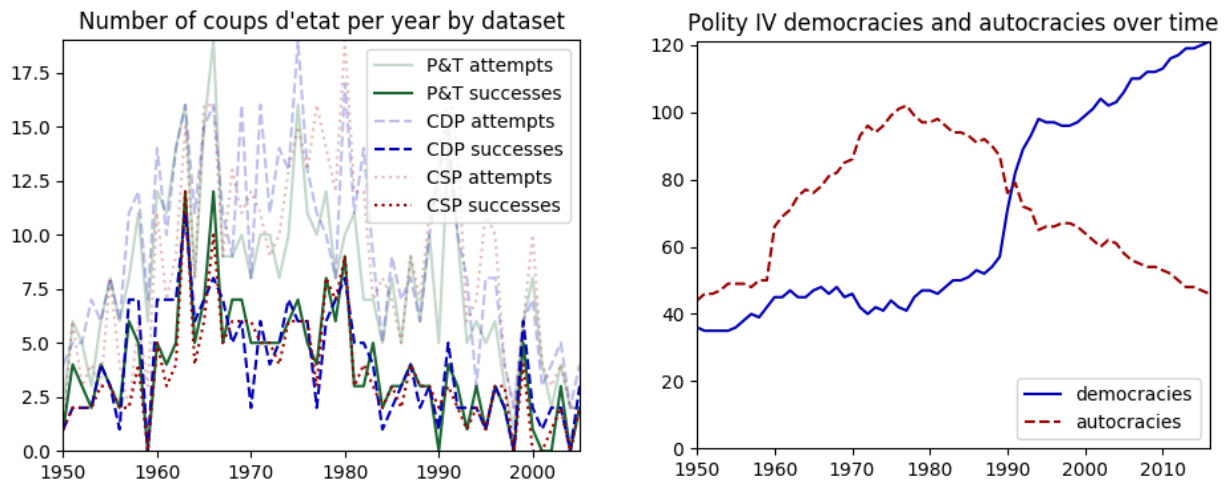
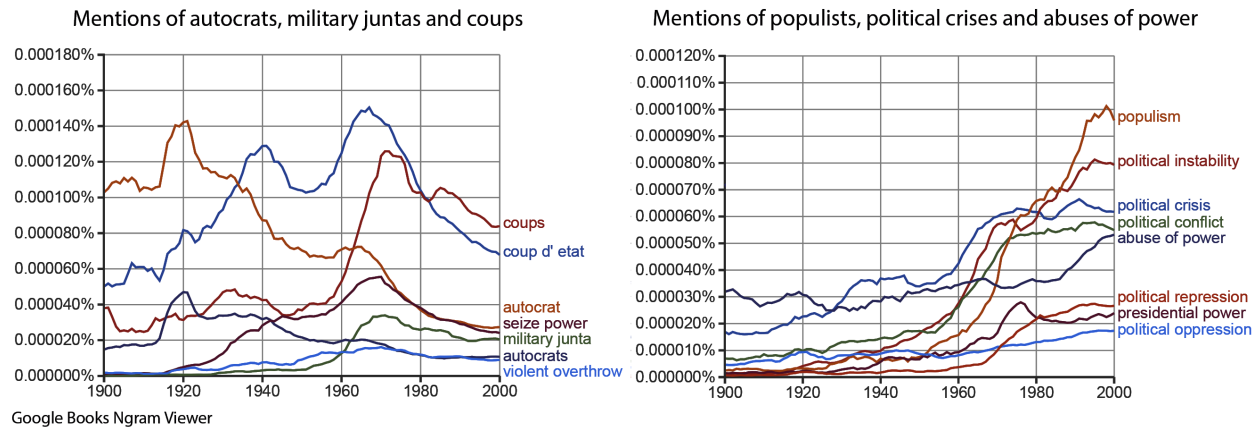


Figure 2 below shows this trend indirectly through public discourse, using data from the Google Books Ngram Viewer, a search tool that shows the number of occurrences of matching phrases in the Google Books corpus of scanned publications over time. The graph on the left shows a sharp drop-off in mentions of autocrats and coups d’etat since the 1970s, while the graph on the right shows increasing references to populism, political crises, and abuses of power, using books published in English in more than 100 countries around the world. Within each graph, I selected terms with similar frequencies, to facilitate readability. “Dictator” and “dictatorship” also declined sharply during the same period, while phrases like “presidential impeachment” also rose. These changes in literary references suggest that many people are concerned about concentrated power in governments other than authoritarian or military regimes.

Figure 2. Literary references to military coups and political conflict over time

Book mentions of key terms related to autocrats and coups d'etat (left) and to populism and abuses of power (right) from the Google Books Ngram Viewer, in English over the 20th century. The recommended/default period up to 2000 is used, since after 2000 the composition of the corpus changed in ways that affect the proportions of words observed (Michel et al 2010).



Along with increased public concerns about unchecked executive power in developing democracies, many countries have also seen a rise in the early removal and punishment of government leaders for abuses of power, through nonviolent means that do not result in overall regime change (Valenzuela 2004; Pérez-Liñán 2007; Kim and Bahry 2008). Yet as in the motivating cases, the early removal and prosecution of former presidents has not always resulted in successor governments that are more restrained. There is growing debate among both academics and practitioners about the merits of prosecuting former leaders, especially in politically unstable states. Cases against former leaders are often highly politicized and pursued on weak legal grounds, and can hinder peaceful leadership transitions by convincing embattled leaders to cling to power (Osiel et al 2000; Roehrig 2002; Olsen et al 2010; Conaghan 2012).

My goal in this dissertation is to study the relationship between these two modern trends: the concentration of executive authority in democratic states, especially presidents' efforts to consolidate power through executive decrees, and opponents' efforts to hold leaders accountable for abuses of power, through prosecution or other sanctions after they depart office. How do former leaders' punishments or rewards affect subsequent leaders' expectations about their own fates? And what is it about those potential punishments or rewards that drives government leaders to seek more direct and unchecked control of the state, or to accept constraints?

1.3 The argument: objectivity and predictability of observed fates

I will argue that government leaders not only care about how likely they are to face arrest, prosecution, or imprisonment after they depart office, they care about what they can do to

avoid that fate. Even if predecessors face charges of abusing power, incumbents may not be confident that they will be spared if they avoid certain actions, because they may perceive the charges to be politically motivated and their predecessors to be innocent of the crimes. I will argue that leaders respond differently to threats of future punishment, depending on whether the punishments appear to be objective and predictable based on the accused leaders' actions in office, or instead appear to be arbitrary and politicized. If former leaders' fates come down to who can manipulate the legal process the most, incumbents may decide that they need to secure more power to protect themselves. If punishments instead seem objectively determined based on real wrongdoing, incumbents may exercise greater restraint to avoid giving opponents cause to prosecute, leading to the accountability that proponents of prosecution seek.

1.4 Theoretical challenges

The motivations for and consequences of expanded presidential authority are highly contentious. Leaders often argue that they need more power in order to get things done, in cases of emergency or due to gridlock in other parts of government (Poguntke and Webb 2005; Pereira et al 2005). Opponents argue that more concentrated and unchecked authority leads to corruption and oppression of dissent (Cooper 2002; Human Rights Watch 2012). This debate is especially visible with regard to executive orders or decrees, which are often constitutionally mandated or publicly justified as necessary for expediency during emergencies or other urgent circumstances. Yet urgent demands can be creatively interpreted, and the lack of input or consent from other branches makes decrees more open to abuse. Presidents are rarely prosecuted for legislation, which is not solely attributed to them, but may be held accountable for unilateral executive action, of which decrees are the most prominent and legally binding form.

Questions about how presidents use their power and why they resort to decrees are difficult to study for multiple reasons. There are many factors that contribute to a president's decisions about which policies to enact and which legal channels to use. It can be difficult to tell whether particular leaders are seeking greater authority to accomplish important policy goals, or to protect their own interests and block rivals' challenges to their power. These motivations are often intertwined. Presidents might increase their chances of reelection by delivering more goods and services to constituents, or by enriching powerful political backers and suppressing dissent, or some combination of the two. Furthermore, presidents are likely to obscure their motivations and claim that they are only pursuing the powers necessary to carry out their agenda during difficult times, even while amassing power for its own sake.

Scholars often assume that government leaders' primary motivation is to retain office (Ames 1987; Bueno de Mesquita et al 2003). Yet most democracies now have term limits or other means of preventing leaders from governing forever, and while leaders may try to alter or circumvent those limits, many do not succeed (van de Walle 2002; Maltz 2007; Ginsburg et al 2011). How leaders use their power while in office might then also depend on how they expect to

be rewarded or punished for their actions after they depart. This is a difficult relationship to analyze, because it involves an expectation of some future event. There might be contemporary indicators of the likelihood that a current leader will face sanction when he/she eventually departs office, but we need to understand how those risks relate to the decisions the leader has yet to make, i.e. how the leader perceives that he/she may be able to alter or improve that fate.

It seems reasonable to assume that incumbents look at their predecessors' fates for indications of the risks they will face themselves; a common argument for prosecuting former presidents is to signal to future leaders that they cannot act with impunity (Olsen et al 2010). Yet as mentioned above, incumbents sometimes respond to such threats by exercising *less* restraint, not more (Osiel et al 2000; Roehrig 2002; Olsen et al 2010). Are incumbents not taking away the lesson that they too might be punished if they amass too much unchecked power and use it to serve their own ends? Or are they learning some other lesson; perhaps that any leader may be at risk of prosecution, *regardless* of whether they misuse power or govern with restraint? If many former leaders are prosecuted, but those cases are not clearly and objectively tied to real abuses, leaders might perceive that the only way to beat such threats is not to appease critics but to gain enough power to beat them.

In sum, we face difficulties dissecting and interpreting both of the phenomena in this study. In order to investigate when or why government leaders consolidate power, we need to distinguish increased use of executive decrees simply to deliver more programs and services to constituents, from decrees that alter the internal allocation of government authority. To analyze how leaders' expectations of future punishment or reward affect their decisions to consolidate power, we need to identify what aspects of former leaders' fates influence subsequent leaders' actions. Previous studies have raised more questions than answers about how we should expect government leaders to react to the prosecution of their predecessors. To answer those questions, we might need to look at more than just how many former leaders were prosecuted. If incumbents care most about how they can avoid similar fates, we need to look at how former leaders' fates were tied to their own actions in office, and whether those fates can be manipulated by whoever wields enough power.

1.5 Methodological challenges

This project's research question involves two distinct phenomena that are not only difficult to study theoretically, they are also both difficult to measure in practical terms. We need to identify indicators of when leaders are seeking to consolidate power, as well as indicators of what they expect to happen to them afterwards, and how (if at all) they expect the decisions they are making in office to affect their future fate. The cases from Peru and Ecuador described above involved dynamic and multifaceted conflicts over presidential authority and leaders' fates, suggesting that a static variable about whether the state fits a particular regime type or whether the president has chosen to be a certain type of leader would be inadequate. Even if we tried to

use a more continuous variable, for instance to classify each regime by the concentration of executive authority at regular intervals of time, it would be difficult to assess how much power a president actually has in any comprehensive way.

We might, however, be able to identify certain *changes* to the allocation of government authority, or certain actions that are more likely to change that allocation, without having to quantify exactly how far each action moves the balance of power. Some actions seem more likely to increase the president's power than others, and to do so in more lasting ways. For instance, the creation of new executive agencies, delegation of additional authorities or resources to those agencies, or revision of internal rules and restrictions, constitute changes to internal government authority structures that are often very durable and have widespread consequences for subsequent policy decisions.

These changes are documented in official records, such as legislation, executive decrees, and other administrative memoranda. Legislation usually reflects the efforts of multiple parties bargaining across the executive and legislative branches, making it difficult to attribute legislation to particular actors without analyzing specific steps in the legislative history of each bill. Administrative memoranda vary more across countries and are less likely to be available in comprehensive, comparable archives. But executive decrees have become a well established and prominent form of decision-making across countries that can be reasonably attributed to the president or chief executive of the state.

Yet decrees are used to enact a wide variety of decisions, many of which do not involve actions that are likely to change the internal allocation of government power. Like laws, decrees might be used to implement existing policies or programs, approving one-time contracts, acquisitions, or public works projects that benefit private parties, but do not determine which government actors have the power to make which future decisions. Decrees might also make changes to taxes and customs duties, or the criminal code, in important ways that still fall short of altering executive decision-making functions. Decrees are also often used to enact symbolic measures, like declaring a national holiday, which have no real impact on public authority.

To distinguish decrees that may change the institutional concentration of power from the many other decrees that involve more temporary, externally-focused policy actions and symbolic measures, we need more than a count of the total decrees issued, as previous studies have used. We need to look at what the decrees actually say they do. For sources of information that are not yet quantified, such as documents in which qualitative distinctions need to be turned into comparable units of analysis, scholars have traditionally relied on hand coding and expert judgment (Schedler 2012; Cruz et al 2016; Marshall et al 2017). Hand coding is labor intensive, which can make large datasets prohibitive to process with limited resources. This limitation is probably one reason why scholars of government institutions tend to rely on very high level measures like regime types (Geddes et al 2014; Marshall et al 2017), often based on key provisions in the constitution, which do not help us study the wider range of day-to-day policy decisions and institutional changes that might matter a great deal to societies today.

However, advances in computational methods are helping to overcome these obstacles, by making it easier to extract events automatically from text, identifying “who did what to whom” without an expert looking at each document individually. These methods are still developing, and can require considerable resources. In political studies, event extraction methods have mainly been used to identify violent conflicts reported in news media. Descriptions of events in news articles vary widely in the language used, making it difficult to consistently identify the same types of events using keyword or phrase searches. It is still more common to use computational methods to identify document-level topics or sentiments (Grimmer and Stewart 2013), than to extract specific instances of events, and the few prominent political event data projects to date generally involve large, well-resourced team efforts (Schrodt 2012).

Yet government documents like laws and decrees might be especially well suited to automated event extraction, since they tend to use consistent formats and precise formal language. Governments in developing regions like Latin America are increasingly making laws and decrees accessible through digital archives, as part of efforts to improve government transparency (Miguel-Stearns 2011), providing new opportunities for more systematic and detailed analysis of policy decisions and actions. Secondary news media might have historically been the only reliable source of detailed information about the actions of government leaders in developing countries. But today, it should be possible take methods that have been used to study large volumes of text from news media, and apply them directly to the more consistent and complete primary source records generated by government decision-makers themselves.

The approach I take to study government decisions in this dissertation involves dissecting events with considerable granularity and detail, aided by digital sources of information and computational methods for processing text and event data. I present new datasets for both the dependent variable (concentrations of power, using executive decrees that involve changes to internal government authority structures) and the main independent variable (leaders’ post-tenure punishments and rewards, including various aspects that might affect future leaders’ decisions). I apply natural language processing and machine learning to extract the main actions and target entities from executive decrees, using open source tools and limited project-specific resources. I also apply machine learning to measure the predictability of former leaders’ fates. Finally, I take a computational approach to the selection of statistical models for the main hypothesis tests, analyzing many permutations of the relevant variable specifications and parameters. I summarize the distribution of results across these tests, to help overcome the biases inherent in selecting only a few models to report, or the arbitrariness of a single significance threshold.

1.6 Case selection

A final point of introduction is the set of countries selected for inclusion in this study, to set the stage for the more detailed discussion of theory, methods, and empirical analysis in the subsequent chapters. For this project, I have chosen a set of countries that have had strong

variation in both the dependent and independent variables, which are sufficiently well documented and available in digital archives to make the analysis feasible. The sample is comprised of the five countries in South America's Andean region: Peru, Ecuador, Bolivia, Colombia, and Venezuela. These countries share many similarities in terms of colonial and post-colonial history, population characteristics and culture, which are not the subject of this study. They have also experienced considerable differences across countries and within each country over time in the political institutions and leadership decisions at the heart of this project.

1.6.1 Introducing the remaining cases: Bolivia

Bolivia's recent political history has many similarities to Ecuador's, especially in a succession of weak presidents until the election of Evo Morales in 2006. The weakness of Bolivia's presidency since the last return to democracy stems in part from an unusual provision in the previous constitution, which allowed the legislature to select the president from among the top electoral candidates if none received a majority of votes in the popular election (Gamarra 1997). This process weakened the president's electoral mandate, since collusion among second- and third-place parties often prevented the top vote-getter from securing the presidency (Slater and Simmons 2012). Gamarra (1997) argues that the National Congress's claim to be the ultimate source of democratic legitimacy in Bolivia, the president's dependence on congressional parties who demanded large shares of limited national resources for patronage, and gridlock imposed through congressional investigations and other obstacles to executive activity, contributed to the gradual erosion of presidential authority through the 1980s and 1990s.

Bolivia also underwent a period of economic crises, which provided justification for numerous executive decrees. Poor economic management, the gradual decline of the traditional mining sector, and changing circumstances in international credit markets produced a major economic crisis in the mid-1980s (Morales and Sachs 1987). Bolivia defaulted on its debts and experienced severe hyperinflation from 1984-1985, further exacerbated by a dramatic fall in world tin prices, one of Bolivia's top commodity exports (Zuazo 1985). Protests, strikes, and pressure from various parties compelled President Hernán Siles Zuazo to depart office early in 1985 (Gamarra 1997). Siles' successor, Víctor Paz Estenssoro, issued a major economic stabilization package by executive decree (Morales and Sachs 1987; Gamarra 1997), then reorganized power to implement the reforms by creating two super-ministers overseeing economic and political affairs, respectively (Gamarra 1997).

Paz Estenssoro's and his successors' positions depended on pacts among party elites, which continued to grow disconnected from their constituents, as elections were superseded by bargaining in the legislature to choose presidential winners and the formation of their cabinets (Gamarra 1997). Eventually, disenchantment with this partisan *pactismo* and continuing economic hardship among a mobilizing indigenous population fueled protests that turned violent, resulting in the killing of dozens of civilians and driving then-president Gonzalo Sánchez de

Lozada out of power in 2003 (Tobar 2003). Interim president Carlos Mesa was also driven out early in 2005, paving the way for the election of an indigenous coca union leader and populist, Evo Morales (Webber 2010, Slater and Simmons 2012). Morales consolidated executive authority in a new Constitution that secured his direct popular election and, through favorable judicial interpretations, enabled him to secure two more terms in 2009 and 2014 (Brice 2009, Ellner 2011, Neuman 2014). He is currently seeking a means to run again in 2019 (Anria 2017).

Another element of the story that is not frequently cited is that early in Morales' tenure, predecessor Sánchez de Lozada was in exile in the U.S., fighting a mounting battle over his extradition and prosecution for the 2003 massacre of protestors (Faiola 2007, Greenwald 2012). Efforts to punish the former president began with campaigns by families of victims and human rights groups, filing petitions demanding legal action (Faiola 2007, Acthenberg and Currents 2015). In 2004, Bolivia's Congress voted to authorize proceedings against the former president and cabinet ministers, and the Attorney General's office filed preliminary charges of genocide with the Supreme Court (Azcuí 2004, Keane 2006, Acthenberg and Currents 2015). In 2008, the government requested Sánchez de Lozada's extradition from the U.S., which was denied in 2012 (Acthenberg and Currents 2015), but a second request filed in 2014 was allowed to proceed in 2016 (Farthing 2016). Meanwhile, in 2009, the Bolivian Supreme Court began a trial of Sánchez and 17 other former officials, with Sánchez de Lozada still abroad (Morgan 2009). In 2011, the court convicted seven of the accused, but Bolivian law prohibits conviction in absentia, such that Sánchez's de Lozada conviction is still pending his extradition (Weinberg 2011).

Sánchez de Lozada is widely perceived to have been protected by the U.S. for so long in part due to his close ties to American political elites, as well as hostile relations between the U.S. and the current populist administration of President Morales. Sánchez de Lozada was educated in the U.S. and was strongly in favor of U.S.-backed free trade policies as well as the U.S.'s war on drugs (Greenwald 2012). He had also hired a Washington consulting firm led by prominent Democratic political operatives to help him secure his election to the presidency in 2002 (Greenwald 2012, Acthenberg and Currents 2015), as well as an advisor who subsequently served as a Senior Director for Western Hemisphere Affairs on President Obama's National Security Council staff, and an attorney who coordinated President Clinton's legal defense during the latter's impeachment trial (Acthenberg and Currents 2015). While Morales did not choose to cultivate the same powerful alliances within the U.S., Sánchez de Lozada's ability to escape justice during Morales' rise to power may have contributed to Morales' own calculations about how to use his power while in office in order to maximize his long-term fate.

1.6.2 Introducing the remaining cases: Colombia

The presidencies in Colombia and Venezuela – the two northern Andean states – have historically been considered much stronger than in their southern neighbors (Archer and Shugart 1997, Uprimny 2003). These last two cases also resulted in fewer post-tenure sanctions during

the period of study, but they help to broaden the context of the study and the scope of other factors that may have influenced executive decrees and power consolidation. The two Northern Andean states have been described as more mature democracies than the Southern three, since they have maintained civilian governments for longer periods of the 20th century (Martz 1999, Uprimny 2003). Yet both also have experienced more political violence and civil unrest in recent decades than the other three countries. Colombia experienced more insurgent and drug-related violence, which both attracted and may have been exacerbated by U.S. antidrug efforts (Petras and Brescia 2000, Peceny and Durnan 2008), while Venezuela has experienced waves of protest and unrest driven by poor economic management and an increasingly oppressive government, much more hostile to the U.S. and international community (Duddy 2015, Gallón et al 2017).

Colombia's latest transition to full democracy passed through an intermediate stage that contributed to the political factors influencing presidential power. A decade of violent conflict (called *la Violencia*) ended in 1958 with a pact between the country's two dominant political parties agreeing to alternate the presidency every four years, as well as splitting congressional seats and cabinet offices (Posada-Carbó 1998, Hinojosa and Pérez-Liñán 2003). The National Front ended in 1974 with the restoration of open presidential elections. While the pact remained in place, the two parties controlled the legislature and government, but the lack of inter-party competition weakened the parties themselves and fed their fragmentation into patronage-based "micro-enterprises" (Posada-Carbó 1998, Hinojosa and Pérez-Liñán 2003). The two parties' mutual veto power within the legislature also may have contributed to a tendency to delegate power to the executive (Hinojosa and Pérez-Liñán 2003).

Presidential powers in Colombia were considered especially strong in the 1980s, although the 1991 Constitution added some constraints, granting Congress the power to repeal or amend executive decrees issued based on delegated authority (Archer and Shugart 1997). However, the president still has considerable agenda setting power, including the exclusive right to introduce major economic bills and propose the budget, along with strong appointment powers and the ability to make significant policy decisions through decrees, including regulating by decree any bills passed by Congress (Archer and Shugart 1997; Uprimny 2003). The president also has vague and broad emergency powers; "state of siege" powers were originally granted to enable presidents to respond rapidly to crises, but have become a regular instrument of decision making used to make numerous important political and institutional reforms by decree (Archer and Shugart 1997; Uprimny 2003).

Colombia is the only country in the dataset which did not experience any early removals or prosecutions of former presidents in the period since democratization. Related pressures existed, however, and at least one president faced a serious scandal that almost removed him from power. In the late 1970s to 1980s, several presidents (Alfonzo López Michelsen, César Turbay, and Belisario Betancur) faced some preliminary legislative investigations, but none reached the House Floor (Archer and Shugart 1997). Colombia also continued to face armed guerillas throughout this period, especially the *Fuerzas Armadas Revolucionarias de Colombia* (FARC) and the National Liberation Army (ELN), countered in the 1980s by paramilitary

groups, some aligned with drug cartels, which played important roles in politics as well. One scandal involving cartel funding to a political campaign in 1994 almost resulted in the impeachment of President Ernesto Samper two years later (Ambrus 1996). Investigations against Samper were carried out by congressional committees and by the Prosecutor General's office, which formally accused him of crimes in February 1996 (Archer and Shugart 1997). Eventually the legislature voted to acquit the president, preventing further judicial action (Ambrus 1996).

While opposition to Colombia's presidents did not reach the level of formal prosecution for abuses of power, there has been considerable variation in their post-tenure legacies and roles in politics and society. Some former presidents secured other government positions, including Andrés Pastrana, who became Ambassador to the U.S., a position which he resigned when then-president Álvaro Uribe appointed the somewhat disgraced Samper to serve as Ambassador to France (a position which the latter ultimately did not fill) (Goodman 2006). Uribe himself served in the Senate both before and after his presidency, and led opposition to his successor Juan Manuel Santos's peace deal negotiated with the FARC (Brodzinsky 2016b). Uribe also spent time abroad immediately after his departure from the presidency, obtaining a visiting scholar position at Georgetown University, despite a letter signed by 150 U.S. scholars asking the university to reconsider in light of alleged human rights abuses and ties to paramilitary groups (Miller 2010). Current President Juan Manuel Santos has faced investigation for receiving campaign bribes from the Brazilian conglomerate Odebrecht, as part of a corruption scandal that has implicated multiple political leaders in the region (Manetto and Palomino 2017).

1.6.3 Introducing the remaining cases: Venezuela

Venezuela's recent history also contains episodes that are consistent with the theory to be developed in this project. From the 1970s through 1990s, Venezuelan presidents gradually expanded their interpretation of fairly weak formal decree-making authorities in the 1961 Constitution. This included the authority to suspend certain liberties during emergency times, short of declaring a full state of emergency. One of the most blatant episodes occurred in June and July of 1994, when President Rafael Caldera suspended multiple articles of constitutional rights, including rights to free transit, assembly, and peaceful protest, as well as economic liberties (Vinogradoff 1994, Crisp 1998). He also used the latter suspension of economic liberties to decree the formation of two new government entities: the Financial Emergency Board and the Exchange Administration Board, creating lasting institutional changes out of what were supposed to be temporary emergency measures (Brooke 1994, Crisp 1998). When Congress voted to reinstitute the constitutional guarantees in July, Caldera immediately re-suspended the same articles (Crisp 1998). In press statements, he also threatened to seek a popular referendum on his right to rule by decree if challenged again by Congress or the courts, which compelled the political opposition to back down (Crisp 1998). Notably, these events occurred at a time when two predecessors – former presidents Jaime Lusinchi and Carlos Andrés Pérez – were on trial for

corruption in office (Freed 1993). Lusinchi's charges were eventually dropped (Rueda 2014), but Pérez would be convicted two years later (Schemo 1996).

If Rafael Caldera held relatively strong sway over the political system in Venezuela in the 1990s, his legacy is overshadowed by a much more prominent personalist leader, Hugo Chavez. Chavez's rise to power had its roots in the decay of the traditional party system in ways that are similar to some of the other cases in this study. Venezuela also experienced fluctuating inter-party and intra-party competition, between and within the two long-standing traditional political parties in the country, Acción Democrática and Social Christian COPEI (Crisp 1998b, Martz 1999). From the 1960s through 1980s, these parties experimented with holding primaries, orchestrating a form of electoral college, and imposing eligibility restrictions to ensure their preferred candidates reached the presidency, even if the winner was officially decided by popular vote (Crisp 1998b, Martz 1999, Carey and Polga-Hecimovich 2005). Martz (1999) argues that this led to "a sharp contrast between organizational preference and popular sentiment" in the 1982 victory of Carlos Andrés Pérez to his second presidential term, which became highly contentious and eventually resulted in his impeachment.

Increasing frustration with Venezuela's traditional parties contributed to the cult of personality that first grew around Chávez after he led an attempted coup d'état in 1992, despite the coup's failure and his imprisonment for two years (Hawkins 2010). He was pardoned after Pérez's removal (Sylvia and Danopoulos 2010), and successfully campaigned on an anti-party populist platform to win the presidential election in 1999 (Hawkins 2010). Like other prominent consolidators of power, Chávez moved quickly to orchestrate a new constitution that permitted his immediate reelection, increased the president's authority over the military and other major instruments of state power, abolished the Senate to create a new unicameral National Assembly, and contained other vague provisions that subsequently gave Chávez room to censor the media and oppress dissent (Kovaleski 1999, Selçuk 2016). Chávez and his successor Nicolás Maduro both governed extensively by decree throughout their time in office, making use both of delegated decree-making powers and emergency powers, and using those decrees to make important policy decisions and substantial changes to the governing apparatus (Carroll 2008, Forero 2010, Diaz-Struck and Forero 2013, López Maya 2014, Brodzinsky 2016a).

Chávez enjoyed strong personal popularity throughout his tenure, but despite publicly designating Maduro as his chosen successor, Maduro has faced much greater opposition (López Maya 2014). The challenges confronted by Maduro's administration may in part represent the legacy of Chávez's consolidation and personalization of control over the Venezuelan state, neglecting systematic institutionalization of authorities and processes in favor of improvised or emergency mechanisms, parallel state agencies with overlapping roles, and poor resource management that led to shortages in publicly provided goods and services (López Maya 2014, Hawkins 2016, Selçuk 2016). Protests have repeatedly broken out since 2014, and heavy-handed responses by police have resulted in deaths, injuries, and large numbers detained, provoking accusations of severe human rights violations (Cawthorne and Ulmer 2016, Rothwell et al 2017). Despite low approval ratings, Maduro continues to cling to power; in 2017 he held elections for a

Constituent Assembly to draft a new constitution that critics argue is a ruse to perpetuate his hold on office (Mogollon and Kraul 2017, McCoy 2017).

Venezuela's Hugo Chávez, Ecuador's Rafael Correa, and Bolivia's Evo Morales are often compared to each other (Ellner 2011), as well as Peru's Alberto Fujimori at least with regard to his style of leadership, if not his political ideology or economic policies (Campbell 2000). The fact that all four were considered political outsiders, however, does not mean that they had similar backgrounds. Morales was a labor union leader and the country's first indigenous president (Slater and Simmons 2012). Fujimori, in contrast, was a university professor and dean, and the son of Japanese immigrants, representing an intellectual expatriate elite (Rapp 2017). Chávez came from a military background, gaining notoriety first by leading the failed military coup in 1992 (Hawkins 2010). The frequent comparison of these leaders as charismatic outsider populists may have less to do with their political identity or origins, and more to do with choices they made in pursuit of power, as well as how they used that power once in office. In other words, these political actors might have responded to practical incentives and pressures in similar ways that transcend personal characteristics, helping us to identify institutional factors that influence the allocation of power across political contexts.

The countries also vary in their political economic models, with Venezuela's, Ecuador's, and to some extent Bolivia's leftist governments controlling major state-owned enterprises, while Peru and Colombia maintain much more liberal market-oriented policies (Flores-Macias 2010, Dargent 2011, Bremmer 2017). These countries have also had considerable variation in foreign relations, especially with the United States, the nearest global power to the region. Relations with the United States have been influenced by each country's economic policies and its approach to countering the drug trade. Peru and Colombia have had the strongest relations with the U.S. (Peceny and Durnan 2008, Taft-Morales 2009), Venezuela has the most antagonistic, Ecuador's and Bolivia's are highly strained (Ellsworth and Symmes Cobb 2017, Ellis 2017, Kwong 2017).

The factors discussed throughout these motivating cases helped shape the theoretical and empirical development of this project, and are incorporated into the project's explanatory and control variables in various ways, as will be explained in subsequent chapters. The inclusion of five neighboring countries with distinct recent political histories and strong variation in the political conditions, presidential power, and post-tenure fates behind the research questions at hand, should help to make the results generalizable to other countries and regions as well.

1.6.4 Summary of country-years covered

In sum, the Andean countries have experienced a mixture of strong and weak presidencies, including a few of the most notorious personalist leaders, but also numerous presidents at other times who did not expand their power in similar ways. Four out of five of the Andean countries have also had presidents removed early, accompanied by investigations and prosecutions of former leaders for corruption or other abuses of power. The region has also

experienced considerable variation in other factors cited in the literature on expansions of executive authority, such as the strength or decline of traditional political parties and varying degrees of legislative opposition, as well as the economic and social crises that drive demands for more urgent government action. Using a set of similar neighboring countries helps control for unobserved factors, while bringing into focus those phenomena that do vary across them and are of interest to the study.

The Andean countries also provide a feasible target of study. Like many developing democracies, they have recently made large volumes of legislation and decrees accessible through digital archives, as part of efforts to improve government transparency (Miguel-Stearns 2011). These countries' similar institutions and common language also make the data collection and measurement more feasible and consistent, since all five countries have comparable forms of executive decrees that use similar styles of legal discourse. The information extraction methods discussed in Chapter 3 take time to develop, and tend to be language and domain specific, making it much more difficult and costly to apply the same processes to multiple regions. I have also been exploring less "supervised" or more inductive approaches to information extraction, that could more easily be applied to additional languages, but those methods are not yet effective enough for practical application, putting off the addition of other regions to future research.

I focus on democratically elected presidents since the end of the last military regimes, where relevant, since those regimes ended toward the end of the 1970s or early 1980s. Peru's last military regime handed over power to an elected government in 1980, Ecuador's last junta did so in 1979, and Bolivia's in 1982. For Colombia, I chose a comparable starting point at the end of the National Front, a governing pact between the two main political parties in which they agreed to alternate the presidency every four years, which ended in 1974. In Venezuela, the last military regime ended in 1958, but the period that I am able to study is constrained by the more limited availability of decrees in public web archives for that country.

For each of these countries and start years, I include in the dataset all presidents who were elected from the stated year onward. In some cases, digital archives of decrees were only available for more recent years (as noted in Chapter 4). In those cases, the dataset covers the years in which I have decree data, but includes post-tenure fate events for any former leaders who were elected since the start of the democratic period noted here, and who were still alive when the decree archives began.

While this study focuses on a single region, many of the ideas behind the proposed theory were initially developed through my experience in other developing regions, including in Asia and the Middle East, and I have sought to identify variables and hypotheses that have broad implications beyond Latin America. Some of the main trends that make Latin America ripe for this study have also been documented in other "third wave democracies" like the former Soviet States in Eastern Europe, including the proliferation of young electoral systems with weak political institutions, the rise of elected populists who proceed to govern with a high degree of unilateral decision-making authority, and frequent constitutional removals and prosecutions of presidents for misuses of power.

1.7 Overview of remaining chapters

The rest of this dissertation is organized as follows. Chapter 2 presents the theoretical framework for this project, reviewing existing literature and introducing new concepts for the dependent and independent variables, as well as the hypothesized relationships between them. Chapter 3 describes the methodology used to measure the dependent variable, from collecting executive decrees to parsing and classifying their main actions and target entities. It includes accuracy tests in relation to human coding, comparing machine learning classifiers and rule-based pattern matching, and arrives at an ensemble approach for the final classification process.

Chapter 4 defines a set of control variables derived from previous theories about presidential power, and presents statistical analysis of the encoded decree categories in relation to these control variables, to set the stage for the new hypothesis tests. The analysis in Chapter 4 provides descriptive statistics about presidents' use of decrees in this project's dataset and the range of factors that appear to influence them. I use the control variables to validate the conceptual interpretation of the decree categories that I have encoded, showing that they represent meaningfully distinct actions leaders take in different circumstances. In particular, I show that the main category of interest – decrees that empower the government executive – are issued when existing theories would tell us to expect presidents to consolidate power.

Chapter 5 and Chapter 6 present the project's main hypothesis tests, defining measures used for the new independent variables – former leaders' punishments and rewards – and the statistical models used to test the hypothesized relationships. Chapter 5 covers the overall categories for post-tenure sanctions, reprieves, and rewards, along with subcategories for different instigating actors. Chapter 6 digs deeper into sanctions, introducing additional measures for different stages of the legal process, estimates of objectivity or predictability, and interactions with near-term crises that might make leaders more concerned about their post-tenure fate. Chapter 7 concludes and discusses policy implications and broader applications of this research.

Chapter 2: A Theory of the Motivations for and Costs of Consolidating Power

In this chapter, I draw on existing literature and lingering debates to develop a theory of how punishing or rewarding former government leaders influences how incumbents choose to consolidate and use their power while in office. More specifically, I address why presidents choose to engage in unilateral decision-making through executive orders or decrees, and why they might use decrees to change the allocation of power within governing institutions. There are many factors that contribute to leaders' policy decisions, including the desire to win reelection or secure a positive legacy, on which most previous studies focus. Separate sets of literature have studied the punishment of former leaders for abuses of power, which might also factor into how incumbents use their decree-making authority. Yet few studies have combined these factors and tested the relationship between former leaders' punishment and subsequent leaders' actions.

The theory presented in this chapter addresses how the threat of future punishment affects government leaders' decisions to consolidate power, in ways that are designed to protect their own interests and block anticipated threats. The intent is not to argue that leaders *do not* seek power for the reasons more commonly assumed – such as to enact their policy agenda and build support for reelection or for a positive legacy. I argue that leaders *also* sometimes seek power for its own sake, or for purposes other than to get more things done in office. In order to identify and study this second motivation, which might be influenced by different factors than the first, we need to be able to distinguish the actions leaders take in greater detail. We may also need to identify different aspects of leaders' fates that not only indicate how likely incumbents are to be punished or rewarded, but how their own actions while in office may influence that fate.

In the following sections, I first discuss the previous literature on presidential power and executive orders, and the literature on the removal and prosecution of government leaders. I identify useful established explanations, as well as remaining debates that the previous studies have not yet addressed or empirically resolved, to situate this project's research question in the broader context of what has come before. I then define the new concepts that I use to distinguish the relevant aspects of the dependent and independent variables in this study, and the proposed hypotheses for the relationships of interest. This chapter establishes the conceptual framework for the empirical inquiry in the remaining chapters, which will define the data sources and precise variable measures used to operationalize the concepts described here.

2.1 Existing literature on presidential power and executive orders

This project relates to several overlapping bodies of work about strong, centralized, and personalist presidencies (or similar heads of government). The literature on “presidentialization”

(Poguntke and Webb 2005; Mughan 2000) and various forms of “plebiscitary” or “delegative democracy” (O’Donnell 1994; Weyland 2001; Mayorga 2006; Green 2009) focuses more on the electoral rise of charismatic outsiders, who win a broad popular mandate to govern as they see fit, unhindered by a traditional party structure or a campaign focused on a specific policy agenda. A small but important body of work addresses the expansion of unilateral executive decision-making, including executive orders in the U.S. (Moe and Howell 1999; Mayer 2001; Cooper 2002; Howell 2003), and executive decrees in Latin America and several other prominent presidential systems (Carey and Shugart 1998; Protsyk 2004; Pereira et al 2005).

Existing studies of presidential power and executive orders mainly rely on traditional methods of data collection and measurement. Studies of executive orders and decrees generally measure their use in terms of the total number issued each year, or the number that meet a certain threshold of salience, which are hand coded from a limited number of government documents and/or secondary sources. Wright (2014) focuses on executive decrees in Andean countries that are enacted based on emergency powers (finding a total of 292 such decrees in three countries over ten years), and groups them by whether the decrees cite social unrest and whether they involve the use of force. Mayer and Price (2002) and Howell (2003) count executive orders that are recognized as genuinely important based on mentions in press, Congressional hearings, federal litigation or court rulings, and in research by legal or presidential scholars.

Most relevant to this study, Mayer and Price (2002) also count an executive order as significant if it created a new institution with substantive policy responsibility. Mayer (2001) recounts historic examples of executive orders in the U.S. that have been used to create new executive agencies, restructure powerful government institutions, or revise internal rules in ways that gave the executive advantages in disputes with Congress. Howell and Lewis (2002) similarly find a rise in new programs and agencies unilaterally created by executive orders in the U.S. These decisions constitute major changes to state institutions and authority structures, beyond the temporary emergency measures usually described as justification for decree powers. Presidential scholars note that while decree authority is usually narrower than legislative authority, presidents often exploit constitutional ambiguity to gradually assume a broader scope of those powers (Moe and Howell 1999; Mayer 2001; Marshall 2008). The potential for decree authority to be creatively reinterpreted and used to make lasting changes to internal authority structures suggests that there is much to learn by distinguishing different actions taken through decrees, which may have different implications for leaders’ objectives and governance outcomes.

Established theories about presidential power point to some common explanatory factors for the use of executive orders or decrees, which I use as a starting point for my own analysis. The main theories generally assume that presidents seek to enact their policy agenda, either to win reelection or to build a strong legacy (Moe and Howell 1999). In either case, the implication is that leaders want more unilateral decision-making power in order to get more things done, and to take credit for those accomplishments. Several studies show that executive orders or decrees increase in volume during periods of weak legislative support for the executive, or high political

gridlock (Howell 2003; Pereira et al. 2005). Mayer (2001) and Pereria et al (2005) also find some evidence for a negative effect of popularity on the use of decrees by South American presidents.

In the broader literature on “presidentialization” or “delegative democracies”, scholars also argue that presidential power becomes more concentrated as electoral campaigns become more personalized through the use of mass media (Mughan 2000; Weyland 2001; Poguntke and Webb 2005; Green 2009), and as government bureaucracies expand and become more complex, requiring a strong central leader to coordinate and oversee an effective policy agenda (Poguntke and Webb 2005; Marshall 2008; Green 2009; Johansson and Tallberg 2010). These arguments imply that increasingly strong presidents are somewhat inevitable in modern contexts (Marshall 2008), but do not explaining variation in presidents’ decisions to concentrate power.

In sum, existing theories about presidential power and the use of executive orders generally assume that leaders seek more power to enact their policy agenda, either to improve their chances of reelection or to build a positive legacy. The main explanatory factors identified so far include legislative opposition or gridlock, and urgent demands for swift action arising from crises or emergency circumstances. However, the existing theories discussed here do not address the possibility that presidents might also seek more power to weaken rivals and preserve their own control over the state, rather than simply trying to get things done in a modern context. It remains unclear whether leaders facing political opposition are really seeking more power to deliver policy results and regain support, or to protect themselves against the legal and political challenges their opponents might launch against them.

One reason why previous studies do not address these different motivations might be that they haven’t dissected the data enough to identify distinct types of actions leaders take in pursuit of competing objectives. In order to answer questions about the effectiveness of punishing leaders for abuses of power, we have to understand how leaders seek to protect themselves from punishment, as well as how leaders seek to build support for reelection and future rewards. To figure out what presidents are actually trying to do with their power, I build and test a theory that distinguishes specific types of decisions taken through executive decrees. I discuss the proposed categories of executive action, and how they relate to previous explanations for executive orders and to my new hypotheses about leaders’ fates, in the sections below.

2.2 Existing literature on presidential removal and post-tenure fates

If leaders vary in their motivations for seeking greater executive authority, do they also vary in the costs of doing so? Government leaders may stand to lose if they seek too much power or otherwise overplay their hand, especially in developing or weakly institutionalized countries. Recent studies in Latin America and other developing regions have chronicled a growing trend of early presidential removals since the last wave of democratization (Valenzuela 2004; Pérez-Liñán 2007; Marsteintredet and Berntzen 2008; Kim and Bahry 2008, Hochstetler and Samuels 2011). Unlike with revolutions or coups, these presidential removals tend to use constitutional

processes, such as impeachment or legislative votes of incapacity, and do not necessarily result in regime change (Pérez-Liñán 2007; Kim and Bahry 2008).

There is evidence that a combination of public protest and conflict with the legislature leads to increased risk of early presidential removal from office (Pérez-Liñán 2007; Hochstetler and Samuels 2011; Kim 2014). Factors contributing to popular mobilization against the incumbent can include political scandals, but also economic downturns, since the chief executive is commonly assumed to be responsible for managing the national economy (Pérez-Liñán 2007; Kim 2014). Regarding legislative opposition, several scholars find that early removals often occur when the president's party controls fewer seats in the legislature (Valenzuela 2004; Pérez-Liñán 2007). However, Kim (2014) also finds that the risk of early removal increases when the incumbent enjoys stronger presidential powers, arguing that stronger presidents tend to engage in less collaboration and compromise. In other words, rivals may challenge the incumbent based on their own concerns about the balance of institutional power, while drawing upon popular dissatisfaction to legitimize their claims.

Even more common than impeachment, former heads of government are increasingly being prosecuted for corruption, human rights violations, or other abuses of power (Osiel et al 2000; Reyes and Gerber 2011; Conaghan 2012). Proponents of prosecuting former leaders often argue that it is necessary to enforce accountability and signal to current or future leaders that they cannot act with impunity (Olsen et al 2010; Conaghan 2012). Yet there remains considerable debate about the consequences of prosecution for democratic institutions and subsequent leaders' actions. Scholars have found that prosecutions of former presidents are often highly politicized, based on questionable legal grounds, and may drive embattled leaders to cling to power in ways that can undermine political stability and peaceful leadership transitions (Osiel et al 2000; Roehrig 2002; Olsen et al 2010; Reyes and Gerber 2011; Conaghan 2012).

Reyes and Gerber (2011) find that prosecutions are more common when the successor government has weak legitimacy and perceives a continued political threat from predecessors, such as when former leaders retain popular support. In other words, prosecution for past wrongdoing may be used by rivals to prevent popular former presidents from returning to power. Conaghan (2012) distinguishes prosecutions driven by independent accountability-seeking institutions (e.g. autonomous agencies, courts, or civil society organizations) from those driven by partisan actors (e.g. legislators or a successor government). In dissecting a set of cases from Ecuador, she finds that prosecutions often unfold through political struggles, changing course with cabinet shuffles and judicial turnover, as various actors lobby to open, continue, or conclude investigations and trials (Conaghan 2012).

The recent scholarship on the prosecution of former presidents has mainly involved descriptive case studies and theoretical explorations of the motivations and conditions under which former leaders are more likely to be prosecuted (Roehrig 2002; Reyes and Gerber 2011; Conaghan 2012). These studies suggest that since prosecutions are often politicized, "victor's justice" may undermine the integrity of the courts and have negative consequences for democratic institutions, rather than resulting in effective accountability (Osiel et al 2000; Reyes

and Gerber 2011). Yet these studies have not empirically tested the relationship between the prosecution of former presidents and subsequent leaders' policy decisions.

A few additional studies about leadership change more broadly indicate that leaders consider the fates of former leaders when deciding how to use their own power while in office. Baturo (2010) shows that previous leaders' fates influence successors' decisions to voluntarily step down or attempt to circumvent term limits. Chiozza and Goemans (2011) find that predecessors' fates affect current leaders' decisions to engage in international conflict, as a means of strengthening their hold on office. They argue that leaders who expect to be forcefully removed from office perceive little to lose and much to gain from engaging in international conflict, while leaders who anticipate a regular removal from office are less likely to engage in conflict because have much more to lose if they do. Both papers use the Archigos dataset (Goemans et al 2009), a global dataset of political leaders which records whether each leader survived, was exiled, imprisoned, or died during the first year after office. Baturo (2010) extends the period to three years after departing office and adds a category for criminal investigation. The variables he uses still constitute a single fate for each leader, and do not distinguish other aspects of the leader's fate like the degree of politicization.

In sum, existing studies about the early removal and prosecution of government leaders raise questions about whether such sanctions are objective responses to real abuses of power, or simply politicized vengeance used to eliminate rivals from the political arena. The literature suggests that there is open debate about whether such sanctions really produce effective accountability, since politicized prosecution might not deter – and might even encourage – future leaders' efforts to cling to power. Yet few studies have actually tested the relationship between former leaders' fates and subsequent leaders' actions. The few studies that do, only test the effects of whether former leaders were punished or not, in broad terms, but do not address the objectivity or politicization of those punishments. They also only compare former leaders' fates to a limited set of actions future leaders might take, such as engagement in war.

In this project, I build on these previous theories about presidential power and the prosecution of former leaders, while developing new conceptual distinctions about both executive action and leaders' post-tenure fates – drawing on new data sources and methods – to be able to address questions that have not yet been resolved, regarding whether politicized punishment of former presidents really deters (or perhaps even provokes) subsequent leaders' efforts to consolidate power.

2.3 Why do leaders seek greater unilateral decision-making powers?

As mentioned in section 2.1 above, common explanations for increasing executive authority include that heads of government need more power to enact policies and programs, manage a complex bureaucracy and agenda, respond effectively to urgent societal demands, and overcome gridlock in other more fragmented parts of the political system. Despite these

justifications, leaders with highly concentrated and unchecked authority are often accused of abusing that power, to enrich themselves and their inner circle, or to target their opponents with force or other forms of oppression. From the perspective of the accusers, leaders are not seeking additional powers in order to more effectively carry out a policy agenda; they are seeking greater control of government to protect their own interests and maintain a firmer hold on power.

The difficulty for researchers is that leaders are often motivated by *both* the desire to get important things done and to protect their own interests, and these two objectives are often mutually reinforcing. We do not expect government leaders to be selfless altruists; leaders should perform well when it is in their interest to do so. Political scientists usually assume that a political leader's primary goal is to stay in office. That is, democratic leaders make policy decisions that maximize their chances of winning the next election, such as by delivering programs and services their constituents desire. Yet political leaders can also take actions to retain office that do not involve delivering policies and programs to win over supporters. For example, leaders might seek changes to electoral laws or representational district boundaries in order to increase their chances of retaining office with the *same* supporters they currently have.

What about leaders who are ineligible for reelection? Many countries today have presidential term limits, which may involve a lifetime maximum or a ban on consecutive terms, both of which are common in Latin American states. Some presidents have been successful at circumventing or removing term limits, by amending or replacing the constitution, or by appealing to the courts for a reinterpretation of existing constitutional provisions (Carey 2003). However, not all presidents attempt to do so, and not all presidents who try succeed. If most government leaders in today's world are unlikely to remain president for life, it seems logical that presidents also care about what will happen to them after they leave office.

Scholars usually characterize this second motivation (beyond retaining office) as the desire to build a positive legacy (Moe and Howell 1999). Leaders make decisions while in office that are designed to establish a strong governing record, which could help them win other opportunities outside the presidency, such as academic positions, leadership of a corporation, a senior role in an international organization, or other honors or awards that would give them continued wealth, status, and influence outside the state. However, leaders might also make decisions to protect their interests after they depart office that do not involve building a positive legacy. For instance, incumbents might try to install loyal supporters in positions of control over the armed forces, state financial resources, and/or courts, weaken the legislature or judiciary's ability to investigate or prosecute political leaders, or simply embezzle as many funds as they can while in office to be able to fall back on when they leave.

We then have two goals that motivate government leaders, and two strategies to achieve those goals. Leaders might seek to maximize their chances of retaining office, and/or maximize their chances of surviving and enjoying other opportunities outside the state. To accomplish those goals, leaders might seek greater powers in order to get more things done, i.e. to enact their policy agenda, or they might seek greater powers in order to weaken rivals and protect their own interests. We can also think of the two strategies in terms of their intended effect on the leader's

opponents. On the one hand, leaders can deliver policies and programs that increase the *desire* to keep them in office and reward them afterwards. Or leaders can amass power in ways that reduce opponents' *ability* to remove or punish them.

Table 1 summarizes these goals and potential strategies. The cell in the upper left represents the policy decisions most commonly studied in both the literature on U.S. and comparative politics. That is, most studies focus on leaders' efforts to win reelection (or the next leadership selection contest), and assume that leaders seek to do so by accomplishing their policy agenda and delivering benefits to constituents in exchange for their support. The cell in the upper right is also frequently mentioned in studies of U.S. presidential power, since second-term presidents are not eligible for reelection and are generally assumed to be motivated primarily by their legacy. These studies also assume that leaders seek to secure their legacy by enacting a positive agenda, demonstrating accomplishments and building support for future rewards.

Table 1. Government leaders' motivations and strategies for use of power in office

<i>Strategy</i>	<i>Objective</i>	<i>Retain government office for as long as possible</i>	<i>Secure positive fate after departing office</i>
<i>Previous literature: Increase others' desire to retain/reward incumbent</i>		Deliver programs and services to constituents, win support for reelection	Build legacy of accomplishments, win support for futures rewards
<i>This project: Decrease others' ability to remove/sanction incumbent</i>		Amass power to block rival authorities' ability to impeach, dismiss, or demand resignation	Amass power to block rival authorities' ability to investigate or prosecute after removal

In this project, I am more concerned about the lower row of Table 1. That is, I am more interested in how leaders might seek to strengthen themselves *against* potential opponents and block others' *capacity* to challenge them, rather than incumbents' efforts to persuade others to support them. The cell in the lower left represents leaders' potential efforts to consolidate power in order to cling to office, which applies to cases in which leaders have some chance of using their control of the state to perpetuate their rule, and when removal from office is their greatest concern. The cell in the lower right reflects similar strategies to increase the executive's control of the state, but applies when incumbents are more concerned about blocking threats that might arise after they depart office, such as future prosecution, incarceration, or exile.

The distinction between the rows in Table 1 is more important for this project than the distinction between the columns. I have shown the two objectives here to fully contextualize this study in relation to prior research, since most prior studies focused on how leaders retain office, rather than what may happen to them afterwards. However, the strategies that leaders use to win support for reelection are often similar to the strategies leaders use to secure a positive legacy.

Likewise, the strategies leaders use to block threats of removal are often similar to the strategies they might use to block future threats of prosecution or imprisonment after they depart. While I focus on leaders' fears of what might happen to them after they leave, I do not assume that leaders react only in such a way that will protect them from prosecution later on. They might also react by clinging to office in order to put off that possible fate. The important distinction is that they are seeking power for its own sake, to strengthen their own control of the state at rivals' expense, in order to block threats (to their office or to their person once out of office) rather than win support for positive future opportunities.

Most political leaders are probably motivated by both goals; they would prefer to retain office as long as they can, but they also care about what will happen to them afterwards. Most political leaders also probably care at least to some extent both about enacting their policy agenda, and about protecting their own power and interests beyond simply ensuring that their agenda is carried out. Yet the two strategies can also conflict. If leaders appease supporters by sharing power and engaging in collaboration to improve policy outputs, they might inadvertently give rivals the power to challenge them. If leaders try to take away rivals' ability to challenge them, they might increase the desire to sanction them. Given these trade-offs, leaders might choose to focus more on one strategy or the other at different times. I turn next to how we might distinguish between these different strategies in leaders' observable actions.

2.4 How to tell if leaders are trying to get things done or seeking power alone?

If leaders might increase their use of decrees either to enact more urgently-needed policies or to protect their own interests in and out of office, how can we tell which is happening? Analyzing the total number of decrees is probably only useful if we assume a single primary motivation for the use of decrees in most contexts. When there are different possible ends leaders might be pursuing through decrees, the total number is probably insufficient. But we might be able to tell what they are trying to do by looking at the specific actions named in those decrees, and the organizations, programs, or other entities those actions target.

For instance, decrees that create new executive agencies or transfer resources to them are more likely to expand the power of the executive and enable a wider range of future executive actions. Decrees that impose regulations on private actors, restructure tax or customs duties, or approve one-time contracts or public works projects, may matter a great deal in terms of who gains or loses. But the latter external actions seem less likely to alter which internal government actors have the power to make future policy decisions.

The main decrees of interest are those that are likely to alter the internal allocation of government power, especially those that increase the president's control over decision-making and resources of the state. The actions I include are the creation of new offices or organizational units, the restructuring of existing offices or assignment of responsibilities or resources to them, the appointment of senior government officials, and the establishment of a state of emergency

(which usually entails special authority for executive agents to circumvent more deliberative processes). The main target entities that I focus on are executive offices, such as the presidency, cabinet ministries, and executive agencies that report to the president, such as the offices of the attorney general or comptroller general. I also look at decrees that target the legislature, judiciary, local government units, and other non-executive public offices, since decrees might be used to reorganize, weaken, or coopt other branches and levels of government as well.

I contrast these potential power-consolidating decrees with other decrees that do not involve clear changes to the internal governing apparatus. These include decrees that approve or execute one-time benefits or services, individual transactions like contracts or acquisitions, or revisions to the civil, penal, or tax code that involve regulations and restrictions on private activity. For decrees targeting entities outside the state, I include private corporations or industries, civil society organizations, and citizens groups, such as workers, youth, women, indigenous groups, or other protected classes.

Table 2 shows how the proposed types of decrees might fit into the cells from Table 1 (where I introduced leaders' different motivations and strategies for achieving them). The general idea is that leaders are more likely to take actions that deliver goods and services to private constituents, when they are focused on winning support for reelection or a positive legacy (i.e. the upper row, containing actions more often studied in previous literature). Leaders are more likely to take actions that restructure *internal* government institutions and decision-making authority when they are focused on protecting themselves and blocking challenges to their power or to their future survival (i.e. the lower row, containing the actions I focus on in this project). As discussed above, the two columns are very similar; the important distinction here is between the two rows. In Chapter 3, I define the coding scheme and methods used to extract specific actions and target entities from each decree, and aggregate them into event data for analysis.

Table 2. Decree actions related to leaders' different motivations and strategies

<i>Strategy</i>	<i>Objective</i>	<i>Retain government office for as long as possible</i>	<i>Secure positive fate after departing office</i>
<i>Previous literature: Increase others' desire to retain/reward incumbent</i>		Popular symbolic measures, programs and services targeting constituents	Lasting symbolic measures, programs and services with positive societal impact
<i>This project: Decrease others' ability to remove/sanction incumbent</i>		Create or reorganize electoral commissions, appoint loyal justices/commanders, assume greater power to block impeachment processes	Create or reorganize agencies, appoint loyal prosecutors/justices/police chiefs, weaken legislature, assume greater control of legal investigations and public funds

2.5 How to verify that these different actions align with different motivations?

We might also be able to tell what leaders are trying to do by analyzing which decrees they issue in the presence of other factors that reflect different motivations or pressures. In this part of the theoretical conceptualization, and in the associated empirical analysis in Chapter 4, I draw on existing explanatory factors from the previous literature on presidential power and executive orders. In particular, I use indicators of legislative opposition or gridlock, economic crises and natural disasters, and time left in office, which we would expect to correlate with increased use of executive decrees. Instead of verifying those variables' expected relationships to *total* decrees, however, I extend the established explanations to consider how they might reflect more specific motivations for the expansion or use of executive power.

The existing explanatory factors were drawn from studies that only implied one motivation for leaders' use of executive decrees: to get more policies enacted, whether because of more urgent demands or because of obstacles to enacting policies through other channels. Since the studies did not address the other potential motivation for additional decrees (e.g. to consolidate the leader's own power), the relationships between the existing explanatory factors and these two motivations are sometimes unclear. For instance, when the president controls fewer seats in the legislature, he/she might turn to decrees in order to get policies enacted that the legislature is refusing to pass. The president might also be wary, however, that an opposition-controlled legislature will be more likely to pursue impeachment and/or prosecution if the president does not bolster the executive branch's powers and weaken the legislature, also through decrees. In other words, the two explanations would be observationally equivalent, if we only look at the relationship between control of the legislature and executive decrees.

There might be related factors, however, that represent similar forms of legislative gridlock, but are less ambiguous about which motivation they should invoke. Highly fragmented parties also represent greater obstacles to enacting desired policies through the legislative bargaining process, since the president would need to coordinate bills with more political actors that lack a unified platform or inter-party discipline. However, high fragmentation – especially among opposition parties – should *reduce* the threat of early removal or prosecution of the president, since opponents would also have to coordinate a unified effort to challenge the president among themselves. Therefore, if we see certain decrees increase only when the opposition is unified – i.e. not when opposition party fragmentation is high, even if small opposition parties still occupy many seats in the legislature – we might interpret those decrees as less about overcoming legislative gridlock and more about the president's self-preservation.

I also analyze two types of external crises that might affect presidents' interests in different ways. Executive decrees are often explicitly authorized to enact emergency measures in the face of urgent economic or security crises, or natural disasters like earthquakes or floods. While some of the Andean states experienced insurgent violence during the period of study, it is much more difficult to identify reasonably exogenous measures of violent threats that could not

have been influenced by the president's own use of power. I instead turn to economic shocks and natural disasters as more exogenous crises, to which the president might react in different ways.

On the one hand, both economic and natural disasters might compel the president to issue more decrees to rapidly get things done and mitigate harm to constituents. On the other hand, presidents are more likely to fear threats to their own office in the face of economic downturns than natural disasters. Earthquakes and floods are unlikely to be blamed on the incumbent, at least in their immediate aftermath (although a poor government response might be criticized sometime later on). But for economic crises, even in the case of exogenous commodity price shocks that the national government cannot control, the president might face political blowback if the shocks are large enough to affect local industries, since the executive is usually considered responsible for managing the national economy (Kim 2014). If we observe decrees empowering public offices during economic downturns, but only observe decrees delivering temporary, externally-focused goods and services in the wake of natural disasters, this would support the interpretation that different motivations are associated with these different types of decrees.

Finally, we might observe differences in the types of decrees that leaders issue as they near the end of their time in office. This relationship is a bit tricky; at first glance we might expect leaders to be most concerned about their post-tenure fate when they get close to departing office. However, the strategies leaders pursue to protect themselves against future threats might change as they get close to departing, especially when they have very little chance of prolonging their tenure. As discussed earlier, leaders might respond to threats of post-tenure sanctions by consolidating power in ways that prolong their time in office, to avoid facing the punishment that might come afterwards. They might enable public offices to block legislative or judicial investigations that could lead to their removal, or put more loyalists into key cabinet and security offices in order to ensure that they do not collaborate with political movements to oust the incumbent. Leaders might also take steps to circumvent term limits, while they still have time, such as stacking an electoral commission or a judicial tribunal that might rule on their eligibility.

However, once a leader gets close to the end of a term with a clear constitutional ban on reelection, having managed to avoid early removal but having failed to secure a chance at a new term, it seems much less likely that the leader would continue to pursue lasting expansions to the power of the presidency, when those powers will most likely fall into a successor's hands. The leader might continue to weaken rival authorities in other ways, such as increasing constraints on the legislature or judiciary, or might even change course and begin enabling another center of power which the incumbent plans to move into him/herself after departing the presidency.

In other words, I expect leaders to consolidate power when they fear what might happen to them after they depart office, but not necessarily at the very end of their tenure. If leaders stand a chance of renewing their mandate, it might make sense to consolidate control of electoral agencies, security forces, and programmatic offices at the expense of rival authorities, in order to fend off challengers and secure a new term. But if leaders reach the final months of a fixed term with no legal path to retain office, in a modern context in which they are unlikely to be able to cling to power by force alone, I expect them to stop further empowering the executive offices

they are fairly certain to lose, and shift focus toward securing their future fate in other ways. If leaders refrain from issuing certain types of decrees toward the end of their term, but only when they are ineligible for reelection, this would suggest that those decrees represent lasting expansions of power, which outgoing leaders would not wish to hand over to their successors.

Table 3. Established covariates that might reveal leaders’ motivations and strategies

<i>Strategy Objective</i>	<i>Retain government office for as long as possible</i>	<i>Secure positive fate after departing office</i>
<i>Prior literature: Increase others’ desire to retain/reward incumbent</i>	<ul style="list-style-type: none"> • More decrees targeting private actors if legislative gridlock (i.e. governing party controls few seats, party fragmentation is high) • More decrees targeting affected areas/industries during any crisis (e.g. econ shocks or natural disasters) • More decrees targeting private actors toward end of term (whether incumbent can run for reelection or not) 	
<i>This project: Decrease others’ ability to remove/sanction incumbent</i>	<ul style="list-style-type: none"> • More decrees empowering executive if opposition is unified (i.e. low party fragmentation) • More decrees empowering executive during severe economic crises (but not during natural disasters) 	
	<ul style="list-style-type: none"> • More decrees empowering executive if leader expects to retain control longer (early in term or when up for reelection) 	<ul style="list-style-type: none"> • Fewer decrees empowering executive when leader is about to leave (at end of term with no chance of reelection)

Table 3 summarizes these expected relationships, showing how explanatory factors from the literature on presidential power can be applied to the same rows used in Tables 1 and 2. I expect leaders to issue more decrees enacting their external policy agenda (top half) when there is gridlock in the legislature from a large but fragmented opposition, but to issue more power-consolidating decrees (bottom half) when the opposition is unified. I also expect both types of decrees to increase in the face of economic shocks, but only external policy measures (top half) in response to natural disasters. In most cases, the ways in which I expect incumbents to use different types of decrees are the same whether they focus more on trying to retain office or on protecting their interests after they depart. However, in the case of decrees that empower the executive at the end of a presidential term, I expect incumbents’ behavior to diverge depending

on whether they have a chance of reelection, a distinction that might be especially helpful for verifying that certain types of decrees do represent lasting changes to executive power.

Using the additional intuition discussed here about how prior explanatory factors align with different potential motivations, yields a set of expected relationships between the established covariates and the specific types of decrees I introduce in this study. Testing for these expected relationships (in Chapter 4) should help to validate that certain actions taken through decrees represent specific motivations and strategies pursued by government leaders, distinguishing the motivations of interest in this study from other motivations assumed in previous studies. In other words, the analysis of these relationships in Chapter 4 will serve as a form of “construct validation” (Adcock and Collier 2001), ensuring that the measures of the dependent variable which I propose to use are associated with theoretically established explanations for the given phenomenon. A good measure of leaders’ efforts to consolidate power to protect their own interests and block rivals’ threats, should capture the expected relationships to the political and economic control variables shown in Table 3.

2.6 How do leaders view the risks of future punishment or reward?

In this section, I turn to the project’s new explanatory variable: leaders’ expectations of future punishment or reward. While previous studies have discussed when and how former leaders might be prosecuted for abuses of power, they haven’t defined what aspects of former leaders’ fates influence future leaders’ actions. In this section, I propose that what matters most to incumbents is not simply whether a predecessor was punished or not, but how the punishment fit the predecessors’ crimes, and how the incumbent believes he/she can best avoid a similar fate.

Why would leaders not always try to expand their own power? It seems unlikely that having more power would be a bad thing for a head of government while he/she remains in office. But what about afterwards: do leaders suffer worse long-term fates if they have concentrated too much power in the executive while they controlled it? There might be a variety of reasons why consolidating power could backfire during or after a leader’s departure from office. Opponents weakened or undermined by the executive might become more determined to remove the leader, and to do so before the end of the leader’s term, resorting to options like impeachment (or even force) that negatively affect the leader’s legacy. The opponents might also become more determined, or have more legitimate grounds, to prosecute the leader for misuse of power after removal from office.

So why would leaders pick a more contentious power-consolidation strategy over a more positive legacy-building strategy? Since leaders have different options for pursuing or protecting their future interests, their choice might depend on which strategy they believe is most likely to improve their fate, given the political environment they face. If leaders expect to be rewarded for a positive record and punished only when they misuse power, leaders should be more likely to focus on enacting their policy agenda, while refraining from assuming too much unilateral power

(i.e. staying in the upper row of Table 1). On the other hand, if leaders fear that a positive record won't guarantee their survival and expect to be prosecuted by rivals *regardless* of whether they commit real offenses, leaders might perceive that they have little to lose by amassing as much power as they can to try to block anticipated threats (i.e. moving to the lower row of Table 1).

An observable indicator of the political climate that incumbents will face is the example set by recent predecessors, i.e. the fates predecessors encountered after they left office. If a former leader is convicted of clear criminal acts undertaken while in office, such as ordering extrajudicial killings or embezzling funds, for which there is strong evidence and consensus on the acts' illegality, a successor might be less likely to engage in similar acts. However, if a former leader is prosecuted on more dubious charges and with weaker evidence, and a successor perceives that the former leader might be innocent of serious wrongdoing, this might encourage the successor to cling to power rather than exercise restraint.

What signs, then, is an incumbent looking for in predecessors' fates? It might matter less how many predecessors were punished, than what types of sanctions or rewards they faced, who instigated those events, and how much consensus surrounded their overall fate. The explanatory factors that I am interested in are difficult to pin down; it is not necessarily obvious what aspects of former leaders' fates will convince successors that their own best option is to restrain themselves, or to cling to power. I propose that what matters is essentially whether former leaders' fates seem to have been objectively determined based on their prior actions in office. That is, does each individual leader's fate match observers' expectations of whether he/she *should* have been punished or rewarded? If the leader's fate does not match expectations of what should have happened, this might suggest to successors that their own fates will be less predictable or certain, reducing any deterrent effect on subsequent behavior.

In the following subsections, I describe what constitutes a sanction or reward in more detail, and discuss aspects that might indicate the objectivity or predictability of those punishments or rewards. Specific variables are operationalized in Chapters 5 and 6; this discussion is intended to introduce the main concepts that I will attempt to measure later on.

2.6.1 Types of sanctions, reprieves, and rewards

The Archigos dataset mentioned in section 2.2 provides a starting point for defining leaders' post-tenure fates (Goemans et al 2009). The dataset encodes whether a former leader was executed, imprisoned, or exiled during the first year after departing office. There are many other potential events, however, that might punish or reward a former leader and influence future leaders' decisions. For instance, even if a former leader was never convicted, facing a congressional inquiry into wrongdoing, criminal charges, or standing trial could undermine the leader's reputation and legacy, and cost him/her real opportunities in private sector or civil society organizations. Outstanding charges might also drive the former leader into exile, and compel him/her to seek protection from powerful allies out of fear of extradition.

Consider the case of Bolivia's Gonzalo Sánchez de Lozada, who has spent over a decade in exile in the U.S. fighting efforts to punish him for the massacre of protestors opposing his unpopular economic reforms in 2003. As described in the introductory chapter, Sánchez de Lozada's legal battles have been driven by campaigns by victims' families and human rights groups, which helped secure preliminary investigations by the Attorney General's office, a congressional vote to authorize proceedings against the former president and cabinet ministers, a formal trial by the Supreme Court with Sánchez de Lozada in absentia, and government requests for Sánchez's extradition from the U.S., which are still pending.⁶

Cases like Sánchez's are complex, involving many steps by multiple actors, sometimes with disagreement and advances followed by setbacks, in the path toward punishment for abuses of power. I consider *sanctions* to include any formal attempts to punish former leaders, including investigation, arrest, criminal charges, requested or obtained extradition, trial and/or conviction. These steps may be undertaken by the executive, legislative, or judicial branches of government, or by foreign governments or international organizations. I will initially consider all of these attempted punishments as sanctions in general, and will then break them down by specific actors involved, as well as by different types of events or stages of the legal process.

If we look at partial or attempted steps toward punishing a former leader, we might also need to look at steps that would serve as setbacks in that process. For instance, a leader might be arrested and detained, but might also be released from detention. A successor government might file criminal charges against a former leader, but a court might dismiss those charges or find insufficient evidence to proceed to trial. Government authorities might request a former leader's extradition from another country, but the foreign country might deny the request. And even if a former leader is put on trial, he/she might be acquitted of the crime. These setbacks would not necessarily be considered rewards; no leader hopes to be remembered for dismissed charges. I consider these setbacks to be *reprieves*, which usually come after some form of sanction.

Political leaders might also care about different types of *rewards*, since their ultimate objective is not only to survive physically, but to retain other sources of wealth, status, and influence as former statesmen and societal leaders. Former presidents might secure new public sector jobs; many former Latin American presidents later served in the legislature, in cabinet positions, or as ambassadors to other countries. Former leaders might also secure prestigious positions or awards from private corporations, universities, or other civil society organizations.

⁶ Mabel Azcui, "El Parlamento de Bolivia autoriza un juicio político contra el ex presidente Sánchez de Lozada," *El País*, October 15, 2004. https://elpais.com/diario/2004/10/15/internacional/1097791227_850215.html

Gretchen Gordon, "The Trial of Gonzalo Sanchez de Lozada," The Democracy Center, May 19, 2009. <https://democracyctr.org/the-trial-of-gonzalo-sanchez-de-lozada/>

Telesur, "US Accepts Extradition Request for Former Bolivian President," February 17, 2016.

<https://www.telesurtv.net/english/news/US-Accepts-Extradition-Request-for-Former-Bolivian-President-20160217-0001.html>

Trial International, "Gonzalo Sanchez de Lozada," January 7, 2014, Last modified 10/1/2017, Accessed 3/5/2018. <https://trialinternational.org/latest-post/gonzalo-sanchez-de-lozada/>

It probably matters to future leaders whether their predecessors were sanctioned or rewarded, and in what ways, but I anticipate that the relationships between general sanctions or rewards and future leaders' actions will be somewhat unclear. The main idea in this theory is that future leaders are trying to discern how objective or how disputable and manipulable their future fate will be. That is, do they think they might be able to get away with abusing power, and are they confident that they will be spared punishment if they do not? I expect that the general category of post-tenure events that will have the clearest relationship to power-consolidating decrees will be *reprieves*, since reprieves involve backtracking on a sanction and therefore signal at least some inconsistency or uncertainty of outcomes.

Table 4 summarizes the hypothesized relationships between post-tenure fate events and subsequent leaders' actions, for the three general categories on their own. Theoretically, some of the relationships might go either way; there are intuitive reasons why post-tenure sanctions or even rewards might encourage some incumbents to issue more power-consolidating decrees, and encourage other incumbents to issue fewer. For instance, sanctions might deter subsequent leaders from amassing power and making themselves culpable for government abuses, or might instead encourage leaders to cling to power to protect themselves.

Table 4. Hypothesized relationships for post-tenure fates and successors' actions

Predecessors' Fates	Incumbents' Consolidation of Power
Sanction	↑ or ↓
Reward	↑ or ↓
Reprieve	↑

Rewards might also go either way, if future rewards can be bought with corruption and control of the state, or if rewards are instead tied to power-sharing, responsible leadership and a positive external policy legacy. I do expect reprieves to more clearly lead to increased power consolidation, since reprieves should begin to tell us more about former leaders' punishments, suggesting that those punishments are contentious or uncertain and that justice can be escaped.

2.6.2 Types of actors initiating a sanction or reward

The next step toward judging whether a former leader's fate was objectively determined or politically motivated might be to identify who was behind the event – that is, what type of actor sought to punish or reward the incumbent. This is the main distinction mentioned in previous studies that qualitatively investigated politicized prosecution. Conaghan (2012) identifies different roles in the prosecutions of former presidents played by partisan actors (i.e.

political elites in the executive or legislature) and “accountability-seeking actors” (i.e. nonpartisan institutions like the judiciary or civil society organizations).

In the cases in this study, the main actors involved in formally punishing former leaders were all public sector entities, primarily the courts, government prosecutors and law enforcement agencies, and to a lesser extent legislative commissions. There has been extensive research and debate about the extent to which judiciaries are really politically neutral or nonpartisan, even in advanced democracies (Domingo 2010; Weiden 2011). It might be difficult to attribute a degree of objectivity to the instigators of a former leader’s punishment, simply based on which branch of government they occupy. However, government leaders might have different concerns about how to protect themselves against threats from different branches of government, so I consider the possibility that the type of actor behind the punishment matters to subsequent leaders.

Another type of actor who might be involved in former leaders’ sanctions, reprieves, or rewards is a foreign government (or foreign corporation or civil society organization in the case of some rewards). While foreign actors might be somewhat removed from domestic political power struggles, foreign governments have their own agendas, which might not align with domestic law or national interests. Rewards bestowed by foreign actors might even have perverse consequences, if former leaders are perceived to be loyal to foreign interests rather than serving their own people. However, foreign authorities’ deference to or concurrence with domestic legal determinations might add weight to those proceedings, such as when foreign governments agree to extradite a former leader back to his/her home country. In those cases, foreign authorities might seem less likely to be persuaded by domestic political vendettas and more likely to require strong evidence, at least that the charges are legitimate cause for extradition, in order to cooperate in forcibly returning a former leader to face justice at home.

Finally, while public authorities are typically responsible for pursuing justice for criminal acts, private organizations are often involved in rewarding former leaders after they depart office. It might matter a great deal to incumbents whether they expect to be able to survive and enjoy other opportunities for wealth and status outside the state, or whether their only options to protect their interests lie within the political system. If former leaders continue to serve in other government offices, incumbents might prioritize entrenching their own personal control as much as possible while they control the presidency, which might help them secure future jobs in other parts of the state, or continue to wield influence from those other offices. However, if former leaders more often retire to private sector positions, incumbents might be more willing to exercise restraint while in office and focus policies outward to build a positive legacy.

Table 5 summarizes the hypothesized relationships between former leaders’ sanctions, reprieves, or rewards by different instigating actors, and subsequent leaders’ actions. This set of hypotheses shows a number of clearer expected relationships than in Table 4, although some relationships still remain ambiguous.

Table 5. Hypotheses for post-tenure fates by instigating actor

Predecessors' Fates	Incumbents' Consolidation of Power
Sanction by legislature	↑
Sanction by government	↑
Sanction by judiciary	↑ or ↓
Sanction by foreign actors	↓
<hr/>	
Reprieve by legislature	↑
Reprieve by government	↑
Reprieve by judiciary	↑
Reprieve by foreign actors	↑
<hr/>	
Reward in public sector	↑
Reward in private sector	↓
Reward by political party	↑
Reward by foreign actor	↑ or ↓

For sanctions, since previous literature suggests that investigations or criminal charges by more partisan legislative and executive actors might appear more politicized, I anticipate that sanctions by those actors are more likely to lead subsequent presidents to issue more power-consolidating decrees. Since the literature suggests that judiciaries are more objective and politically neutral, I include a weak hypothesis that sanctions by judiciaries might lead to more restraint (or fewer power-consolidating decrees) by subsequent presidents, although judiciaries may also be politicized, so I allow for potential relationships in both directions. Sanctions by foreign actors are the only ones that I do not expect to appear politicized, at least in the case of domestic legal proceedings in which foreign sanctions only take the form of compliance with extradition requests, so the hypothesis in the fourth row of Table 5 only appears in the column on the right (i.e. fewer power-consolidating decrees).

For reprieves, I expect much stronger relationships in the first column of Table 5; that is, I expect all categories of reprieves to lead subsequent presidents to consolidate power. For rewards, I expect that former leaders retiring to private or civil society positions will *discourage* power consolidation and encourage incumbents to focus on building a positive legacy. However, former leaders' return to public office or political party leadership might signal to subsequent leaders that their best options for future rewards reside within the state, encouraging them to build as many levers of personal influence as they can while they control the presidency. Foreign rewards might also signal to incumbents that they can escape justice if they use their power to amass resources and secure foreign allies. The hypothesized relationships for rewards therefore depend strongly on which actor (or sector) provides the reward.

2.6.3 Different types of sanctions, stages in the legal process

Even after separating out different instigating actors, there remains more ambiguity about the consequences of sanctioning former leaders, and sanctions are the events that I am most interested in, given the current debates about whether and how to prosecute former leaders for abuses of power. Not all sanctions are the same, and we might find more precise relationships if we distinguish specific types of punishment events. In particular, incumbents might interpret and react to former leaders' fates differently as cases against them proceed through different stages of the legal process. Early steps like investigations, arrests, and pretrial detention might appear more arbitrary and politically motivated than later steps like a formal trial and conviction. These distinctions also relate to different instigating actors, since initial investigations are usually undertaken by legislative offices or executive agencies, while a conviction – handed down by a court – usually comes after those earlier steps have been completed.

Sanctions at different stages of the legal process might also capture some cumulative information about what preceding events should have already happened. This relates to the reasons why I expect reprieves to be so controversial and to lead to consolidations of power. Reprieves imply that some attempted sanction has already occurred, and that the reprieve signals a reversal of that process. In contrast, convictions can only occur after investigations have been concluded, charges have been filed, and a trial has been held. In this way, convictions may signal a series of *consistent* events, in which multiple actors in the legal system have arrived at a degree of consensus as to the former leader's guilt. Convictions generally entail a higher burden of proof than initial arrest or pretrial detention, and even if the verdict is still influenced to some extent by politics, the momentum or consensus that builds throughout the legal process may indicate greater objectivity and certainty with regard to the ultimate fate, convincing successors that justice will be more difficult to manipulate or escape.

Table 6. Hypotheses for specific types or stages of sanctions

Predecessors' Fates	Incumbents' Consolidation of Power
Exile	↑
Detained pretrial	↑
Formal criminal trial	↓
Convicted of crimes	↓

Table 6 shows the hypothesized relationships for these more specific types of sanctioning events at the beginning and end of the legal process. For more preliminary efforts to sanction former leaders, especially pretrial detention, I expect subsequent leaders to respond by issuing more power-consolidating decrees to protect themselves from similar actions. I also include exile as a specific type of sanction, which might occur early in response to a growing threat of possible

sanctions, since leaders usually flee abroad to escape justice before they have been arrested and tried. As I discuss in Chapter 6, the potentially preemptive nature of self-exile can make it difficult to distinguish when a former leader is living abroad by choice or out of fear of punishment, so I shade that hypothesis as weaker than the others.

For the final steps in the legal process, especially formal trial and conviction of criminal acts, I expect subsequent leaders to respond by *reducing* power-consolidating decrees. These are the events in which I expect there to be the greatest consensus as to the former leaders' guilt, and the strongest signal to successors that their punishment will be deliberated and objective, such that they stand a good chance of being spared if they do not abuse power, but that they cannot escape justice if they do.

2.6.4 Consistency or predictability of former leaders' fates

The last section suggested that certain post-tenure events might imply that preceding steps have already been taken, which were either consistent (in the case of convictions) or inconsistent (in the case of reprieves), although the hypotheses are stated only in terms of one observed event at a time. To more directly analyze the objectivity of a former leader's sanctions or rewards, we might need to look at the distribution of all post-tenure events that happened to the same leader over time. For instance, if a former leader faced both attempted punishments and rewards from different groups or at different times after departing office, this might indicate greater contention or uncertainty about his/her fate.

It might also matter how well former leaders' fates align with their actual decisions in office. For instance, if leaders are consistently punished who abused power in certain ways, successors should find it easier to predict how their own actions may lead to punishment, and act accordingly by refraining from those actions. However, if former leaders are punished who governed in very different ways, including some who did not appear to overstep their authority or otherwise give clear cause for sanction, successors might perceive that their fates are more arbitrary or politically motivated. In the latter case, successors might perceive that they have less to lose by amassing power and using it to pursue and protect their own interests, since they may not be spared punishment even if they refrain from doing so. They might perceive that their best chance of surviving after office is to secure enough personal influence and resources to beat back any legal challenges by rivals, manipulating the legal system to escape punishment.

Table 7 summarizes the hypothesized relationships between these concepts of consistency or predictability across multiple post-tenure events and subsequent leaders' actions, stated in abstract terms. I expect these relationships to be the strongest, in principle, since they capture the underlying expectations behind incumbents' decisions about their use of power. However, the consistency or predictability of leaders' fates may be more difficult to measure than the type of event or actor involved in an individual sanction or reward. In Chapter 6, I define more specific calculations used to measure the predicted probability of a former leader's

post-tenure fate, to test whether incumbents react differently to former leaders’ punishments that seem expected or consistent with the former leader’s actions, versus punishments that are unexpected and do not seem to be objectively determined based on real abuses of power.

Table 7. Hypotheses for consistency or predictability of sanctions

Predecessors’ Fates	Incumbents’ Power Seeking
Former fates consistent with expectations , leaders punished who should be	↓
Former fates inconsistent or unexpected , leaders punished whose actions do not suggest they should be	↑

2.6.5 Interactions between near-term pressures and post-tenure fates

As noted in Section 2.3, leaders might face trade-offs between focusing on the policy goals that they have been empowered to pursue, and protecting themselves against threats of early removal or retribution by rivals if they have to depart office. How leaders allocate their priorities might depend on how concerned they are about their record or legacy, on the one hand, and how much they stand to lose later on if they do not take enough action to protect themselves, on the other. As a final set of hypotheses, I take the near-term control variables that might indicate a leader’s risk of being removed early, and test whether their interactions with former leaders’ fates have a stronger effect on incumbents’ use of decrees. In other words, former leaders’ fates might have heterogeneous effects on current leaders’ actions, depending on how imminently the incumbents fear losing office. Or, viewed the other way, leaders might react to near-term crises that increase their likelihood of removal from office in different ways, depending on their expectations of what will happen to them afterwards.

For instance, leaders facing economic crises might seek greater powers both to get things done and to protect themselves from removal. But they might limit themselves to delivering more services targeting private actors in order to directly mitigate the crisis, if they are more concerned about their legacy than the threat of future punishment. They might instead seek more permanent reconfigurations of internal government authority, if they are more concerned about protecting themselves against rivals who would use the crisis to pursue their removal and elimination from the political sphere. In sum, we might observe a difference in decrees when economic crises coincide with the politicized punishment of former leaders. Table 8 summarizes this final hypothesis, proposing that power-consolidating decrees may increase most strongly when economic crises coincide with sanctions of former leaders.

Table 8. Hypothesized interactions between crises and post-tenure fates

Predecessors' Fates	Incumbents' Power Seeking
Economic shocks x sanction of former leaders	↑
Economic shocks (but no sanction)	↓

These interaction tests might be useful if we are especially concerned about how leaders react to crises, and how to ensure that they respond well. Leaders will sometimes have to face trying times, in which if they do not act quickly and effectively, they might lose their jobs. It would not be surprising to discover that leaders seek to expand executive authority in times of crisis, but it would also not be a very interesting conclusion. However, if we find that leaders react *differently* to crises depending on some other more controllable factor, practitioners might be able to encourage leaders to better serve the country's long-term interests even when leaders face short-term risks. Policy practitioners might not be able to reduce or prevent the crises themselves, but might be able to shape other conditions that could discourage leaders' short-sighted reactions, such as institutionalizing assurances of objective and depoliticized justice for former leaders accused of misuses of power.

2.7 Summary of theoretical discussion

In sum, in this chapter, I have reviewed existing theories from previous literature on presidential power and the use of executive orders, and the motivations and consequences of prosecuting former government leaders. I presented a theory that brings these two issues together, identifying more specific aspects of each that may be important to distinguishing leaders' motivations and strategies, in the face of different pressures or threats. I introduced the main concepts I plan to use to distinguish executive decrees that represent efforts to consolidate power, and suggested ways to use established explanations from previous literature on executive authority to validate these decree categories, by testing for expected relationships to certain control variables that reflect different motivations leaders might have for issuing decrees.

I then introduced the main concepts behind the independent variable in this study, leaders' expectations of future punishment or reward, based on their predecessors' post-tenure fates. I defined a set of variables and hypothesized relationships in several stages, beginning with the most observable general categories of sanctions, reprieves, and rewards, and the potential implications of different instigating actors. I then introduced several additional ways that sanctions in particular might be broken down in order to get at the more latent aspects of objectivity or predictability at the heart of the proposed theory. I also suggested that leaders

might respond differently to post-tenure sanctions in the presence of near-term crises like economic shocks, introducing hypotheses about the interaction between these factors.

In the remaining chapters, I define the empirical strategy used to test this theory, discuss data sources and methods for measuring the concepts introduced here, and present statistical results for the hypothesis tests.

Chapter 3: Methodology for Measuring Executive Action Through Decrees

3.1 Overview of methodological decisions

In designing any study, we face a series of high level decisions about how to get from research questions to meaningful, compelling answers. What is the exact phenomenon we wish to study? What sources of information are available to observe it? How can we use those sources to measure the phenomenon of interest? How will we use the measures we have derived to test hypotheses about the causes and/or consequences of the phenomenon of interest, and how do the decisions we have made about data and measures affect the results we are able to obtain? These decisions are driven by the substantive motivations of the research, as well as by available resources and practical constraints.

In this chapter, I present two methodological contributions to the study of how presidents expand and use their authority. First, I make a case for using government documents – in this case executive decrees – to directly observe policy decisions and institutional changes. Government action and political behavior have more often been observed through secondary news reporting, or through researcher-initiated collection efforts like surveys and polls. Yet official public records constitute rich sources of authoritative information about government action, if we can process and encode that information into analyzable data.

Second, I apply and test several computational approaches to automatically categorize decrees based on their main actions and target entities. The focus on certain actions and targets is driven by the theoretical discussion in Chapter 2, regarding how presidents might use different types of decrees to pursue different objectives, including those they use to strengthen their own executive offices. I develop a process for assigning action and target labels to each decree, making use of existing natural language processing (NLP) tools like part-of-speech tagging and dependency parsing, as well as machine learning classifiers, to construct an effective document encoding system with limited project-specific resources. At the end of this chapter, I evaluate the accuracy of the resulting data in relation to human coding.

3.2 Sources of information on government action

Researchers studying government action and political institutions often collect data from secondary news media, surveys and polls, or abstract indicators compiled and coded by experts. For the Polity IV indicator for “constraints on the chief executive,” coders give each country an annual score based on how the different branches of government are selected, who initiates legislation and other policy proposals, and how much power each branch has to appoint and/or

veto actions by the others (Marshall et al 2017). This composite indicator represents a high level of abstraction, using expert judgment to combine information about constitutional provisions and observed tendencies into a single seven-level score. Because of its high-level composition, the scores rarely change, and do not capture individual events or disaggregate different types of action, which might have different implications for governance and political outcomes.

Data-driven approaches that make use of large volumes of raw text have become more common in recent years, in part due to the explosion of information available in digital news and other online media (Schrodt 2006; Grimmer and Stewart 2013). These data sources are often too large, and the events or topics of interest too sparse, to be hand coded, requiring some form of automated classification or information extraction to measure variables of interest. With regard to government actors, scholars often focus text-as-data efforts on government actions directed outward, such as violent conflict between states, which are frequently observable in secondary news reporting (Bond et al 2003; Schrodt 2006; Raleigh et al 2010). News reports tend to be informative and straightforward, structured with the most important details first, and to be useful for monitoring crises and detecting major new developments (Tanev et al 2008).

3.2.1 Shortcomings of news reporting and advantages of public records

However, news media pose challenges to measuring events consistently over time. News reporting is often redundant, so that in order to accurately count specific events, researchers must deconflict reports about the same event. News reporting is also influenced by many factors other than whether the events occurred, such as the publication's resource constraints, reporters' access to participants, the shifting interests of the target audience, and the editors' or publishers' business objectives (Kepplinger 2002; Althaus et al 2011; Weidmann 2015). In terms of scope, news reports are not designed to account for every minor policy decision, public works project, or change to legal rules, nor every single battle in a conflict (Ortiz et al 2005). Since news is designed to sell, it is more useful for studying large-scale attention-grabbing events like new wars or regime changes, rather than everyday political decisions and interactions.

There are many important research questions, however, that revolve around the day-to-day business of government and the evolution of public offices and authorities. The best or only source of those day-to-day activities may be official government records. Governments are increasingly making their records publicly available in digital archives, including legislation, executive decrees, and other documents related to policy-making processes and debates (Cardie and Wilkerson 2008). These documents offer new opportunities to study government authority and action in more systematic and detailed ways. Text analysis tools have been used to identify topics and sentiment in legislation and congressional debates (Thomas et al 2006; Purpura and Hillard 2006; Gerrish and Blei 2012; Grimmer and Stewart 2013), and to extract party names and other key details from court rulings to assist legal researchers in finding relevant prior cases (Brüninghaus and Ashley 2001; Jackson et al 2003; Cheng et al 2009).

There is much more that we can learn from government records; previous efforts only scratch the surface of the information they contain, and new methods for processing text in general offer additional options to explore for government documents. There are also many advantages to using official records rather than secondary news and social media, when analyzing political events, which have been under-explored. First, laws and decrees are primary source documents that themselves enact the policy decisions or institutional changes they report. This means that each law or decree constitutes a separate action, which do not need to be deconflicted for redundancy, and researchers can assemble a reasonably complete set of relevant decisions enacted through a particular channel, if a consistent archive is maintained.

The language used in government documents is also more formal and consistent than news reports, since they are designed to be authoritative, rather than attention-grabbing or easy to skim. Laws and decrees use exact legal terminology, with official organization names written out in full, and the same verb phrase regularly used to represent the same type of policy action. For instance, in the executive decrees in this project’s dataset, the verbs *nominar* (“to name”) and *designar* (“to designate”) are consistently used for official appointments, whereas news reports might use a wider variety of colloquial terms like “to pick” or “to tap” an individual for a job. There are also fewer components that need to be identified in the text for each event, because the agent or “doer” can usually be interpreted as the authority enacting the law or decree. Researchers may only need to find the main verb and the main object (but not the subject) of that verb, turning questions of “**who did what to whom**” into “**what was done to whom**”.

3.2.2 Challenges to analyzing government documents

However, researchers still face considerable challenges when seeking to collect and process a large volume of full text government documents, especially in developing countries. Despite progress in the availability of public records, there are still considerable differences in quality, completeness, and usability across different archives. There are often multiple archives for the same type of document in each country, which do not have identical contents. Decrees might be published in an official gazette or daily registry, a subset might be available on the president’s website, or that of a cabinet office dedicated to public records, and additional collections might be available as part of the legislature’s archives or maintained by an office of the judiciary. In Peru, decrees appear in the official gazette (*Diario Oficial*) *El Peruano*⁷, the Digital Archive of Legislation of Peru⁸ maintained by the Congress, and the Peruvian System of Legal Information maintained by the Ministry of Justice and Human Rights⁹.

⁷ <http://www.elperuano.com.pe/>

⁸ <http://www.leyes.congreso.gob.pe/>

⁹ <http://spij.minjus.gob.pe/>

These archives differ in terms of years covered, search functionality, formats, and information included about each document. Laws and decrees are considered public domain in many countries, and transparency laws often require government records to be made publicly available.¹⁰ Yet to cover the costs of archival services and manage server load, some databases charge fees and/or impose download limits, even while the same documents are available for free or without limit on another website. As with news media, norms are still developing regarding database access for those seeking to crawl, scrape, or otherwise mine archives for research purposes, especially in countries outside the United States (Truyens and Van Eecke 2014).

Finally, it is more common to find lists of metadata (e.g. titles, dates, sponsors, status) than full text files in these archives, especially going back more than a few years. When historic full text documents are available, they are often scanned images that are not machine readable. Optical character recognition (OCR) software for non-English text still produces many errors, especially when used on grainy images. Even when machine readable, full text laws and decrees are also difficult to parse and interpret in a fully automated system. Legal norms often contain substantial non-operational preamble language, and are sometimes annotated with legislative histories, alternative terms, or other explanatory notes throughout, which do not follow a standard format across jurisdictions or different archives.

Despite these challenges, government documents offer tremendous potential value for studying the complex institutions, decisions, and interactions that define public policy and produce governance outcomes. The opportunities and challenges discussed above all contributed to the decisions I made regarding the document collection and processing used in this project, as described below. These resources represent an underutilized source of information that can enable us to tackle difficult questions in ways that may not be feasible, or might be more biased or noisy, when working with secondary news reporting or other media alone. Public records are useful not only for narrow applications in legal research, where these records have traditionally been used, but for much broader questions about political, social, and economic systems and developments in which governments play important roles.

3.2.3 Documents chosen for this project

I have collected information on almost 75,000 executive decrees issued in the five Andean countries over the past one to three decades (exact years and documents per country are summarized in Chapter 4). Most countries have a single type of executive decree, but I included any relevant subtypes wherever the constitution grants multiple forms of executive decree-making authority. This means that for Peru, I include the three main types of executive decrees (*decretos supremos*, *decretos legislativos*, and *decretos de urgencia*) to maximize coverage of the various policy decisions that presidents make through decrees, and to make the data

¹⁰ For instance, Peru's LAW No. 27806 of 2003, "Law on Transparency and Access to Public Information," http://www.peru.gob.pe/normas/docs/LEY_27806.pdf

comparable to the other countries. The current Constitution of Bolivia also contains multiple provisions for decree-making authority, some of which refer to *decretos supremos* and others of which simply refer to the power “to decree” (*decretar*) certain policy decisions, which are labeled as *decretos presidenciales* in the official gazette.¹¹ I include both in the dataset.

I focus on decrees, rather than legislation or other government records, because it is easier to attribute decrees to a particular political actor and test theories about that actor’s incentives and constraints. Bills and laws arise out of complex bargaining processes among multiple actors, including both the executive and legislative branches, whereas decrees are unilaterally promulgated by the president’s office. Decrees might be drafted by subordinate offices, such as the cabinet ministry responsible for a specific policy area, but decree authority resides in the president, who formally signs them (even if co-signed by relevant ministers). Controversial decrees, even those drafted by ministries, are often attributed to the president in news media and opposition critiques.

Due to the limited availability of machine-readable full text documents, I have chosen to use the titles of decrees, which are most widely available across countries, archives, and years. The titles also tend to have similar linguistic structure even across countries, centered around a recognizable policy action phrase and an organization or population which the law or decree governs. These titles can be thought of as combining some of the brevity and directness of news reporting with the authority and formality of official records. They are designed to clearly identify the decree by the gist of its main provision(s), in correct legal terms and with enough detail to distinguish it from other similar documents, while fitting on a few lines. Example decree titles are shown in Table 9 below.

Titles might be inadequate for more detailed information extraction, such as when seeking to map the history of specific sub-law provisions. But titles generally suffice to identify the main actions and target entities that I seek to measure for this project. I conducted a brief verification exercise, hand-coding a random sample of 100 full-text decrees available in Peru, after hand coding their titles only, using the project coding scheme defined below. The full-text and title-based labels were in agreement over 90% of the time. Only six of the 100 decrees contained enough additional information in the full text document to identify a different type of target entity (as defined below), and in only one of the 100 decrees did the main provisions in the body of the text indicate a different category of main action (or type of decision being made) than was indicated in the title. Of these coding differences, only one would have affected the data used in the project, since the other differences occurred between categories that were not selected to be used as the main dependent variable in this study.

¹¹ *Constitución Política del Estado Plurinacional de Bolivia*. 2009. Asamblea Constituyente. La Paz, Bolivia.

Table 9. Examples of decree titles in dataset

	<i>Title (original)</i>	<i>Title (manually translated)</i>
Ecuador, 3/8/2007	Reforma al ERJAFE y modificación de ciertos ministerios y secretarías	Reform of ERJAFE (Statute of the Legal Administrative Regime of the Government Executive) and modification of certain ministries and secretariats.
Ecuador, 8/19/2010	Nómbrese al Dr. Pedro Dávila Guevara, Gobernador de la Provincia de Imbabura, al aceptarse renuncia en tal cargo, a Sr. Luis Salazar Buitrón.	Appoint Dr. Pedro Dávila Guevara, Governor of the Province of Imbabura, upon acceptance of the resignation of Mr. Luis Salazar Buitrón.
Ecuador, 7/29/2014	Ratificase el Protocolo de Montevideo sobre compromiso con la democracia en el Mercosur.	Ratify the Montevideo Protocol regarding commitment to democracy in Mercosur.
Peru, 9/29/2011	Autorizan Transferencia de Partidas a favor de los Pliegos Ministerio de Defensa y Ministerio del Interior en el Presupuesto del Sector Público para el Año Fiscal 2011.	Authorize transfer of funding allocation in favor of the Ministry of Defense and Ministry of Interior in the public sector budget for fiscal year 2011.
Peru, 7/6/2003	Establecen Límites Máximos Permisibles de Radiaciones No Ionizantes en Telecomunicaciones	Establish maximum permissible limits for non-ionizing radiation in telecommunications.
Peru, 9/12/2007	Declaran el domingo 21 de octubre de 2007 “Día del Censo Nacional” para efectos de los Censos Nacionales: XI de Población y VI de Vivienda.	Declare Sunday, October 21, 2007 “National Census Day” for the purpose of the National Censuses of Population (XI) and Housing (VI).

3.3 Established methods for extracting actions and entities from text

To turn narrative records into comparable units of measurement, researchers need to encode information from the text into quantities or categories of interest. Social scientists have traditionally analyzed the contents of documents by hand, which enables them to identify complex and nuanced social concepts. However, hand coding is labor-intensive and highly subjective, which limits transparency and replicability for others in the research community, and may lead to inconsistency even within the same project (Mikhaylov et al. 2012). Because of these costs, social scientists have begun to use more computational methods for automatically coding and classifying text documents into data for analysis (Cardie and Wilkerson 2008).

In political science research, the most common approach to automatically categorizing text utilizes the frequency of words appearing in each document, regardless of grammar or word order, i.e. “bag-of-words” (BOG) techniques (Hopkins and King 2010; Grimmer and Stewart

2013; Biagioli et al 2005). These word frequencies are often used to assign a label or category to each document as a whole, describing some characteristic or perspective of the author, such as political ideology or sentiment toward the document's theme (Grimmer and Stewart 2013).

For known categories, researchers hand label a set of training examples, then train “supervised” machine learning models to predict which combinations of words should be assigned to which categories (Biagioli et al 2005; Thomas et al 2006; Purpura and Hillard 2006). There are many supervised classification algorithms and popular ready-to-use tools available today, but these classifiers can only be used to assign known labels, and require large quantities of pre-labeled training documents to achieve accurate results (Grimmer and Stewart 2013). These classifiers also preserve whatever biases the pre-labeled documents contain.

Alternatively, if researchers want to learn unknown categories, “unsupervised” or inductive approaches may be used to look for new patterns in unlabeled documents. A widely-used method is Latent Dirichlet Allocation (LDA) for topic modeling (Grimmer 2010; Gerrish and Blei 2012; Nguyen et al 2015b; Roberts et al 2014). Topic modeling is also typically performed using bag-of-words features, so that the learned topics are represented by distributions of words. In this way, bag-of-words features are often used to reveal documents' themes, sentiments or tones, but they provide insufficient information to indicate which elements in the text play which functional roles.

Going beyond document-level word frequencies, a less common but growing approach to text-based data is automated event extraction, using more complex processes to parse and extract “who did what to whom”. Traditional efforts rely on keyword searches and pattern matching, using dictionaries of named entities and verb phrases. In political science, the most common use of event extraction has been in English-language studies of violent conflict (King and Lowe 2003; Schrodtt 2006). Legal scholars have developed similar systems, using shallow parsing and manually constructed dictionaries and phrase patterns, to automate the indexing of relevant entities and provisions in case law (Brüninghaus and Ashley 2001; Jackson et al 2003; Cheng et al 2009). These systems tend to be designed and built wholesale for each project, requiring relatively large teams and considerable resources to implement.

In recent years, open source NLP tools have made the task of event extraction more accessible. These tools include *part-of-speech taggers*, which tag words as verbs, nouns, prepositions, etc., and *syntactic parsers*, which label words' grammatical relations like the subject, direct and indirect objects of a given verb (Klein and Manning 2003; de Marneffe et al. 2006). Instead of searching for fixed-order phrases or sentence patterns, researchers can construct logical rules, using automatically labeled parts of speech and grammatical relations, to identify which actors played which roles in events (Schrodtt 2014). For instance, a researcher might look for any form of the verb *attack*, and assign the roles of *attacker* to its subject and *victim* to its direct object. Researchers have also used supervised machine learning to assign semantic roles, by first annotating training corpora at the word level (Gildea and Jurafsky 2002), although such annotation is especially labor-intensive (Kim et al 2008; Caselli et al 2011).

One of the most prominent and long-lived projects that reflects these methodological developments is the succession of systems that began with the Kansas Event Data System (KEDS) (2006), and later spawned TABARI and PETRARCH as next-generation automated event coding systems (Schrodt 2014). KEDS provided for automated event coding based on manually constructed rules, patterns, and entity dictionaries. TABARI and PETRARCH made incrementally greater use of shallow and then deeper syntactic parsing, while still relying on different types of structured pattern recognition (Schrodt 2014). KEDS, TABARI and PETRARCH are designed to extract events in a specific ontology: the Conflict and Management Event Observations (CAMEO) coding scheme. CAMEO is an expert-designed taxonomy for automatically generating data on political conflict events, with origins in earlier coding schemes for interstate wars and diplomacy, to which it added more post-Cold War phenomena like ethnic and low-intensity conflict involving sub-state actors (Schrodt 2006). These are valuable and efficient resources for extracting known event templates within a major area of political research, but are still difficult to replicate in new languages and domains.

Computational linguists have also begun to experiment with unsupervised or inductive approaches to learning event templates, which do not require hand-written patterns or hand-labeled documents. Some researchers have used pipeline approaches in which they separately cluster verbs into event types and those verbs' noun objects into event roles (Chambers and Jurafsky 2011; Ahn 2017). Others have used probabilistic generative models to jointly learn event types and entity roles (Chamber 2013; Cheung et al 2013; Nguyen et al 2015a; Sha et al 2016). Like topic modeling, these inductive methods often produce unstable results and are not yet very effective at inducing the same types of events or roles that researchers might choose to specify; the best accuracy scores are usually below 50% when compared to hand-labeled test documents (i.e. better than random guesses for more than two labels, but producing more noise than correctly identified events for social scientific study). There are also not yet available off-the-shelf tools for social scientists to use these methods for new empirical applications.

3.4 Project strategy for document classification

The task of automatically encoding text into comparable measures for analysis involves several decisions: 1) what the goal is, in terms of the categories or labels we are seeking to encode; 2) what information (features) we can use from the original text to determine those categories or labels, and 3) what encoding process to use to transform the input features into the output labels for our data. These decisions involve a number of trade-offs, including what type and how much training data is needed, how much effort is required to implement the encoding process, and how well the resulting data captures the concepts of interest for the research goals.

Since I am using document sources and studying policy actions for which there are not yet established best practices, I have chosen to compare multiple methodological options for each of the questions above. I have selected several algorithms used to classify and extract

information from text records, including a rule-based pattern matching system and probabilistic machine learning classifiers. I have also chosen to compare different types of input features to the classifiers – “bag-of-words” term frequencies versus more structured features informed by grammatical relations – and to evaluate the models’ accuracy for labeling both a high-level categories and more fine-grained subcategories. The chosen approaches are summarized in Table 10 and explained in the sections below.

Table 10. Summary of Methodological Options Selected for Comparison

Step 1. Output Labels		
	pros	cons
<i>Document labels</i>	Easier to hand-label training data, good for topics/tones	Does not indicate which components fill which roles
<i>Phrase labels</i>	Enables extraction of info on event roles and entity relations	More labor to annotate training docs or write extraction rules
<i>Hybrid: several labels per doc</i>	One main action label per document, One target entity label per document	
<i>Few high-level categories</i>	Simpler classification task, need fewer training docs	Might oversimplify, categories too abstract to learn or be useful
<i>More low-level subcategories</i>	Preserves more distinctions, might matter for research goals	Requires more training docs to cover all categories
Step 2. Input Features		
	pros	cons
<i>Bag-of-words</i> (doc term freqs)	Easy to extract, existing tools, little preprocessing required	Discards word order, grammar which might be important
<i>Structured features</i> (main verbs and noun objects)	Enables extraction of info on events/relations, may produce more accurate doc-level classification too	More effort to extract, requires parsing (existing tools, but <100% accurate), more complex features will be more sparse
Step 3. Encoding Process		
	pros	cons
<i>Rule-based pattern matching</i>	Requires less labor to get easy cases right, might get decent scores w/ fairly simple rules, very fast	Hard to specify less common/subtler cases, tail takes more effort w/ diminishing returns, might be an upper limit on potential accuracy
<i>Supervised machine learning</i>	Should be capable of highest accuracy with enough training data	Need lots of labeled training data, can only reproduce human coding, biases in input replicated in output
<i>Unsupervised learning</i>	Needs no labeled training data, can learn new clusters or categories	Hard to evaluate, open-ended possibilities, much more fine-tuning required to produce useful output

I haven't listed the steps in the order in which they emerge in the final process, i.e. inputs to encoding model to outputs, but in the order in which I defined and implemented them (and in which it makes the most sense to introduce them). For any applied project, one first needs to decide what the objective is, i.e. what categories one is trying to encode, then what raw information one can use to assign those categories, and finally what algorithm is most appropriate to get from the raw information to the chosen categories.

I include in Table 10 several options in gray text that I have decided not to implement in this project, since they do not fit the substantive objectives or are not yet sufficiently well developed for practical application (as in the case of unsupervised approaches to event extraction). The options that I have chosen to implement and compare are shaded in blue. The table shows advantages and disadvantages for each of the options listed, to motivate the methodological choices I have made, and to show the trade-offs among the options that I have chosen to compare in the accuracy tests at the end of this chapter. Again, each of these options and the ones I have chosen to implement and test are discussed in the sections below.

Step 1. Project coding scheme (i.e. desired output labels)

As introduced in the theoretical framework in Chapter 2, I am interested in identifying what type of decision or change the president enacted through each decree, and who the primary target of that action was. The desired information involves a bit more than an overall document theme, but I do not need to identify every aspect of each decree event, such as the duration or target location. I have chosen a coding scheme that balances the research objectives with feasibility constraints; the resulting scheme falls between a single document-level category (e.g. a topic) and a full event schema (e.g. with an agent, target, time, place, and instrument or means). I seek to identify two pieces of information for each document: a *main action* and a *target entity*, each of which are grouped into a limited set of categories.

Table 11 shows the chosen categories. In them, I have preserved two levels of granularity to explore what works best: a higher level with only a few broad action and entity groups, and a lower level with more subtypes. The higher-level categories are easier to classify automatically, based on a finite number of labeled examples, but the lower-level categories preserve more information about distinctions between decrees that might be important to the theory. The choice of coding scheme should take into consideration both the feasibility of reliable classification and the usefulness of the resulting categories in empirical analysis. In Chapter 4, I discuss how I select which of these categories to use in the main hypothesis tests, using expected relationships to control variables discussed in Chapter 2, to validate which categories of decrees are issued when previous theories suggest that we should expect presidents to consolidate power.

Table 11. Coding Scheme with main action and target entity categories

Category	Subcategories	Example Terms
Action		
<i>enable</i>	<i>enable: empower,</i> <i>enable: appoint,</i> <i>enable: finance,</i> <i>enable: modify</i>	create, authorize ... appoint, nominate ... transfer, fund ...
<i>regulate</i>	<i>regulate: restrict,</i> <i>regulate: rules,</i> <i>regulate: modify</i>	require, prohibit ... regulation to implement a law ...
<i>other</i>	<i>other: enact,</i> <i>other: modify</i>	execute, distribute, ratify ...
Target		
<i>gov (exec)</i>	<i>gov: executive</i>	presidency, cabinet, ministry ...
<i>public (non-exec)</i>	<i>public: legislature,</i> <i>public: judiciary,</i> <i>public: local,</i> <i>public: personnel,</i> <i>public: institutes,</i> <i>public: other</i>	legislature, congress ... court, tribunal ... municipality, province ... civil servants, diplomats ... (usually public) schools, hospitals ...
<i>private</i>	<i>private: business,</i> <i>private: other</i>	corporation, industry ... youth, workers, voters ...

Step 2. Text preprocessing and feature extraction (i.e. input features)

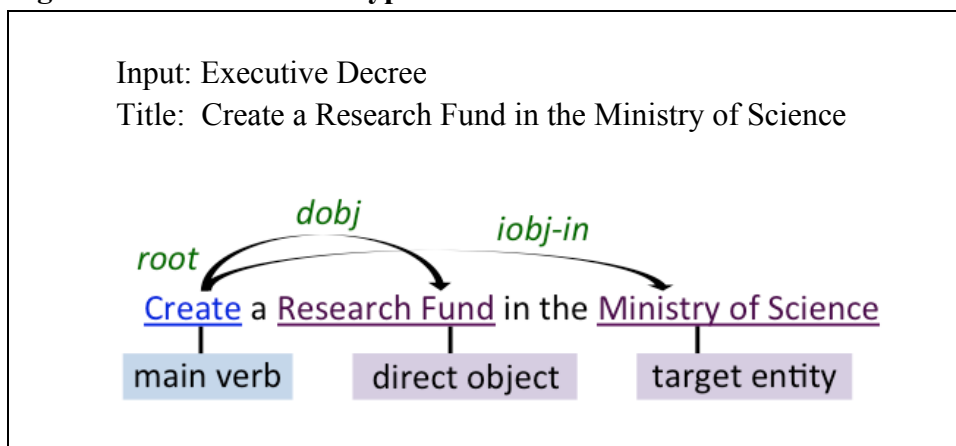
For a computer to be able to automatically categorize or label documents, we can't feed it raw text, because the machine would not know how to make mathematical decisions using human language. Instead, we have to feed the machine a numeric representation from the text that it should use in its calculations; then the machine can learn which combinations of numeric values or "features" should be assigned to which labels. Transforming raw text into numeric features is often considered a step in "preprocessing" the text, to make it ready for classification or other forms of analysis. We have many options for what features to use, which differ in terms of how easy they are to extract from the text and how well they capture the information we need.

Bag-of-words features: The most common features used in text analysis are counts of words that appear in each document, regardless of grammar or word order, i.e. "bag-of-words" approaches. Bag-of-words features require very little preprocessing of the raw text, beyond tokenization (i.e. segmenting words), which can be done using whitespace for Spanish as well as English text. The main step is to construct a document-term matrix, i.e. a vector of term counts for each document. I use the top 1000 most frequent terms in the corpus as the columns in this

matrix. Following common practice, I “lemmatize” words to their root form (i.e. converting verbs to the infinitive, nouns to singular male form) and weight the resulting vectors by inverse document frequency, to emphasize terms that are more distinct to specific documents.

Structured features: To extract features that capture more linguistic structure from the text, I also perform several preprocessing steps involving NLP tools and resources. These are the same steps I use to prepare the text for the rule-based pattern-matching system as well. The first step is to apply part-of-speech tagging and dependency parsing, for which I use the Stanford CoreNLP toolkit (Manning et al. 2014), version 3.7.0, using the Spanish language model for the PCFG parser (Klein and Manning 2003; Spanish models by Jon Gauthier). These tools produce a tree structure for each sentence rooted at its main verb, as shown in Figure 3.

Figure 3. Parse tree for a typical decree title



After parsing, I identify each document's main verb, using the highest-level active verb in the parse tree. This verb is usually labeled by the parser as the sentence “root”, or attached to an enabling verb like “propose” or “declare”, as in “Propose to *create* a new office ...”. I use my own list of Spanish enabling verbs, based on inspection of decrees in the dataset, to distinguish common procedural verbs from the active verbs that indicate what the decree actually does. Actions may also be stated in nominal form, as in “Propose the *creation* of a new office ...”. I use a Spanish spell-checking dictionary and morphological rules to convert noun forms of verbs back to their verb infinitive (Rodríguez and Carretero 1996). Throughout this chapter, I refer to these action terms as “verbs” for readability, although they might appear in noun form as well.

After identifying the main active verb in each decree title, I identify that verb’s noun objects, to use in labeling the decree’s target entity. The targets might be direct objects, as in “Create a *new agency*...”, or indirect objects that appear after a preposition like “to”, “of”, or “in”, as in “Transfer funds *to the ministry*...”. In the latter case, the direct object (“funds”) is a general resource, which might be transferred to different types of organizations. The direct object clarifies what action is being taken (i.e. a financing operation), but not which entity is being

funded. The indirect object following “to” (the ministry) is the organizational target of the action. For simplicity, in this preprocessing step, I extract all direct and indirect objects of the main action verbs to include as potential target entities in the decree’s structured features.

Finally, identifying specific verbs and noun objects allows us to incorporate additional information about these terms from external references, which is especially useful for grouping the terms into patterns in the rule-based system. For instance, it is much more time consuming to define every possible phrase that might refer to an attack literally, like “Russia attacked China” or “North Korea attacked Japan”, than it is to specify a template like “[Country A] attacked [Country B]” and then label noun phrases in the corpus as country names so that they can be matched to the looser template. This is similar to the way that the KEDS and TABARI verb patterns and actor dictionaries are constructed (Schrodt 2014). However, defining a dictionary of all possible terms in the corpus with their associated entity types would also be labor-intensive, and may be unnecessary, when lists of known entities like country names already exist.

Instead, I use WordNet, a lexical database of over 100,000 word senses with definitions and hierarchical relationships. A Spanish language version of the Multilingual Central Repository (MCR) (Gonzalez-Agirre et al. 2012) is available through the Open Multilingual WordNet interface in the Python Natural Language Toolkit (NLTK) (Bond and Paik 2012). WordNet includes many geographic and political unit names, as well as common nouns for government branches, agencies and offices, programs and services, and other concepts useful for distinguishing the target of a policy decision or institutional change. Each word sense in WordNet includes a list of higher-level parent terms (i.e. “hypernyms”), which help to identify what type of entity it is. For instance, the term “congress” is listed under “legislature”, which is listed under “assembly”, which is a type of the more general term “group”.

In this project, I use WordNet in two ways. For the machine learning classifiers, I look up each decree’s main verb and noun objects in WordNet and include the word’s hypernyms in the decree’s structured features. In this way, the structured features include more information than the “bag-of-words” models, taking into account external knowledge about words that play certain grammatical roles in the text. Figure 4 shows an example of two feature vectors – bag-of-words and more structured features – for the same raw text.

Figure 4. Bag-of-words vs. structured features for example text

<i>Document text</i>	Bag-of-words	Structured features
“Create a Research Fund in the Ministry of Science”	create: 1, research: 1, fund: 1, ministry: 1, science: 1 ...	verb=create: 1, object=fund: 1, object=ministry: 1, object hypernym=government department: 1, object hypernym=administrative unit: 1 ...
<i>Note: other features that do not appear in this decree title would have values of zero</i>		

For the rule-based pattern-matching system, as I developed the project’s coding scheme, I also created a list of WordNet hypernyms that should be assigned to each of the chosen categories. This enables me to look up the target words from each decree in WordNet and use the hypernyms listed there to identify the right category in the coding scheme, rather than listing the actual words exhaustively in each project-specific category. (The pattern-matching rules are defined in more detail in Step 3b below.)

Step 3. Classification algorithms (i.e. encoding process)

As mentioned above and shown in the third part of Table 10, I have chosen to compare two approaches to identify each decree title’s main action and target entity (i.e. to turn the input features into the output labels). The first is the more common approach to document classification today, using “supervised” machine learning, in which a machine is fed a set of input features and a correct output label for each document in a labeled training set, and learns a function for which combinations of input features should be assigned which label. Supervised machine learning is popular because there are easy-to-use off-the-shelf tools to do so, one only needs to supply the labeled training data. However, these models require large amounts of good training data with a clear and consistent coding scheme in order to learn how to accurately replicate human classification. The amount of training data needed increases when the coding scheme includes many categories, some of which appear infrequently in the training set.

The second approach is to use rule-based pattern matching, to deterministically assign certain statements in the text to certain categories, rather than using a learned probabilistic model. It can be easier to write a few rules to tell a machine directly how to identify the most common patterns in the text, than to provide enough training examples for the machine to figure those patterns out on its own. For instance, if I know that decrees should be labeled as *enabling* the *government executive* whenever the root verb is “create” or “establish” and its direct object is the name of an executive agency, it might be much more efficient to define a rule for this pattern, rather than label enough decrees with variations of this pattern until the machine learns that the *enabling* category goes with those specific verbs and the *government executive* label goes with the agency names (and not some other words that also appear in the labeled examples).

As will be shown in the evaluation section that follows, this approach can perform better at classifying a larger number of more infrequent categories, when there is insufficient labeled training data for a computer to learn all of the correct combinations of features. However, the process is deterministic, and will only be able to label the patterns we specify, so there may be an upper limit to how well it can identify every possible configuration of verbs and nouns that reflect a given action in the text. It becomes more difficult to write rules for subtler and more nuanced cases, especially if those cases are more rare and harder to find by manual inspection. These are some of the reasons why machine learning classifiers have become more popular, and why they are often capable of the highest accuracy, once enough documents have been labeled to

thoroughly cover a manageable number of categories. Below, I describe the specific classification models that I will use, and then evaluate them in comparison.

Step 3.1 Supervised machine learning classifiers

For supervised machine learning, I have chosen existing document-level classifiers available in the Python scientific computing package *scikit-learn*. I test Naïve Bayes, support vector machines (using a linear kernel), logistic regression, and random forests. These algorithms mainly constitute linear classifiers (generalized or log-linear in the case of logistic regression), with the exception of random forests, which are compilations of decision trees. Each algorithm uses a different approach to transfer input features into output labels, updating the parameters of that transformation to best fit the correctly labeled documents supplied as training data. Some models may work better for certain types of data or features than for others.

For instance, a linear classifier will use a linear transformation of the input features to calculate the probability of each output category, such as a weighted sum with different learned weights for more and less important features, akin to linear regression. The Naïve Bayes algorithm assumes that the input features are independent, so that with bag-of-words features, the presence of a word like “transfer” in the document might contribute a certain amount of probability to the output classes, and the presence of another word like “funds” might contribute a separate amount, but their interaction is not accounted for.

Decision trees instead arrive at labels through a sequence of decisions based on individual features, such that a path down the tree represents the interaction of multiple feature values. For instance, a decision tree might first look for the word “transfer” in a document, and then only if that word is present, look for the word “funds”, assigning the label of a *financing* action if both are present at the same time. Because decision trees assign different labels to specific combinations of features, they have a tendency to overfit, i.e. to memorize certain idiosyncratic patterns that appear in the training data. Random forests overcome this limitation by constructing multiple decision trees using different features from the overall training set.

Since these off-the-shelf classifiers operate at the document level (or on a single vector of features for each labeled text), I run all classifiers twice on the full set of documents, once to assign a main action label to each decree and once to assign a target entity label. I repeat this process using the bag-of-words features and the more structured features as inputs. For bag-of-words models, the same vectors are used to classify both actions and target entities. For the structured features, slightly different vectors are used that roughly paralleling the information I use in the rule-based approach below. I use both main verbs and their noun objects to classify action types, since “establish a fund” is an enabling action, while “establish a regulation” is a restricting action. I only include the noun objects (and their hypernyms) as input features in the structured feature models when assigning target entity categories.

Step 3.2 Rule-based information extraction

As mentioned above, I have also constructed a rule-based pattern-matching system, to more directly assign specific verbs and noun objects to action and target categories, rather than relying on probabilistic or learned functions to predict the intended action or target based on document-level word frequencies. The pattern-matching system is in some ways a more traditional approach to automatic event coding, searching the text for certain verbs and noun objects that represent the actions or events the researcher is looking for. In other ways, however, this process involves more advanced text processing than typical document-level classifiers.

In its simplest form, rule-based pattern matching might be done with keyword or keyphrase searches. Today, however, we can do much more than just look for identical fixed-order phrases or sentences in the text we seek to label. As discussed in Step 2, I use grammatical parsing and lexical resources to identify more information about specific words that appear in specific positions in the text. In this version of Step 3, I use that information to specify which configurations of main verbs and their noun objects should be assigned to which action or target labels in the coding scheme. In other words, I look for patterns not in the raw text, but in the elements of the grammatical parse tree, and use not only exactly matching words but groups of words with certain hypernyms in WordNet, to produce more flexible and efficient coding.

For each label in the project coding scheme, I define one or more rules that contain conditions a decree title must meet to receive that label. Each rule includes a condition for the document's main action verb, plus potential conditions for that verb's objects. An object condition specifies a particular dependency relation (e.g. the verb's direct object or an indirect object) and a WordNet hypernym for the noun that should appear in that grammatical position. If a decree title has a main verb matching the permitted verbs for a given rule, and that verb also has a noun phrase with the right hypernym in the right object position, the document is assigned the action category corresponding to that rule. Rules for target entity categories follow the same format, but any verb is allowed, only the noun object conditions matter.

Figure 5. Example rules to match verb-object clauses to action or target labels

category	subcategory	=	verb	+	dependency relation	:	hyponym
Action							
<i>enable</i>	<i>finance: enable</i>	=	"financiar"	+	[any object]		
<i>enable</i>	<i>finance: enable</i>	=	"transferir"	+	<i>dobj</i> : [assets]		
Target							
<i>gov</i>	<i>gov: executive</i>	=	[any verb]	+	<i>iobj</i> : [government department]		
<i>private</i>	<i>private: business</i>	=	[any verb]	+	<i>iobj</i> : [industry corporation]		

Figure 5 shows examples of the rules used. For instance, if a decree title contains the verb “*financiar*” (to finance), the decree will be labeled with the main action category of *enable: finance*. That verb is sufficiently clear on its own that no object condition is needed. Alternatively, if a decree title contains the verb “*transferir*” (to transfer), this could refer to several types of transfers, so the decree will only be labeled *enable: finance* if the verb also has a direct object that falls under the hypernym for “assets” in WordNet (such as words for money, funds, resources, etc.). For the target entity labels, any main verb will do. If a decree title has a main verb with any object (including indirect objects following “to”, “for”, “in”, etc.) that falls under the WordNet hypernym for “government department” (such as a cabinet ministry, agency, treasury, etc.), the decree will be labeled with the target category for *gov: executive*.

3.5 Evaluation of accuracy in relation to human coding

The most common way to evaluate automated document classification is to test a model’s accuracy against a baseline of hand-coded examples. This type of test is appropriate if our goal is to train a machine to replicate the same coding decisions that we have made by hand, meaning that it makes the most sense for supervised classification. Accuracy tests are straightforward to implement and interpret; a machine that performs well may be very valuable for automating previously labor-intensive processes.

I test all models on a random sample of 1300 decrees from the project dataset, which I hand labeled using the project coding scheme. I began by labeling decrees from Peru – the largest sample in the dataset – to develop the project’s coding scheme and initial rules, then refined the categories and rules as I incorporated decrees from the other four countries. In order to ensure that I had enough labeled documents for accurate classification and proper evaluation, I ultimately labeled 500 decree titles from Peru plus 200 decree titles from each of the remaining four countries, resulting in a total of 1300. This hand-coded sample is larger than often used for document classification, since the dataset covers five countries and there are still modest differences in policy decisions and phrases used in decrees across even nearby countries.

To ensure that no model has overfit (or simply memorized) the training examples, we need to conduct accuracy tests using a “held-out” test set. For the supervised machine learning models, I use 10-fold cross validation, iteratively reserving a different 10th of the labeled documents, training on the remaining 9/10^{ths}, testing on the held-out 10th, then averaging the scores from all 10 folds. Each fold has the same number of documents (the labeled set is divisible by 10). For the rule-based extraction system, I wrote the rules while hand-coding the first 100 decrees from Peru, refined the rules on the next 50 Peruvian decrees, then made minor revisions when labeling the first 100 decrees from each of the remaining countries. I froze the coding scheme and rules while hand coding the remaining 350 out of 500 decrees in Peru and the second 100 decrees in the other four countries. I evaluate the accuracy of the rule-based extraction system on the latter 750 decrees that were labeled after finalizing the rules.

3.5.1 Results of accuracy tests, comparing models

I sought to test all combinations of the methodological options summarized in Table 10 that made sense, varying the input features, the classification algorithm, and the granularity of output labels. In Table 12, I report scores for precision (correct guesses out of total guesses), recall (correct guesses out of total true labels), and F1-scores (harmonic mean of precision and recall), reported in percentages for readability. The top half of Table 12 shows scores for labeling decrees with the few high-level categories in the project coding scheme, while the bottom half of the table shows scores for labeling the more fine-grained subcategories. The highest scores at each of the two levels are highlighted in gray.

For both levels of categories, the support vector machine (SVM) model with a linear kernel, trained on bag-of-words features (scores on the left in Table 12), performs best, achieving F1-scores of 82.9% for the few high-level categories and 75% for the more fine-grained subcategories. The random forest and logistic regression models perform similarly on the high-level categories, achieving about 80-81% F1-scores, and about 5% lower on the more fine-grained subcategories, achieving about 69-70%. The Naive Bayes models perform much worse, though they vary more with the choice of input features.

Table 12. Accuracy against human coding, comparing classification algorithms

Approach	Classification algorithm	Bag-of-words			Structured features		
		P	R	F1±95%CI	P	R	F1±95%CI
<i>Fewer high-level categories</i>							
Supervised machine learning	Naïve Bayes	46.5	46.5	46.5 ±1.8	62.6	62.6	62.6 ±1.8
	Random Forest	80.5	80.5	80.5 ±1.8	79.4	79.4	79.4 ±1.0
	Logistic Regression	81.3	81.3	81.3 ±1.4	79.5	79.5	79.5 ±2.3
	Support Vector Machine	82.8	82.9	82.9 ±1.5	76.6	76.7	76.7 ±1.6
Deterministic	Rule-based pattern matching	-	-	-	77.4	80.9	79.1 ±1.9
<i>More low-level subcategories</i>							
Supervised machine learning	Naïve Bayes	25.3	25.3	25.3 ±1.9	49.0	49.0	49.0 ±2.2
	Random Forest	70.6	70.6	70.6 ±1.7	69.2	69.3	69.3 ±1.5
	Logistic Regression	69.5	69.5	69.5 ±2.6	70.3	70.4	70.3 ±2.1
	Support Vector Machine	74.9	75.0	75.0 ±1.9	66.3	66.3	66.3 ±2.0
Deterministic	Rule-based pattern matching	-	-	-	70.9	74.9	72.9 ±2.3

P = precision, R = recall, F1 = F1 score

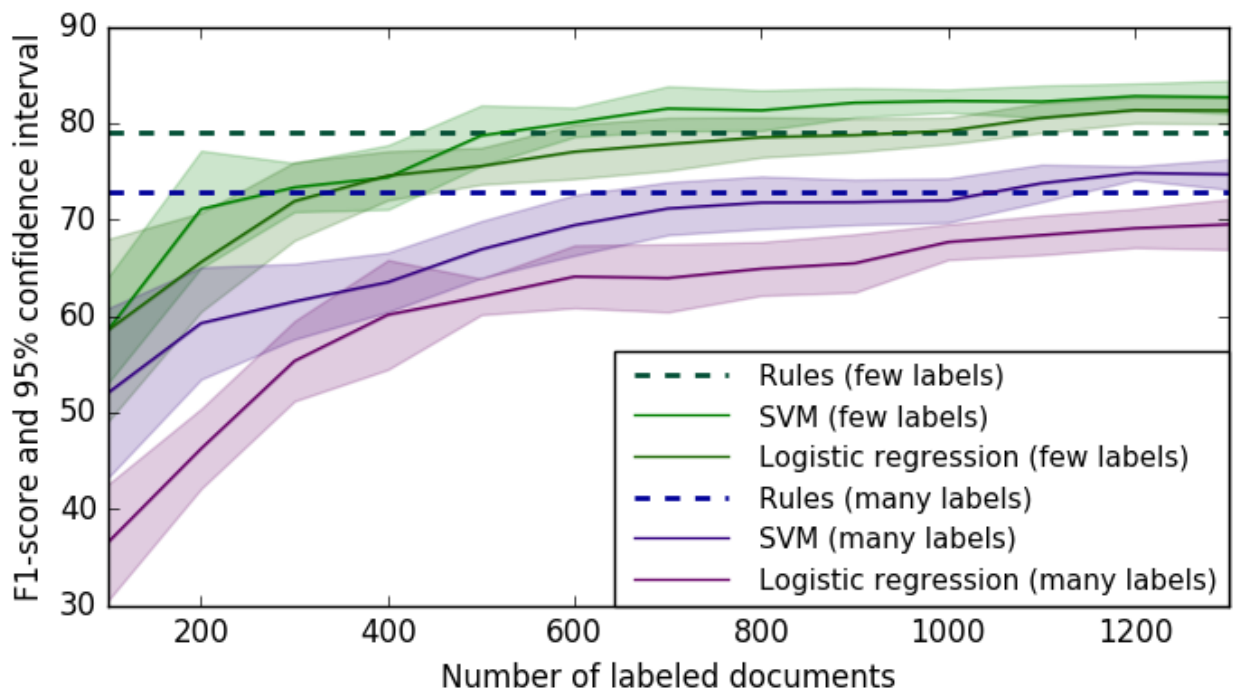
In particular, Naive Bayes is the only model to perform considerably better when using the more structured features (on the right) than when using simple bag-of-words term

frequencies (on the left). Its scores jump from 46.5% to 62.6% on the high-level categories and from 25.3% to 49% on the low-level categories, when using structured features instead of bag-of-words. Naive Bayes is a fairly simple linear algorithm that does not account for interactions between features unless explicitly included, so providing specific verb and noun features appearing in certain positions in the text may help to better identify label nuances. The structured features do not appear to help the other classifiers, which seem to have inferred at least as much useful information from the bag-of-words vectors alone.

The rule-based pattern-matching system (in the bottom row of each half of the table) does not quite match the performance of the SVM model at either level of the coding scheme, but it does beat the other machine learning classifiers for labeling the more fine-grained subcategories. The rule-based system achieves an F1-score of 72.9% in the lower half of the table, while the random forest and logistic regression models only achieve scores of about 69-70%. This suggests that even with 1300 hand-labeled documents, there still isn't enough training data for most of the machine learning classifiers to learn how to accurately label all of the coding scheme's more fine-grained subcategories.

Figure 6 shows how the accuracy of the top two machine learning classifiers increases with the size of the training corpus. (The size refers to the total labeled documents used; in each test, 9/10^{ths} of those documents were used for training and 1/10th was held out for testing.)

Figure 6. Accuracy of supervised learning by size of training corpus



For the few high-level categories (the green lines in the upper part of the graph), the machine learning classifiers reach the performance of the rule-based system at around 500 labeled documents (for the support vector machine) and 900 documents (for logistic regression). At 1300 labeled documents, both classifiers surpass the rule-based system with statistical significance, using the 95% confidence interval calculated from variation across test folds. However, for the more fine-grained categories (the purple lines in the lower part of the graph), even with all 1300 labeled documents, the rule-based system still outperforms the logistic regression classifier. The support vector machine doesn't surpass the performance of the rule-based system until there are at least 1000 labeled documents for training and testing.

Machine learning classifiers need enough labeled examples to distinguish each category included in the coding scheme, so more training documents would be needed to improve the classifiers' scores for the more fine-grained categories. However, the rule-based system specifies exact combinations of verbs and nouns to be used for each fine-grained category, and I simply aggregate these rules to assign the corresponding high-level categories. As suggested in some literature on automatic information extraction, rule-based systems often have lower up-front costs (Chiticariu et al 2013); it is easier to explicitly instruct the computer how to find the most common examples of what one is looking for. However, it is much harder to continue manually specifying rules for more rare patterns or more complicated distinctions, and there is often a point at which it is easier to simply label examples of subtler or more nuanced cases and then let the computer figure out which combinations of words are more likely to go with which labels.

With enough training data, as long as the number of training examples is high compared to the number of labels or categories that the machine needs to learn, the machine learning classifiers may be able to achieve higher accuracy than rule-based classifiers on their own. If researchers need to preserve more labels or categories in the data encoded for analysis, however, and the input text is already well structured, deterministic rules might still be the most efficient and effective means of achieving the desired results.

3.5.2 Results for ensemble classification

The fact that of the bag-of-words machine learning classifiers performed best overall (at least for labeling the higher-level categories) does not mean that the other options developed and tested for this project (more structured features and rule-based patterns) are not useful as well. Computational social scientists have suggested that the best approach to document classification is to combine multiple classifiers into an “ensemble” learning process (Grimmer and Stewart 2013). Table 13 shows precision, recall, and F1-scores for several combinations of the classifiers tested individually above. For each ensemble, I run all of the included classifiers on each decree and assign the action and target label that received the most classifier “votes”.

Table 13. Accuracy against human coding, ensemble classification

Machine learning models	Bag-of-words models only			Bag-of-words structured feature models			+ Bag-of-words structured feat models + rule-based system		
	P	R	F1±95%CI	P	R	F1±95%CI	P	R	F1±95%CI
<i>Fewer high-level categories</i>									
<i>Best individual model</i>	82.8	82.9	82.9 ±1.5	-	-	-	-	-	-
NB, RF, LR, SVM	81.5	81.6	81.6 ±1.9	83.0	83.0	83.0 ±1.7	84.1	84.2	84.2 ±1.5
RF, LR, SVM	82.7	82.7	82.7 ±2.0	83.4	83.4	83.4 ±1.4	85.2	85.3	85.3 ±1.3
<i>More low-level subcategories</i>									
<i>Best individual model</i>	74.9	75.0	75.0 ±1.9	-	-	-	-	-	-
NB, RF, LR, SVM	72.4	72.4	72.4 ±1.9	72.6	72.7	72.6 ±1.7	74.6	74.7	74.6 ±1.5
RF, LR, SVM	73.1	73.2	73.1 ±1.8	75.7	75.8	75.7 ±1.3	76.2	76.1	76.1 ±2.1

P = precision, R = recall, F1 = F1 score

In Table 13, the top row in each section includes all four machine learning classifiers (including Naïve Bayes and support vector machines), while the bottom row only includes the two machine learning classifiers that performed best individually (logistic regression and random forests). In both upper and lower halves of the table, the columns on the left show scores for the machine learning classifiers with bag-of-words features alone, the columns in the middle add machine learning models trained with structured features, and the columns on the right include the rule-based pattern matching system as well.

Perhaps unsurprisingly, including all possible classifiers – even Naïve Bayes – performs worse than using the better individual classifier on its own. In the second row of Table 13, the ensemble with four bag-of-words classifier models only achieves an F1-score of 81.6%, while the best individual classifier achieved 82.9%. Combining just the three high-performing classifiers (in the third row) just about matches the top performing individual classifier. However, the ensembles perform incrementally better from left to right, when including not only the bag-of-words models (on the left), but also models trained on the more structured features (in the center), as well as the rule-based system (on the right). The same is true in the lower half of Table 13 for the fine-grained subcategories.

This suggests that in an ensemble approach, including multiple models that use different types of information in different ways works better than only using fairly similar classifiers, even if those classifiers performed best individually. Greater variation in the included models might help take advantage of their different strengths, filling in or compensating for the cases in which the other classifiers are weaker, to achieve the best possible accuracy scores overall.

For the data to be used in the remainder of this project, I use the ensemble in the second row and farthest right column of Table 13, combining the rule-based system with the top two supervised machine learning classifiers, each trained separately on bag-of-words and more structured features, for a total of 5 models incorporated into the votes. I also limit the resulting labels to one action and one target per decree, using the rule-based classifier to break ties. This combination achieves an 85% F1-score on the higher-level categories and 76% on the more fine-grained categories (the two top scores highlighted in gray). In the next chapter, I use decrees labeled with the lower-level categories to provide descriptive statistics about the full range of policy actions and targets included in them. I use the higher-level categories, however, which the ensemble codes with higher accuracy, for the hypothesis tests in Chapters 4 through 6.

Table 14 breaks down the scores from this final ensemble approach by each of the high-level action and target entity categories. While the overall F1-score is 85%, there is variation across categories, and the two categories most important for this project perform well (highlighted in gray). The ensemble achieves a 92% F1-score for decrees with a main action in the *enable* category, and 84% for decrees targeting the *government executive*. The latter is slightly below the overall average, but is still impressive given that this category was originally one of the lower-level subcategories of *public sector* targets. (In Chapter 4, I discuss why I separated executive offices into their own high-level category, since they capture the most relevant trends for the phenomenon of interest.)

Table 14. Evaluation scores for ensemble classification by high-level category

Models	Features	Label Type	Label	P	R	F1
Rules,LR,RF	BOG, Structured	Action	<i>Enable</i>	89.5	95.1	92.2
"	"	Action	<i>Regulate</i>	88.7	78.2	83.0
"	"	Action	<i>Other Act</i>	80.4	79.1	79.4
"	"	Target	<i>Gov Exec</i>	88.0	80.8	84.0
"	"	Target	<i>Public Other</i>	83.8	77.3	80.1
"	"	Target	<i>Private</i>	81.5	90.4	85.6

P = precision, R = recall, F1 = F1 score

The fact that these particular categories receive high accuracy scores is not surprising, since they are the most well defined in the project, and are the concepts I spent the most time investigating and clarifying. I originally considered analyzing regulations or restrictions on power as well, but *regulating* actions proved more difficult to label. The Spanish words for “regulating” or “regulations” are often used for administrative rule-making, defining procedures for implementing previously enacted laws or programs, rather than addressing restrictions on authority per se. The categories for *other* actions and *private* targets are default categories in the

rule-based system, for decrees that do not match any other pattern, which explains their lower precision scores. The tests above indicate that the ensemble classification process developed in this project is able to categorize decrees in the dataset with reasonably high accuracy, especially those enabling the government executive. This should produce a reliable measure of the project's dependent variable, at least as defined in the coding scheme and the hand-labeled training documents and patterns used for classification.

3.5.3 Caveats to evaluating data measures by accuracy alone

A downside of supervised classifiers is that they can only learn categories that are known and sufficiently well defined by the researcher beforehand. The resulting data will preserve whatever assumptions were encoded into the training data, whether those assumptions arise from human coder interpretations or from real-world processes that generated the training data. The same is true of classifiers that are highly accurate when tested against human coding, whether they learn to reproduce the same coding scheme in a rule-based, supervised or unsupervised way.

While accuracy tests can evaluate the effectiveness of replacing human coding with an automated process that produces the same output, accuracy tests can't tell us whether the resulting data capture the most useful underlying concepts for the actual research objectives. In the next chapter, I turn to empirical analysis as a second form of validation, using bivariate regressions on control variables to assess whether the data extracted from decrees in this chapter captures meaningful trends with expected relationships to other real-world phenomena.

Chapter 4. Description of Decree Usage and Contextual Validation of Decree Categories

In this chapter, I present descriptive statistics of the decrees collected and categorized, analyze trends in the composition of those decrees in relation to established explanatory factors from the literature on presidential power and executive orders, and use those established factors (or control variables) to validate the conceptual interpretation of the decree categories to be used in subsequent hypothesis tests. The main goal of this chapter is to verify that a certain subset of decrees captures patterns we would expect to see not when leaders are simply seeking to enact more policies in general, but when they may be seeking to consolidate power for themselves.

4.1 Summary of decrees collected by country

After collecting decrees from the five Andean countries and automatically categorizing their contents using the process in Chapter 3, what do the resulting data look like? Table 15 summarizes the total decree titles collected and years covered for each country in the dataset. Figure 7 shows the decrees for each country aggregated by month and plotted over time.

Table 15. Total decrees and years covered for each country in the project dataset

Country	Years Covered	Total Decrees	Source
Peru	1980 – 2016	30,314	<i>Sistema Peruano de la Información Jurídica (SPIJ)</i> , Ministerio de Justicia, http://spij.minjus.gob.pe/
Bolivia	1982 – 2016	13,927	<i>Gaceta Oficial de Bolivia</i> , http://www.gacetaoficialdebolivia.gob.bo/
Colombia	1999 – 2016	16,260	<i>Presidencia de la República</i> , http://historico.presidencia.gov.co/decretoslinea/index.htm
Ecuador	2000 – 2016	9,711	<i>Registro Oficial de Ecuador</i> , http://www.registroficial.gob.ec/
Venezuela	2009 – 2016	3,458	<i>Gaceta Oficial, Tribunal Supremo de Justicia</i> , http://www.tsj.gob.ve/gaceta-oficial

While no public records archive is perfect, as discussed in Chapter 3, I selected the most consistent and comprehensive archives of decrees available in each country. To the best of my knowledge, the numbers shown in Table 15 represent the total decrees issued by each executive

during the years covered. There is some variation in how frequently decrees are issued in each country, but all are of the same order of magnitude on average. We can see that decrees are commonly used forms of decision making in all five countries – executives often issue more than one a day. Decrees are also issued fairly continuously; there isn't a clear seasonality like we would see with legislation, where congresses are in or out of session during different parts of the year. Yet the lines in Figure 7 are not very smooth; there is considerable month-to-month variation, as well as longer-term peaks and troughs that span multiple years.

Figure 7. Executive decrees issued by month for each country in the dataset

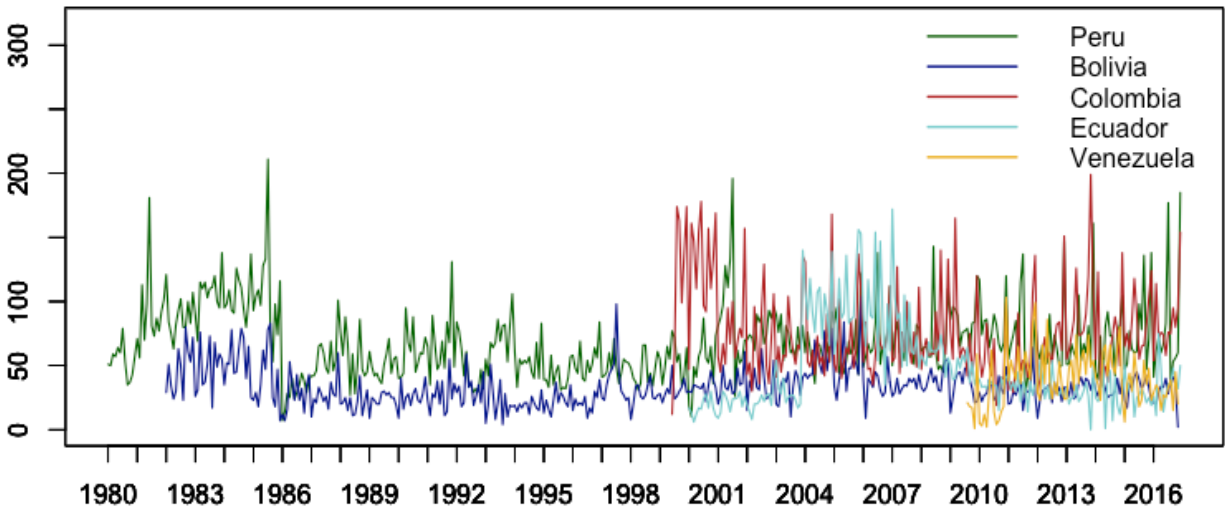
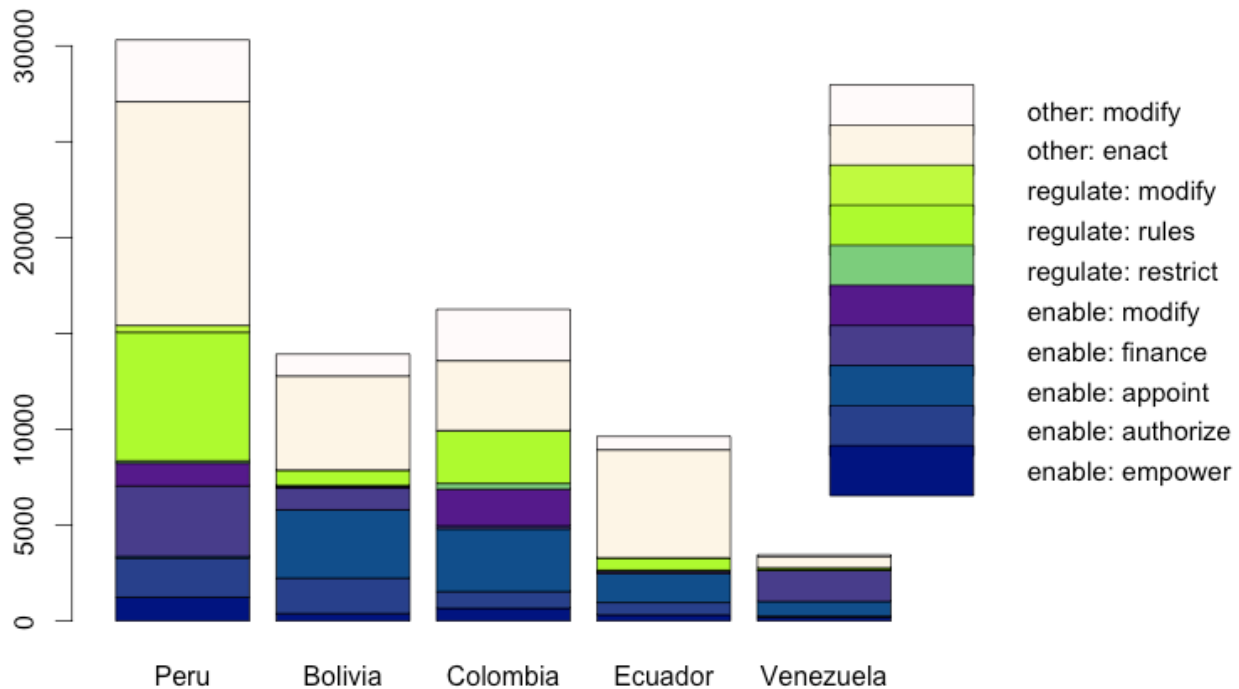


Figure 8 below shows the break-down of decrees by main action for each country, resulting from the coding process defined in Chapter 3. There is some variation in the composition of decree actions across countries, but all have at least a few thousand *enabling* decrees (in dark blues stacked at the bottom), which are the main actions of interest in this study. For most countries, around a third of the decrees fall into the *other* categories (in pale shades at the top), involving one-time policy implementation activities and symbolic measures.

There is more variation in *regulating* decrees (in greens in the middle of each bar), with presidents in Peru and Colombia issued many more decrees regulating or restricting organizations or activities, than the other three countries. Regulations most often target private actors, such as businesses, industries or workers (as shown in Figure 10 below), so the different frequencies may reflect the countries' different political economic models. Peru and Colombia have the most market-oriented economic policies, while Venezuela and Ecuador have the most socialist-style economies that include major state-owned enterprises, and Bolivia falls somewhere in between (but much closer to Venezuela and Ecuador in terms of foreign

investment¹²). Countries in which more actors outside the state are engaged in economic activity might experience more regulatory activity, as the government’s primary means of pursuing its economic policy objectives, while countries with more direct control of economic actors might carry out their economic policies more often through direct administrative action.

Figure 8. Decrees in dataset by country, broken down by main action category

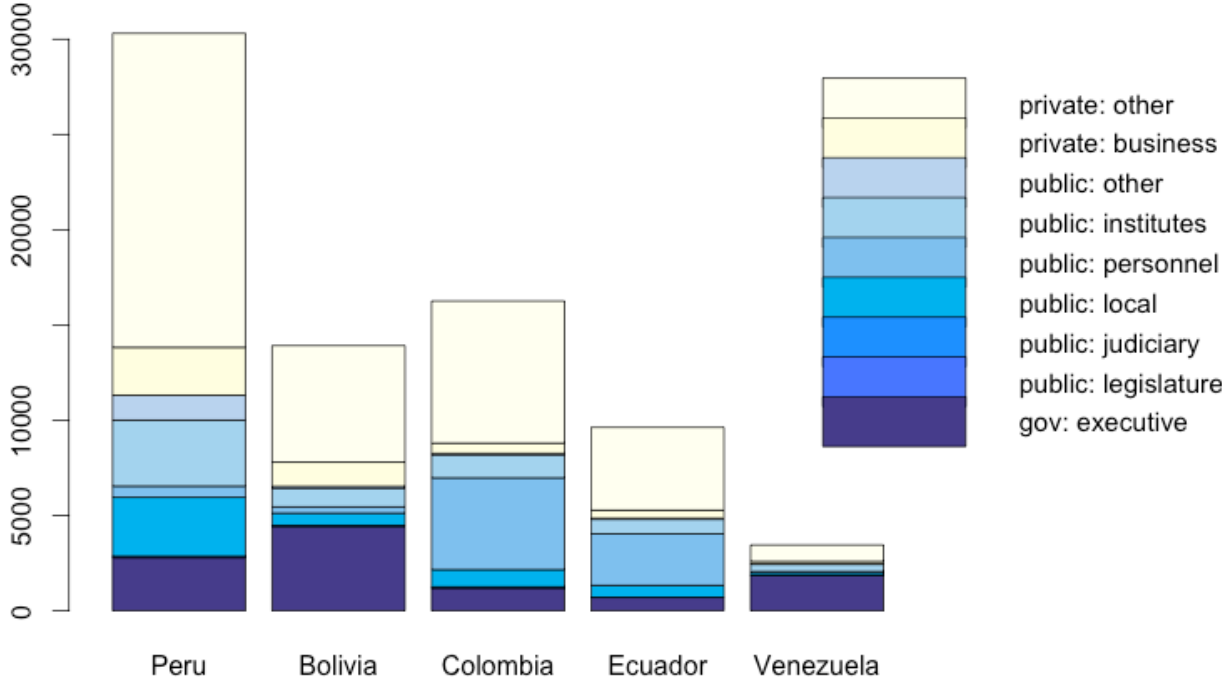


Venezuela had far fewer years worth of decrees publicly available in digital archives, and the composition of those decrees is the least diverse. The majority fall into the enabling category, primarily transfers of funds, along with some appointments. These decrees are coded correctly; inspection of the documents reveals that far more decrees in Venezuela were used to make additional budgetary allocations or supplementary credit transfers to specific ministries, agencies, and other government offices. This might indicate an inefficient or inadequate annual budgeting process, or greater arbitrary personal influence and cooptation of public resources for specific interests. Venezuela is also the country in the study that has the most ongoing political conflict and instability today. Keeping it in the dataset provides greater variation in the political conditions that might affect leaders’ motivations and pressures, especially for the analysis of decree categories in relation to control variables in this chapter, to demonstrate that decrees enabling the executive are used to consolidate power in a variety of political contexts.

¹² Foreign direct investment, net inflows (BoP, current US\$), World Bank, <https://data.worldbank.org/>

Figure 9 shows decrees by country broken down by target entity category. The fraction of decrees targeting private actors is largest in Peru and smallest in Venezuela, but in all countries at least a third of the decrees targeted public sector actors. The number of decrees targeting government executive offices in each country – the main target of interest in this project – varies from about 700 in Ecuador to over 4,000 in Bolivia.

Figure 9. Decrees in dataset by country, broken down by target entity category

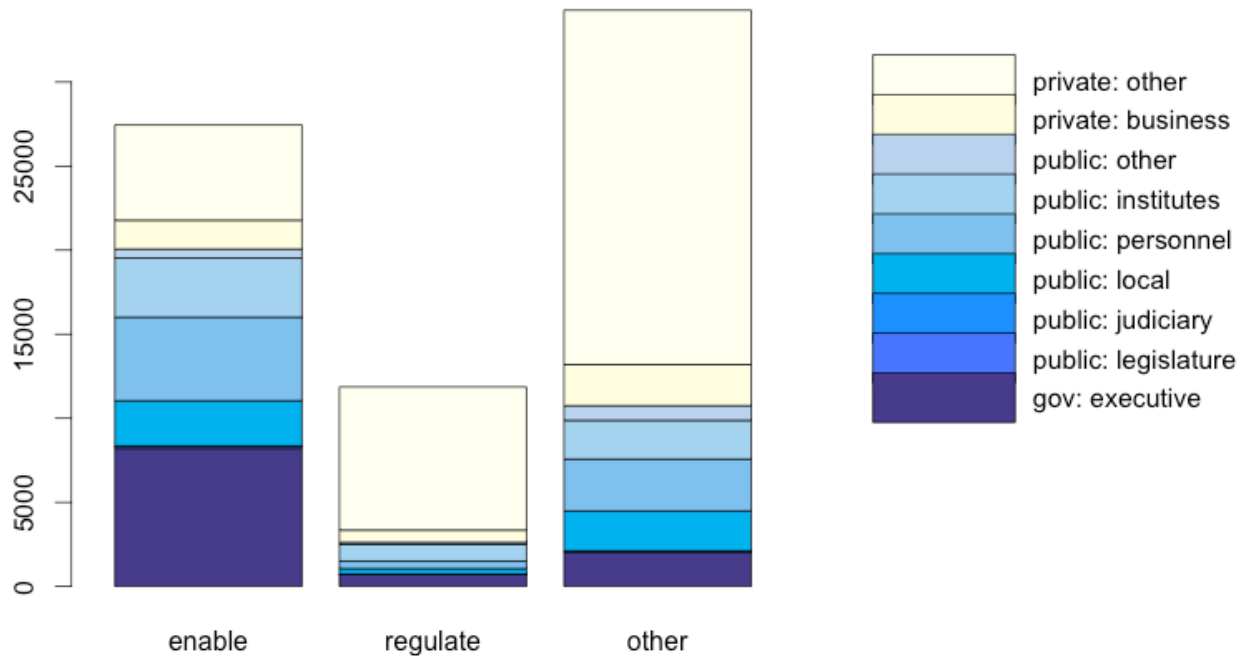


There are actually very few decrees that target the legislature or judiciary directly in any of the countries in the dataset. This is one reason why I focus on decrees that enable the executive as the primary means through which presidents appear to use unilateral decision-making to consolidate power, rather than decrees explicitly placing restrictions on rival authorities. (The analysis presented later in this chapter, on the relationships between the different decree categories and established control variables, provides more solid evidence that decrees enabling the executive best represent leaders’ efforts to consolidate power.)

Figure 10 shows how the main actions and target entities appear in combination, breaking down each high-level action category by the targets of those decrees. We can see that *enabling* actions more often target public entities, especially the government executive (the primary combination of interest in this study), while *regulations* overwhelmingly target private entities. Other decrees executing one-time policies or enacting symbolic measures also target private actors more than public. There may be some relevant decrees targeting executive authorities that ended up in the *other* action category, that involved language not seen in the training examples

enough to encode more precisely. This should not be a problem for the main analysis, since the majority of relevant decrees – especially those with the clearest articulations of power consolidating actions – are included in the *enabling* category, and the residual *other* decrees do not factor into the main analysis in Chapters 5 and 6.

Figure 10. Decrees in dataset, by combination of main action and target entity



In sum, there are 74,090 decrees in the project’s dataset, of which 8,497 were labeled as enabling government executive offices in the final ensemble classification process. Total decrees and decrees in each category vary across countries and over time. This dataset provides much more detailed and dynamic information than we could obtain with abstract annual indicators or total decrees on their own, for analyzing government leaders’ policy decisions and actions, as they pursue their objectives and respond to pressures or threats over the course of their tenure.

4.2 Using established correlates to distinguish motivations for decrees

To begin to evaluate what this data can tell us about government leaders’ decisions and actions, I first compare the decrees in different categories to a set of control variables derived from the literature on executive power, to test for certain expected relationships. The goal here is to evaluate whether the encoded categories of decrees (i.e. specific actions and target entities) offer useful distinctions about how leaders use decrees under different circumstances. More

specifically, I want to show that the category comprised of decrees enabling the executive does a good job of capturing the trends we should see when we expect leaders to consolidate power.

As discussed in Chapter 2, leaders might increase their use of decrees for multiple reasons, including to simply get more policies enacted in the face of urgent demands or gridlock in other decision-making channels. Most of the literature on presidential power assumes that leaders issue more executive orders or decrees to enact their policy agenda, in order to win reelection or build a positive legacy. I propose that leaders also sometimes use unilateral decrees to make more lasting changes to the allocation of government power, in order to weaken rivals and protect their own control of the state. How can we tell, then, whether leaders are just trying to get things done or are pursuing power for its own sake?

In Chapter 2, I proposed that leaders would take different actions – e.g. creating or reorganizing public offices, versus delivering programs and services to private constituents – depending on which objective they are prioritizing. Based on that logic, I have sought to measure leaders’ efforts to consolidate power by categorizing decrees’ stated actions and target entities. To justify this approach, it would be useful to find other indicators that show leaders take the particular actions I have encoded, when we would most expect them to consolidate power. For instance, we can look at circumstances in which we would expect incumbents to feel the need to protect themselves in office, versus when we would expect them to prioritize executing their agenda, and see which decrees they enact during each of those times.

This analysis represents a form of measurement validation often called “construct validity”. This is distinct from “criterion validity”, in which the researcher tests for correlation between one measure of a particular concept and other indicators of the same phenomenon (Adcock and Collier 2001). The accuracy tests in Chapter 3 can be thought of as serving a similar purpose to criterion validity, showing that the automatically categorized decrees provide reliable and consistent measures of the decree actions and targets that a human coder assigned to them. In the remainder of this chapter, I now test for expected relationships to *other* variables which are theoretically associated with leaders’ efforts to consolidate power. This type of evaluation assesses the downstream usefulness of the categorized decrees: do they capture meaningful empirical relationships to other real-world phenomena?

For this analysis, I aggregate decrees by month for each of the action-target combinations encoded in the last chapter, then divide by the total decrees for that month, so that the dependent variable is the percentage of decrees issued in a given month that contain a particular type of action and target entity. By regressing the *fraction* of decrees in each category on the expected covariates, we can observe how leaders change the *composition* of their decrees in different circumstances, beyond changes in the total volume of decrees issued at a given time. In the main hypothesis tests in Chapters 5 and 6, I will use whole counts of power-consolidating decrees. Here, I use proportions to explore and validate the categories in comparison to each other.

In the following subsections, I present bivariate regressions of each decree category (as a fraction of the monthly total) on each expected covariate (defined below), using ordinary least squares with country and year fixed effects. I cluster standard errors by presidential term, since

there may be serial correlation between months within the same president's tenure. I do not combine multiple explanatory factors into the same model, since many of them are correlated, and testing select combinations of factors can produce unstable results. (In Chapters 5 and 6, I take a more computational approach to combining multiple explanatory factors into regression models for the main hypothesis tests.)

The objective in this chapter is simply to compare pairwise relationships between different subsets of decrees and expected covariates, rather than test a comprehensive theoretical model and interpret the results in absolute terms. Some tests may capture spurious correlations, when important variables are omitted, but there should be even stronger relationships in the tests in which more directly related variables are included.

4.2.1 What actions do leaders take when facing legislative opposition?

In the existing literature on presidential power discussed in Chapter 2, one of the main explanatory factors found to contribute to increased use of executive decrees is greater opposition control of the legislature. I use this as the starting point to analyze the decree categories identified in this project. If the president's party controls fewer seats in the legislature, the president should be less successful in passing legislation, which might make the president more inclined to enact decisions through decrees instead. In this section, I use variables from the Database of Political Institutions (DPI) compiled by the World Bank Development Research Group and the Inter-American Development Bank (Cruz et al 2016).

Figure 11. Relationship between total decrees issued and legislative control

Estimates from bivariate regression of total decrees on government parties' (top row) and opposition parties' (bottom row) share of seats in legislature, respectively.

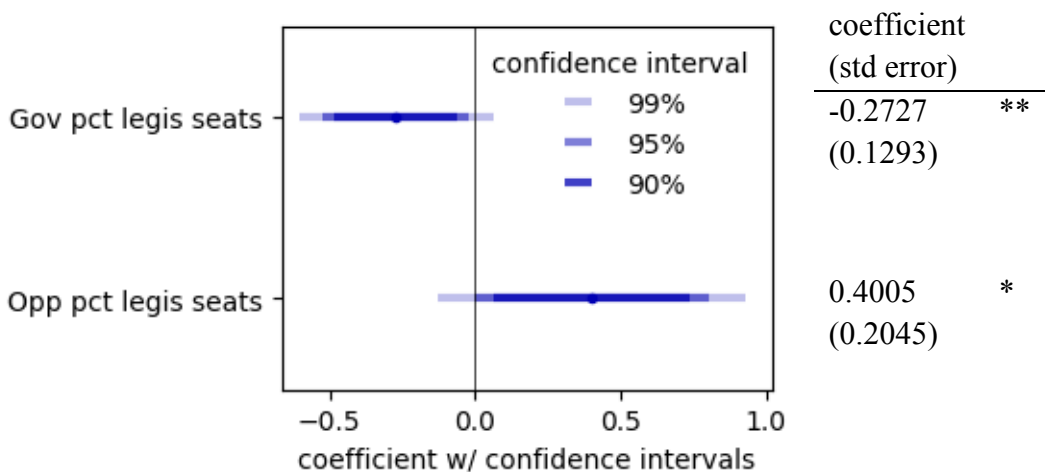
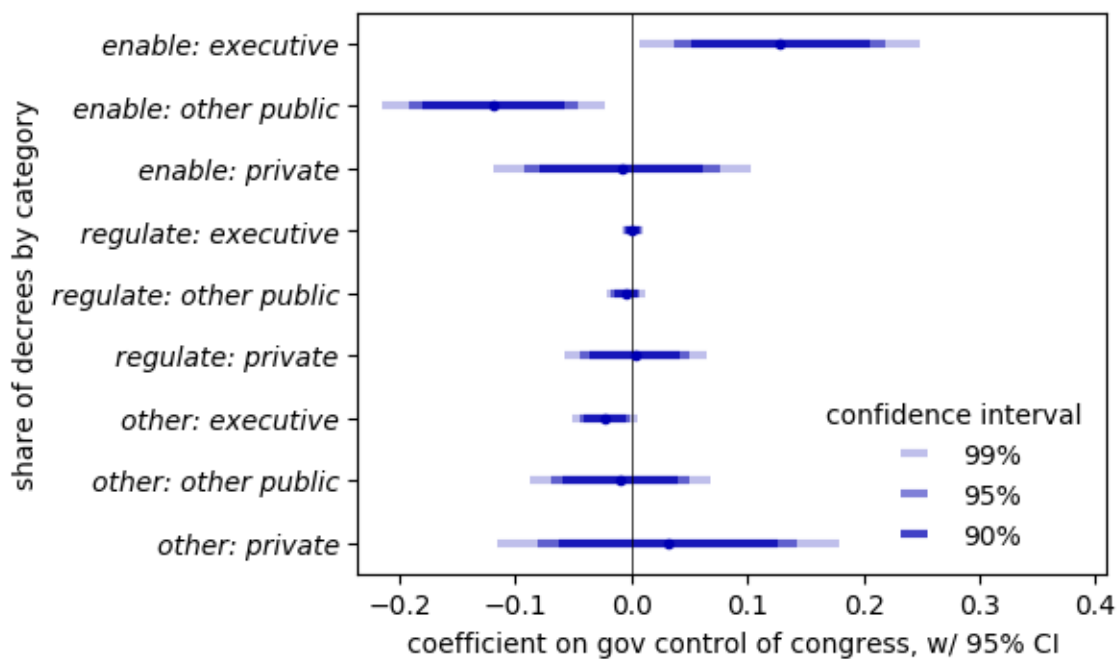


Figure 11 shows a significant negative relationship between total decrees and the governing party's share of seats in the legislature, as well as a significant positive relationship between total decrees and the main opposition party's share of legislative seats. However, the president's *motivations* for issuing more decrees in the face of legislative opposition are ambiguous. The incumbent might be concerned about not getting enough programs and services enacted through legislation. But the incumbent might also be concerned about the legislature turning against the president and taking other action, such as investigating misuses of power or pursuing impeachment. In the latter case, the president might choose to issue more decrees, not simply to overcome gridlock, but to expand the power of the executive to protect it against potential challenges to the leader arising from an opposition-controlled legislature.

Figure 12 below shows coefficients and confidence intervals for the fraction of decrees in each of the different action-target categories, regressed on the share of seats in the legislature controlled by the president's party. While most types of decrees have a negative relationship to executive control of congress, the share of decrees enabling the executive *goes up* when the president's party controls more legislative seats. This suggests that presidents do not issue more decrees enabling executive offices when we would expect them to be compensating for limited control of congress. In other words, they do not appear to be using these decrees to directly substitute for legislation. Since the coefficient refers to the *share* of decrees that enable executive offices, this does not mean that presidents issue more absolute decrees enabling the executive when they control more seats in congress. Rather, they tend to issue more of the *other types of decrees* when it is harder for them to execute their agenda through the legislature.

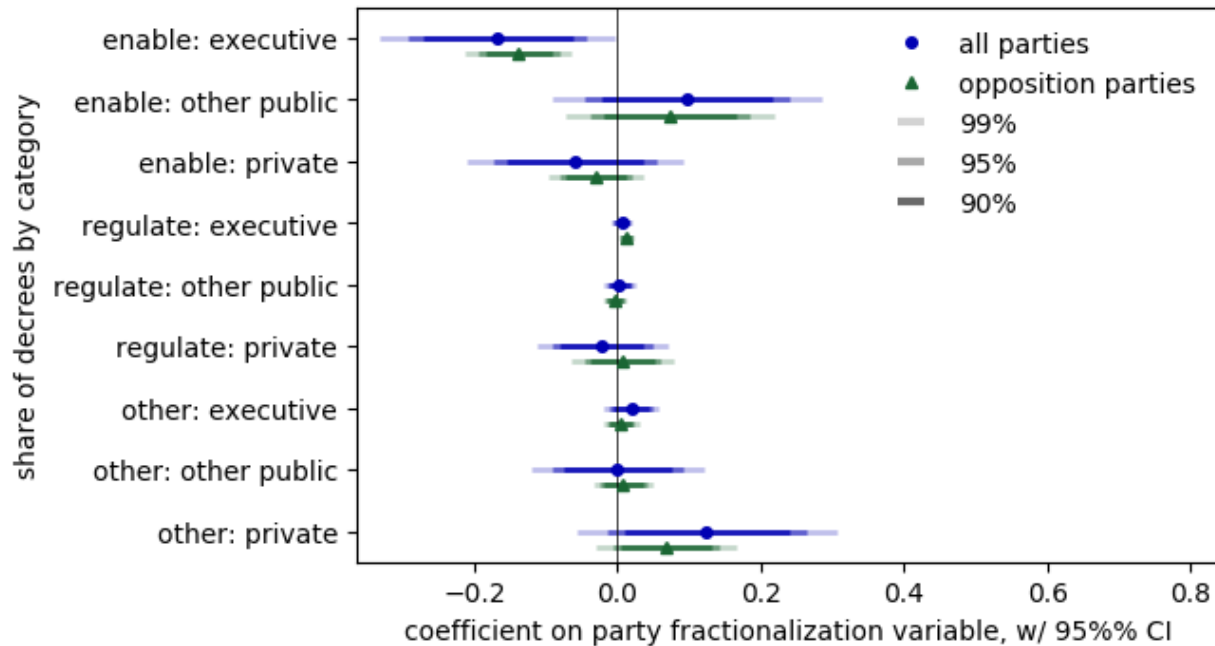
Figure 12. Bivariate regression of decree categories on governing party control of congress



How can we distinguish between legislative opposition that simply blocks the president’s policy agenda, and legislative opposition that threatens the president’s control of office (or future fate)? As discussed in Chapter 2, a more useful indicator is the degree of fractionalization among political parties, especially opposition parties. If legislative seats are divided among many small parties, it is probably harder for the governing party to coordinate the legislative coalitions necessary to pass executive-sponsored bills. However, highly fragmented opposition parties are less likely to threaten the president’s power, since it should also be harder for them to coordinate effective action among themselves to investigate, censure, or impeach incumbent officials.

Figure 13 shows regression coefficients and confidence intervals for each decree category when regressed on the DPI measures for overall party fractionalization (blue circles) and opposition party fractionalization (green triangles). In both cases, we see a clear difference in terms of which decrees presidents issue when facing a fragmented party system. When all parties, including opposition parties, are highly fractionalized, presidents are more likely to issue decrees in the *other* category (executing programs or services, or symbolic actions) targeting private actors. This suggests that presidents prioritize enacting those measures through decrees when fragmentation and gridlock make it more difficult to accomplish everyday tasks through legislative processes. Most other categories do not show significant relationships here.

Figure 13. Bivariate regression of decree categories on party fractionalization



However, decrees enabling executive offices are highly significant and *negatively* correlated with party fractionalization, especially opposition party fractionalization, indicating

that presidents feel the need to issue more of executive-enabling decrees when the opposition is unified (i.e. low fragmentation). In other words, presidents shift more of their focus toward empowering the institutions they control, not when it is harder to coordinate legislative action with many small parties, but when it is more likely that collective action may be taken *against* them. This supports the interpretation that decrees enabling executive offices represent leaders' efforts to consolidate power, not merely to get things done, but to protect themselves in the face of potential challenges to their hold on power or their future political interests.

Notably, decrees enabling (or restructuring) other public sector actors do not show the same trend. If anything, the second category moves in the opposite direction (although the relationship is not significant). Decrees regulating (or modifying restrictions on) the executive and other public actors also do not capture the trends we would expect to see when leaders face threats to their power. This is one of the main sets of evidence that the decrees enabling executive offices (rather than decrees enabling or restricting other public sector actors) best represent leaders' efforts to consolidate power.

4.2.2 What actions do leaders take during disasters and economic shocks?

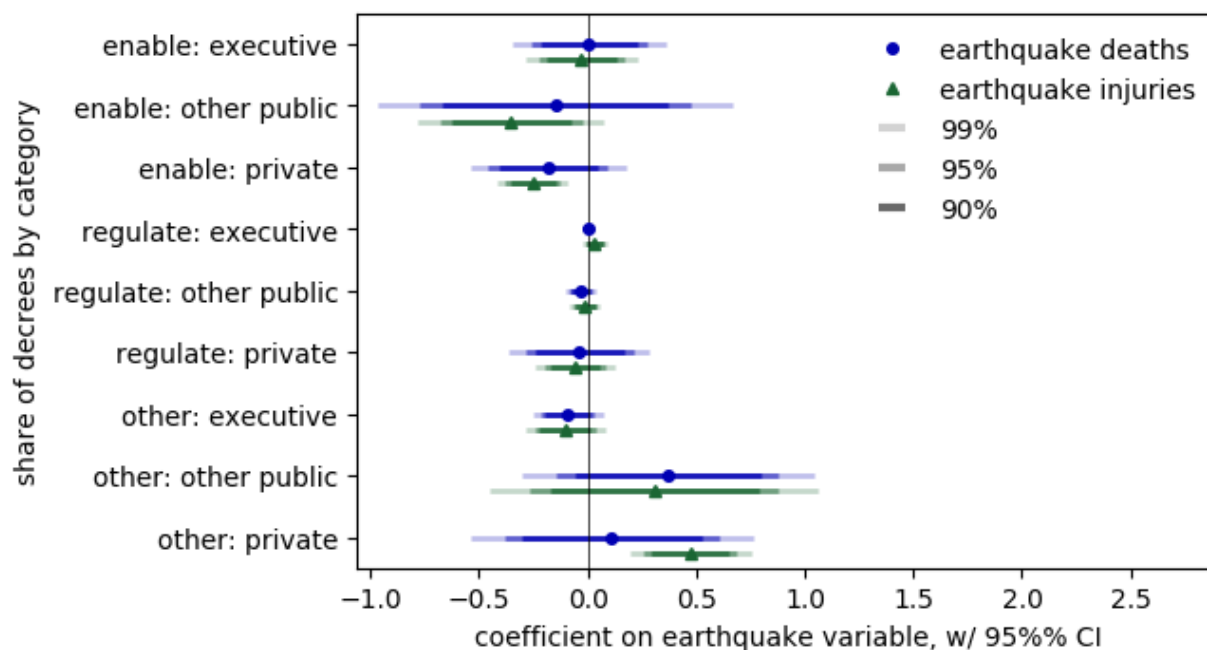
Another common justification for executive decrees is the need to take swift action during emergencies. I consider two types of crises: economic downturns and natural disasters, which occur with sufficient frequency in the Andean region, and are easier to define and measure than other forms of social conflict. Violence and civil unrest are also popular justifications for swift executive action, but these types of conflict are highly endogenous to government leaders' own decisions, which would make the direction of influence difficult to interpret.

Natural disasters do not occur in response to political developments, making them a nice baseline example of an emergency situation in which leaders are unlikely to face blame for the incidents themselves. Economic downturns are also often influenced by the president's policy decisions, but there are certain types of economic shocks that scholars consider reasonably exogenous, especially global price shocks for primary export commodities on which the national economy depends, but which the national economy is not large enough to significantly determine. While these shocks are sufficiently exogenous for social scientific analysis, constituents may not be aware of the origins of an economic downturn, so that leaders might still face blame and heightened risk of removal or sanction when these shocks occur.

For natural disasters, I use data from the Global Significant Earthquake Database maintained by the U.S. National Oceanic and Atmospheric Administration (NOAA)¹³. The database contains consistent measures of magnitude, physical damage and human costs across major earthquakes worldwide over more than a century. Figure 14 shows bivariate regressions of the various decree categories on earthquake deaths and injuries, which reflect not only the magnitude of the earthquake but how badly it hit populated areas.

¹³ National Centers for Environmental Information (NCEI), <https://www.ngdc.noaa.gov/hazard/earthqk.shtml>

Figure 14. Bivariate regression of decree categories on earthquake disasters



The categories with the largest positive relationships involve *other* actions, targeting non-executive public offices and private actors. Other public offices include local governments, such as cities and provinces. Since earthquake effects are usually localized, targeting local areas with immediate response measures would make sense. Overall, these results suggest that the *other* action category – which includes one-time programs and services – probably covers the main measures leaders use to respond to exogenous disasters like earthquakes, which call for urgent action but do not provide immediate justification for leaders’ removal.

Meanwhile, we do not see any relationship to decrees enabling the executive during natural disasters, which is consistent with the expectation that natural disasters should not affect government leaders’ concerns about threats to their power. I also explored alternative data sources covering other natural disasters – such as floods and disease outbreaks, captured in the EM-DAT Emergency Events Database (Guha-Sapir 2017) – to use as control variables in the more comprehensive hypothesis tests in Chapters 5 and 6. The results were nearly identical, however, since natural disasters do not appear to be associated with decrees enabling the government executive, nor with former leaders’ fates.

For economic shocks, I use data from the International Monetary Fund (IMF) on the global prices of each country’s top three primary export commodities. For each month in the dataset, I average the percentage price change across these primary export commodities for the previous quarter, six months, and year (in three separate variables), since it might take time for downturns in export prices to affect local industries and government responses.

Figure 15 shows bivariate regression estimates for these export price changes and total decrees. The coefficients are all negative and at least somewhat significant (at the 95% level for the 3-month average change, at the 90% level for the other two), indicating that leaders do issue more decrees during exogenous economic downturns. However, as with legislative opposition, leaders' motivations for issuing more decrees during economic downturns are ambiguous. On the one hand, they might simply be taking action to mitigate the crisis or alleviate hardship, by delivering more services or benefits to private actors in affected industries or regions. On the other hand, leaders might fear that they will be blamed for the downturn, and that political opponents might take advantage of the crisis to remove them. If so, incumbents might issue more decrees to consolidate their own power and protect themselves against challenges.

Figure 15. Bivariate regression of total decrees on export commodity price shocks

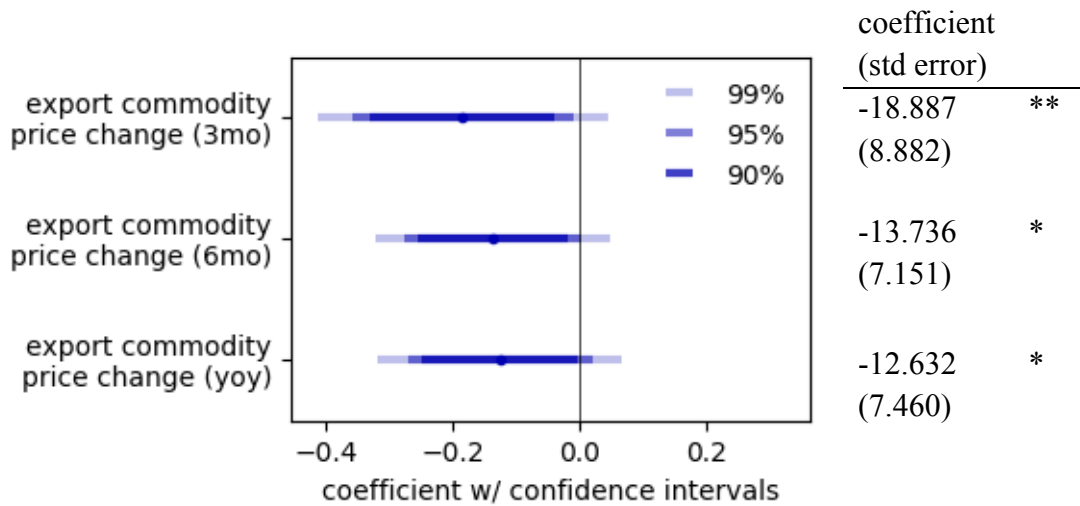
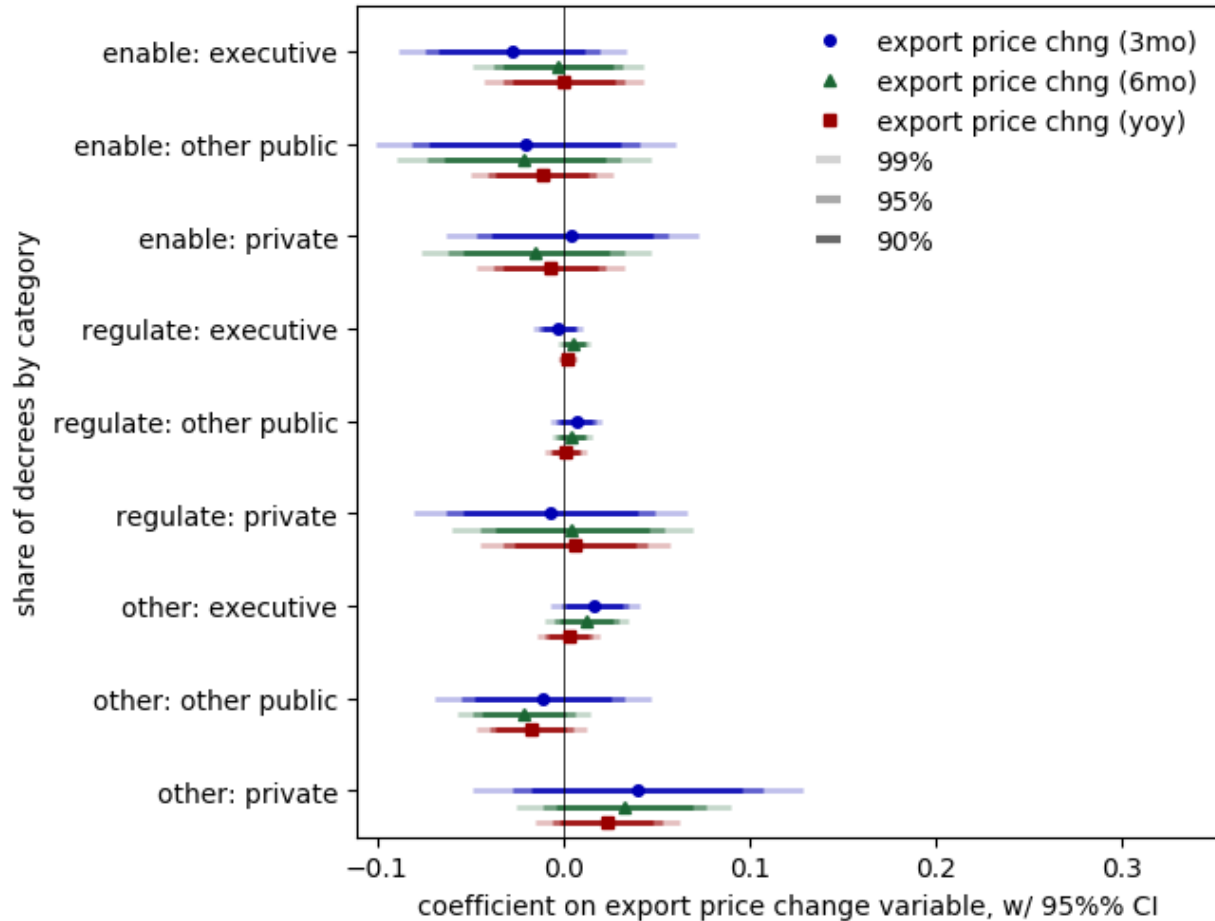


Figure 16 below shows the breakdown of decree categories, regressed on the same export commodity price changes. Unfortunately, we do not see a clear picture in terms of which types of decrees are issued more often in the face of economic downturns. We know that overall decrees are significantly correlated with export commodity price changes, but the results are insignificant for most of the specific decree categories. The coefficients on most of the *enabling* categories are negative, and on most *regulating* categories are positive. This might suggest that leaders do more enabling (of both public and private entities) when the economy performs poorly (perhaps to stimulate recovery), and more regulating when the economy performs well, but weakly so.

Decreases in global commodity prices are a fairly indirect source of variation in the incentives or pressures a national government leader might face. These exogenous shocks might affect leaders differently, depending on other factors, such as how likely their opponents are to blame the crisis on the incumbent and use it to challenge the leader's power. In Chapter 6, I will test for heterogeneous effects of exogenous price shocks when interacted with this project's main

explanatory variable: former leaders' fates. We might see some leaders react to economic downturns by consolidating power, while others do not, depending on how they perceive the risks of removal from office and punishment afterwards.

Figure 16. Bivariate regression of decree categories on export commodity price shocks



4.2.3 What actions do leaders take as they near the end of their term?

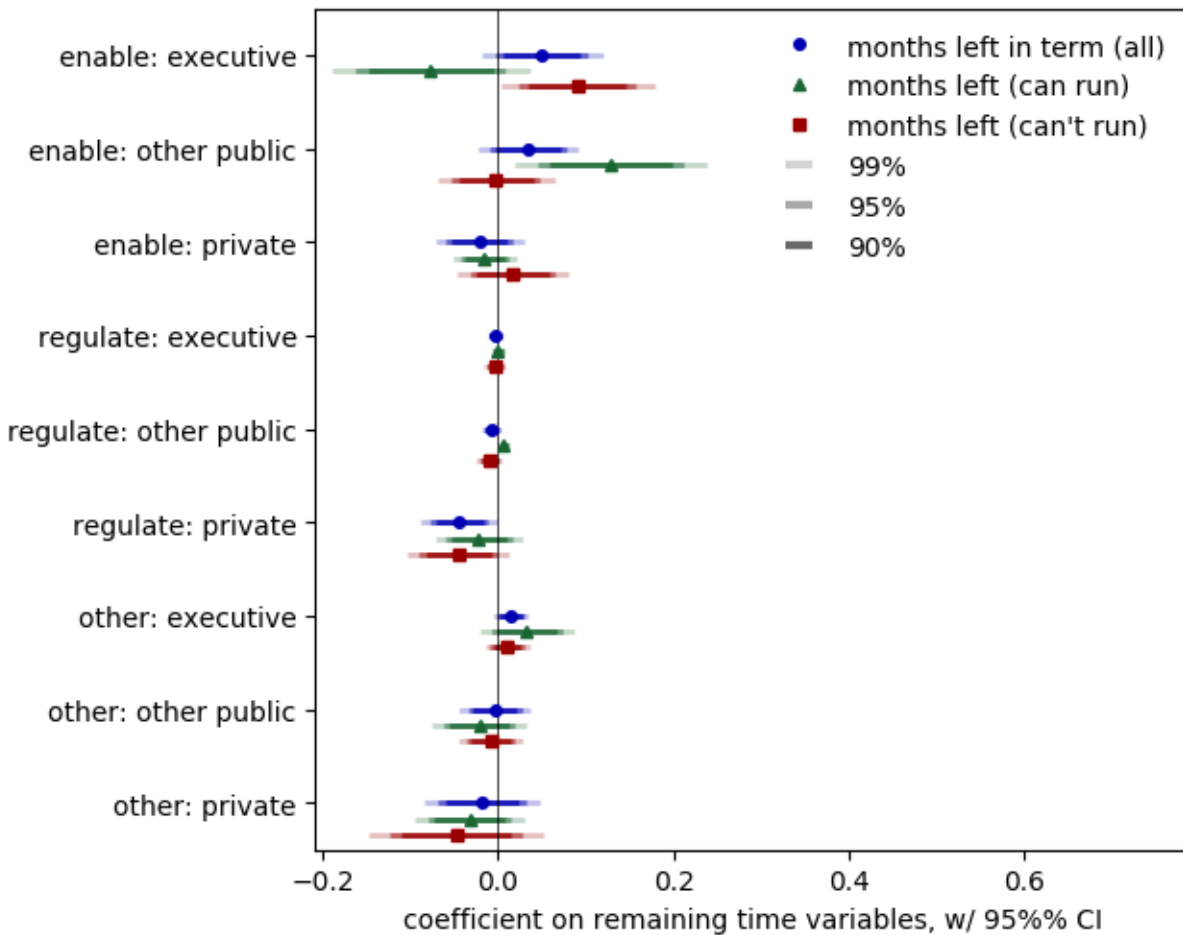
As a final question regarding leaders' objectives for issuing decrees, we might ask which types of decrees they enact as they near the end of their time in office. If leaders' primary motivation for issuing decrees is to accomplish their policy goals, we might see the frequency of decrees increase when the leader has little time left to complete that agenda. The pattern should be the same whether the leader is eligible for reelection or not, since a successful policy record should be useful both to win a new term and to secure a positive legacy.

However, if a leader is worried about what will happen after departing office, and uses decrees to alter internal government authority structures in his/her favor, those decrees might

look very different depending on whether the incumbent expects to remain in the presidency or not. A leader facing the end of his/her tenure, who is not eligible for reelection, should be less likely to enact lasting institutional changes that strengthen the presidency or executive branch, since such changes would most likely only empower a successor.

Figure 17 supports this logic, showing regression results for the different decree categories when regressed on the number of months left in the president’s term, with different bars separating the cases in which the president was eligible for reelection (green triangle) or not (red square). I coded these variables based on each country’s constitutional provisions on presidential term limits. In some cases, presidents were initially barred from reelection, but managed to change the constitution or obtain a judicial ruling allowing them to run again. In those cases, I recorded remaining months left under the “can’t run” variable up until the legal change, then moved remaining months left to the “can run” variable thereafter.

Figure 17. Bivariate regression of decree categories on months left in term



For most of the decree categories in Figure 17, the coefficients maintain the same sign, regardless of whether the president was eligible for reelection or not. In either case, as presidents near the end of their term, they appear to direct their attention more toward private actors. In particular, they appear to issue more decrees *regulating* (including revising taxes and other rules) or delivering *other* actions (including one-time programs and services, and symbolic acts) to private organizations and constituents. The fact that the coefficients are negative (if not all significant) regardless of the leader's eligibility for reelection, supports the interpretation that when leaders increase those types of decrees, they are simply trying to accomplish their policy agenda, to win support among voters for reelection or build a positive legacy.

The second and third bars diverge most, however, for decrees enabling executive offices. When the president is allowed to run again, decrees empowering the executive increase as time in office gets shorter. The coefficient is very large and statistically significant at the 90% level. When the president is constitutionally barred from reelection, however, decrees empowering the executive *drop off* toward the end of the leader's tenure. The highly significant positive relationship indicates that incumbents are much more likely to issue those decrees earlier in their tenure. This strongly supports the interpretation that the decrees enabling executive offices produce meaningful and lasting increases in the president's authority. Presidents appear to seek more of those powers when nearing a battle to retain office, since control of major agencies can help them run and win reelection, but they strongly avoid doing so when it is fairly certain that those powers will fall into someone else's hands. Again, we do not see this same relationship for decrees enabling other public offices; if anything the relationship is reversed.

4.3 Finalizing the decree categories to use in main hypothesis tests

The analysis presented in this chapter supports the use of decrees – specifically those enabling executive offices – to represent leaders' efforts to consolidate power, when we would expect them to be protecting their own interests rather than simply trying to get things done. To conduct this analysis, I used established covariates from the previous literature on presidential power, in order to identify which subsets of decrees reflect more traditional assumptions about leaders' motivations, and which decrees reflect the motivations of interest in this project.

As mentioned in Chapter 3, I initially considered using all decrees that enable or restructure authority among any public sector entities, as potential power-consolidating decrees. However, in both the analysis of how leaders react to opposition party fragmentation and how they shift course toward the end of their term, the decrees enabling other public offices show the opposite relationships to those enabling the executive. I also considered using decrees that either modify regulations on the executive, or that increase restrictions on rival government branches, as another measure of leaders' efforts to consolidate power. However, regulating decrees also fail to capture the expected trends that would indicate efforts to consolidate power. Upon inspection of the hand-coded decrees used for classification, *regulating* verbs are often used for

more mundane rule-making acts to implement legislated programs, and the language used to articulate true restrictions on authorities or activities is more difficult to encode, since it is often simply signaled by negation, e.g. “X office *shall not* do Y”.

In sum, in circumstances in which we would expect leaders to seek to protect themselves against threats to their power, we observe more use of decrees to expand the authority and resources of the offices the executive controls, rather than explicitly decreeing changes to the powers of rival authorities. The analysis in this chapter has shown that the decrees enabling executive offices capture by far the most consistent trends with what we would expect to see when leaders are motivated to consolidate power. Decrees enabling other public offices do not show the same trends, nor do decrees involving regulations on either set of public offices.

Because of this analysis, I decided to use the decrees containing *enabling* actions that target *government executive* offices only, in the main analysis in Chapters 5 and 6. Since the *executive* was originally a subcategory of all *public sector* actors in the project coding scheme, I went back and revised the coding scheme to include not two but three high-level target categories (government executive, all other public sector actors, and private actors). This allows me to use the high-level classification for both decree actions and targets, in the final data to be used for analysis, since the machine learning classifiers are more accurate when encoding only a few high-level labels than when encoding all of the other subcategories as well.

This represents an iterative process in which both the accuracy of the classification algorithms and substantive analysis of relationships to expected covariates informed the final categorization to be used in the main hypothesis tests. It is important to note that the decision about which decree category to use for the final analysis did not involve the new explanatory variables to be introduced in Chapter 5. This decision was made solely based on control variables derived from pre-existing theories about executive authority. If anything, the use of control variables to select and refine the main dependent variable specification should favor the pre-existing explanations in any subsequent analysis, setting a high bar for testing my own theory on the same data. In the next chapter, I turn to the new set of explanatory variables, former leaders’ fates and incumbents’ expectations of their own prospects for future punishment or reward.

Chapter 5. The Relationship Between Power Consolidation and Leaders' Fates

The final piece to the puzzle is the addition of this project's main explanatory variable: the fates of former presidents, which may signal information to subsequent leaders about how their actions will be punished or rewarded. I am interested not only in *whether* former leaders were punished, but in what ways and by whom. I am also interested in the potential back-and-forth between sanctions and reprieves, as well as the role of post-tenure rewards. To dissect these different types of developments in each former leader's fate, I first needed to collect new data in the form of events, which might occur at various points for years after a leader has left office.

In developing an explanatory variable to be used for rigorous hypothesis tests, this is also the point at which it is important to address concerns about potential endogeneity between the dependent and independent variables. We may find a statistically significant relationship between the two, but that correlation may not provide useful lessons to apply in practice, if there is a strong possibility that the outcome influenced the purported cause, rather than the other way around. As discussed in the introduction and theory chapters, it would be extremely difficult to manipulate powerful government institutions in randomized experiments. Studies of government power usually depend on historical observations, and the complex interactions among interdependent governing branches and parties inevitably produces concerns about endogeneity between the actions and reactions that comprise institutional struggles for power.

In the data collection and analysis of leaders' post-tenure fates, I seek to address concerns about endogeneity in several ways. First, I only include the fates of *former* leaders in the explanatory variables. Sitting presidents might also face legislative or judicial investigations into wrongdoing before they depart office, or at least preliminary inquiries, depending on the extent of legal immunities assigned to their positions. However, I do not include any in-office sanctions against an incumbent leader in the project's dataset, so that the actors being sanctioned are different from the actors making decisions about power consolidation.

Second, I lag all explanatory variables by at least one month in the baseline dataset, and in the final regression models, I include additional lag periods farther back in time, to ensure that the sanctions and rewards against former leaders occurred before the decrees were issued that the sanctions are used to explain. Third, I distinguish sanctions by different instigating actors, to test for different hypothesized relationships that are each consistent with the theory, which helps ensure that the evidence I find in support of the theory does not depend exclusively on a single relationship between two actors, which might be interpreted the other way. Finally, I test for a high degree of robustness to alternative specifications, using random permutations of regression model parameters, to reduce the possibility that the results could be driven by spurious correlation arising from an arbitrary configuration of covariates.

I discuss each of these steps in more detail in the remainder of this chapter, then present the first set of empirical results, using the main categories of post-tenure fates defined below.

5.1 Data sources for leaders' fates

For measuring leaders' expectations of survival after departing office, a useful example comes from the Archigos dataset of political leaders, mentioned in Chapter 2 (Goemans et al 2009). Archigos includes the fates of former leaders in the first year after departing office, coded by general categories for death, imprisonment, or exile. The dataset only contains one fate per leader, rather than a series of potentially conflicting events, and does not include information on the actors involved. Archigos focuses more on violent leadership transitions, and contains more detail on how leaders were removed from power, including whether government actors, military, or rebel forces participated, with or without foreign support. I focus on attempts to hold leaders accountable for abuses of power in democratic contexts that do not necessarily involve violent conflict, and my theory involves more aspects of how leaders were punished or rewarded over time. My approach is consistent with the general model established by Archigos, however, for studying political leaders' behavior when facing threats to their future interests.

As with the dependent variable, I have chosen to create new measures of the main explanatory variables in this study, in the form of event data rather than a static state for each leader's ultimate fate. I initially explored a variety of sources to find reasonably comprehensive, consistent, and reliable information on former presidents' fates across the five Andean countries in this study. I began with news reports, reviewing international English-language newspapers and newswires, as well as regional and national Spanish-language newspapers. As noted in Chapter 3, news media are common sources of data for studying political events that occur outside of or between government actors. However, English-language coverage of local events in developing countries is spotty and inconsistent, and local Spanish-language newspapers vary widely in the availability and searchability of digital archives, most only going back a few years.

Furthermore, some types of post-tenure events are more likely to appear in news reporting than others. Sensational developments like criminal convictions and incarcerations are most likely to make international headlines, while intermediate steps like legislative investigations that do not lead to criminal charges may not, and positive fates like a leadership position in a private or civil society organization may be even harder to find. No channel of secondary reporting is likely to contain every development in former leaders' punishments or rewards. For the purposes of this research project, it seemed most important to prioritize consistency, across countries and leaders over time, as well as across different types of events that might occur after a leader's departure from office.

More comprehensive accounts of both punishments and rewards may appear in historic biographies of former leaders. The next step in my exploratory process was to turn to potential biographical sources, including scholarly case studies, historical narratives, and memoirs.

Again, the challenge was to identify comparable sources across leaders, since there are major differences in the attention that historians have paid to former Latin American presidents. Biographies and academic sources also tend to include selective historic events, based on the authors' chosen narratives or theories of historical salience. Books also tend to focus on current or very recent presidents (at the time of writing) and those leaders' experiences leading up to and during their presidency, rather than revisiting former leaders later on.

Due to these observations and considerations, I decided to use Wikipedia as the primary source of data on former leaders' fates. I use both Spanish-language and English-language Wikipedia sites, which both have individual biographical articles on all presidents in this study. I considered other websites that contain encyclopedia-like entries on multiple Latin American leaders. But such websites are often maintained by research institutes or organizations with at least some ideological bias, if not outright political agenda, which can skew the inclusion of positive versus negative events reported for different leaders. In contrast, Wikipedia is a widely used free online encyclopedia that is edited by a large and dispersed collective of volunteer contributors, and its accuracy and reliability have been favorably assessed as comparable to more traditionally edited reference sources (Giles 2005; Okoli 2009). A comparison of Wikipedia to work by academic historians found that Wikipedia's prioritization of neutrality may reduce the depth of analysis or interpretation of historic events, but also makes Wikipedia articles "more factualist," and particularly accurate in capturing names, dates, and events in biographies of historical figures, where the unit of analysis is clear (Rosensweig 2006).

The Wikipedia articles on former Latin American presidents contain consolidated summaries of each leader's major life events, including the kinds of information (e.g. dates, locations, and perpetrators) necessary for comparative analysis. Since Wikipedia is not written or edited by scholarly experts, some critics do not consider it authoritative for research that hinges on the accuracy of specific details, such as in-depth case studies. However, with the growth of automated information extraction and similar data collection efforts, some level of noise may be considered acceptable (and unavoidable), as long as there is no expectation of systematic bias in the data. While no Wikipedia article includes every single instance of punishment or reward that occurred to a given political leader, I found that the articles shared similar lengths, organizational structures, and types and magnitudes of biographical events mentioned across former presidents.

Working with three undergraduate research assistants, I collected data on all post-tenure events mentioned in each leader's English and/or Spanish Wikipedia entry. We consulted the Wikipedia citations and searched for other sources (e.g. news reports) to confirm or add key details, mainly exact dates or actors involved. The resulting data contain major developments in former presidents' post-tenure fates, which were selected for inclusion based on a common, politically neutral reference that provides consistent overviews of public figures' life stories.

5.2 Coding scheme for post-tenure events

I defined the following types of post-tenure events for inclusion in the dataset, summarized in Table 16 and discussed below. As introduced in the theory in Chapter 2, the main types of post-tenure events in this project’s coding scheme are *sanctions*, *reprieves*, and *rewards*.

Table 16. Summary of post-tenure fate events, actors involved, and additional factors

<i>Fate</i>	<i>Examples</i>	<i>Actor/Agent</i>
<i>Sanction</i>	<ul style="list-style-type: none"> • Investigation • Arrest, detention • Prosecution, conviction • Exile, extradition 	<ul style="list-style-type: none"> • Legislature • Executive • Judiciary • Foreign government
<i>Reprieve</i>	<ul style="list-style-type: none"> • Denied extradition • Dismissed charges • Acquittal • Pardon 	<ul style="list-style-type: none"> • Legislature • Executive • Judiciary • Foreign government
<i>Reward</i>	<ul style="list-style-type: none"> • Other gov post • Corporate role • Academic position • Honor, award 	<ul style="list-style-type: none"> • Public sector • Private sector • Political party • Foreign organization

For sanctions, I am interested in any attempt to formally punish a former president for his/her acts in office, including through a congressional investigation, judicial inquiry, criminal prosecution, or related actions like stripping a public official’s immunity or ordering his/her arrest and detention. For rewards, I am interested in any formal positions, honors, or other awards that demonstrate former presidents can remain free and enjoy other sources of wealth, status, and influence outside of the presidency. These may include attaining other government offices or leadership roles in business, academia, or other civil society organizations. I include a third category for reprieves that do not refer to positive rewards, but to the suspension or abatement of previous sanctions. Reprieves include dismissed charges or acquittal in court, release from detention, or a foreign government’s refusal of a request for extradition.

For each sanction, reprieve, or reward event, I code the primary agent of the event, in the following categories. For sanctions and reprieves, I code whether the legislature, government executive (e.g. the attorney general of a successor administration), judiciary, or foreign state were behind the act. For rewards, I code whether the former leader won a new position in (or honors bestowed by) a public office (in any branch of government), a private sector (or civil society) organization, or a foreign entity (including private and multinational organizations).

5.3 Structure of the post-tenure event data

I add the post-tenure fate variables to the project's dataset as monthly event counts, in the month after a former leader was sanctioned or rewarded. This means that each country-month observation in the dataset contains counts of different types of decrees issued by the incumbent president, along with counts of different types of sanctions or rewards that might have occurred to the incumbent's predecessors during the previous month. Lagging the post-tenure events ensures that developments in former leaders' fates occurred before incumbents issued the decrees that may be associated with those events, within a given observation in the dataset.

For instance, in December 2007, Peru's Alberto Fujimori received his first conviction for abuses of power. Alán García occupied the presidency at that time. The following month, in January 2008, García issued 61 total decrees, 2 of which are coded as enabling the government executive. The observation for Peru in January 2008 includes the counts of García's decrees, along with a 1 in the variable for former leaders sanctioned (for Fujimori's conviction from December), and a 1 in the more specific variable for sanctions by the court (for the same event). There were no reprieves or rewards of former Peruvian presidents in December 2007, so those variables contain zeros in the January 2008 observation.

If another former president (such as Alejandro Toledo) had also been sanctioned in December, or if Fujimori had received two convictions in the same month, the sanction event count variable would have been 2. If Fujimori or another former president had also received a reward in December, the count of reward events would have been at least 1. In other words, the post-tenure event counts are not strictly binary, although the events are sparse enough there are very few months in the dataset with more than one sanction, reprieve, or reward at a time. The fate variables are also not mutually exclusive. If a former leader was sanctioned and then reprieved within the same month, both variables would equal 1 for that month. If a sanction and a reprieve happened to two different former presidents in the same country and the same month, the data would look the same.

This aggregation is driven by the assumption that incumbents should react similarly to punishments of any predecessor, if those punishments themselves are comparable (i.e. taking into account more specific aspects of sanction events defined below). It might be the case that there are unobserved diadic factors involved in whether Fujimori believes García's fate to have been more relevant to his own than their mutual predecessor Fernando Belaúnde Terry's. But to test for different reactions to sanctions by predecessor would require further slicing of already sparse data, without concrete intuition about its necessity. I instead choose to treat all predecessors' sanction or reward events equally, but to break them down based on additional observable characteristics that might make some sanctions more comparable than others.

Finally, the event counts are not cumulative; they are only recorded once each in the baseline dataset. Most reprieves happen after some attempted sanction, but these events are not explicitly paired; they are each simply recorded in the country-month in which they occurred. On average, reprieves in the dataset occurred 26.6 months after the same leader's last sanction, with

the shortest interval between a sanction and a reward being one month and the longest being 97 months (or about eight years). If a former leader was sanctioned and then later reprieved, there would be a 1 for sanctions in the month in which the sanction occurred, then a 1 for reprieves in the later month in which the reprieve occurred, but zeros in between. My goal in the main analysis is to test for successor presidents' reactions to any sanction or reprieve of a predecessor, regardless of what else happened to those leaders. In Chapter 6, I will address combinations of post-tenure events and the consistency of each former leader's overall fate.

As mentioned at the beginning of this chapter, while a president might be investigated for corruption while still in office, I only include post-tenure events for *former* leaders, to reduce the risk of endogeneity between incumbents' actions and the concurrent retribution they might bring upon themselves. Therefore, for any given country-month observation, the leaders who faced the sanctions or rewards in the post-tenure event counts are always different individuals from the leaders who issued the decrees in the decree category variables.

In the baseline dataset, I have coded all post-tenure events as individual occurrences, rather than conditions that persist over a period of time. Most sanctions and reprieves only occur at one point in time, such as when charges are filed or dismissed. There are some types of rewards that might persist for a period of time, especially when a former president occupies another position in the public or private sectors. For consistency, I record all of these events one time, in the month in which the new job or appointment began, since that is when the award or accomplishment would be new news and might have the greatest influence on successors' expectations. In Section 5.7 and later in Chapter 6, I present alternative analyses using events recorded in every month in which they were in effect, for certain types of sanctions and rewards that do persist over a period of time. The multi-month event counts add more statistical power to the sparse data, but are generally consistent with the results for the single-month event counts.

5.4 Descriptive statistics of presidents' fates in dataset

Table 17 and Figure 17 below summarize the number of individual post-tenure fate events recorded for former leaders in each country in the dataset. The dataset contains a total of 133 unique post-tenure events occurring to former leaders, during years in which I also have data on their successors' executive decrees. There are 49 instances of formal sanctions undertaken against former leaders in the dataset, such as investigations, prosecutions, or convictions. There are also 19 instances of reprieves, such as dismissed charges, and 65 instances of rewards, such as securing a corporate or academic position, honor or award. These events occurred to 42 former presidents who were still alive during the periods for which I was able to collect decrees.

Table 17 describes the sparseness of the data. The first column shows the total number of events collected in each of the main categories of post-tenure fates. The second column shows the number of country-month observations with at least one event in each category, since some events occur in the same month. The third column shows the remaining country-months with

zero events. The sparseness of the one-time event counts in the baseline dataset means that the vast majority of country-months have values of zero in each of the post-tenure variables.

Table 17. Former leaders’ post-tenure events in the dataset, observation event counts

<i>Fate</i>	<i>Total events (sum of all country- month counts)</i>	<i>Country-months with at least one event (count > 0)</i>	<i>Country-months with no events (count = 0)</i>	<i>Country-months with count > 0 for lagged 12 month period</i>
<i>One-time event counts (recorded only at the start)</i>				
<i>Sanction</i>	49	44	1322	336
<i>Reprieve</i>	19	17	1349	143
<i>Reward</i>	65	58	1308	444
<i>Multi-month event counts (recorded in every month for their duration)</i>				
<i>Reward</i>	65 (unique events)	530	836	682

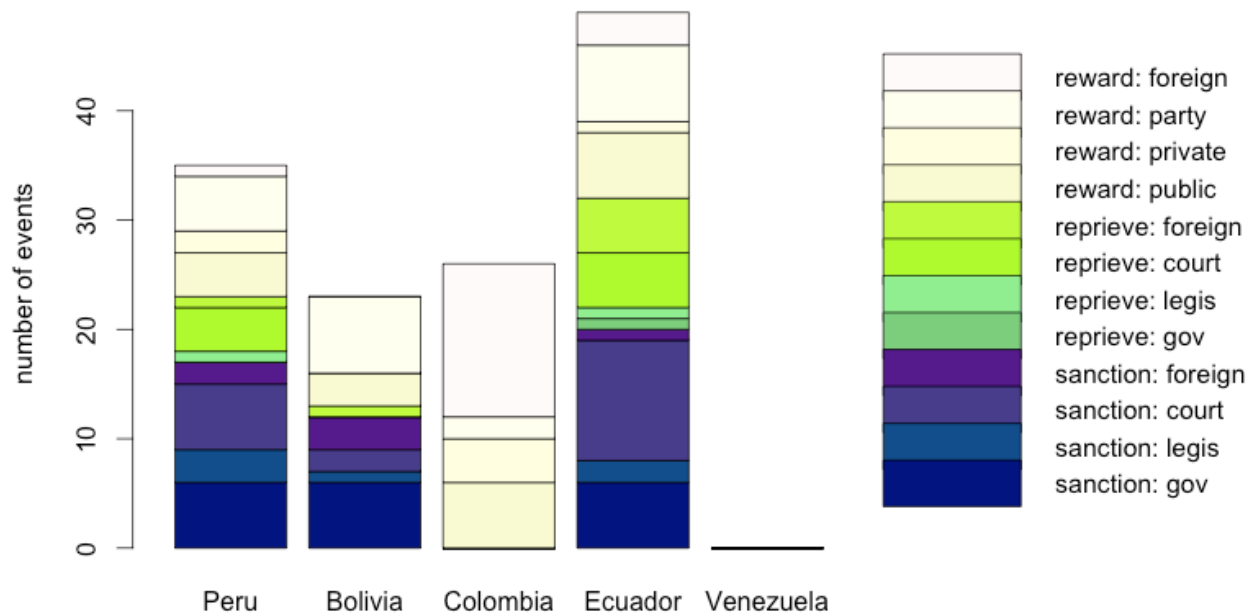
However, it is common to lag explanatory variables in time-series datasets, often by more than one period, if the theorized relationships are expected to play out over longer periods of time. Since the observations in this study are very short – one month instead of more common country-year observations in comparative political studies – we might expect to observe leaders’ reactions unfolding over more than one observation period. I also include tests of the hypothesized relationships when increasing the lag period for the post-tenure event variables up to 12 months before the decrees were issued in each observation. This maximum period is consistent with other studies that test the effects of explanatory variables one year prior to the outcome of interest. For the explanatory variables with longer lag periods, I average the event counts over the entire lagged window.

The last column in Table 17 shows the number of country-month observations with at least one post-tenure event, in each category, over the maximum lag period of twelve months prior to the observation month. There are now positive values in around 10% of country-month observations (for sanctions) to 33% of observations (for rewards). These lagged windows are intuitive and consistent with time-series data used in other studies, especially country-year studies. The varying lags also reflect the fact that, on the one hand, decrees are designed to be expedient forms of decision-making, which leaders might issue very rapidly, even within a single month. On the other hand, leaders may take time to adjust their expectations after a new event, and the new expectations may continue for a period of time, especially since sanctions and rewards of former leaders are infrequently observed. Lagging the explanatory variables therefore also helps to overcome the challenges of sparse data, extending recent post-tenure events to cover more observations in which those events might be associated with subsequent decrees.

Figure 18 shows the total number of unique sanctions, reprieves, and rewards by country, broken down by the main actors involved. The number of post-tenure events for each country in

the dataset is influenced by multiple factors, including how many presidents recently served in office and were still alive during the period in which I was able to collect their successors' decrees. Ecuador had the highest presidential turnover during the 1990s to early 2000s, and also had the highest combined total of post-tenure sanctions, reprieves, and rewards. While Peru had slightly fewer presidents than Bolivia (8 compared to 10) during the same period of study, Peru had one of the most contentious presidents (Fujimori), who is responsible for a large number (but not all) of the former leader sanctions in Peru's recent history.

Figure 18. Former leaders' post-tenure events, by country and by instigating actor



There is strong variation among former presidents' fates in three out of five of the countries in the study. In Colombia, no presidents in recent history have been removed early or formally sanctioned for illegal acts, although many were rewarded after departing office. The inclusion of Colombia allows me to address the relationship between different types of positive rewards and subsequent leaders' actions, including in countries in which there may be little overt threat of sanction. The large number of rewards received by former leaders in Colombia suggest that there is some tradeoff between sanctions and rewards, although some former leaders in other countries were both sanctioned and rewarded by different actors or at different points in time.

In Venezuela, several former presidents faced sanctions in the 1990s to early 2000s, but I was only able to access public archives of decrees since 2009, and no new sanctions or rewards of former presidents began after that date. Two former Venezuelan presidents who had earlier been tried for corruption were still living in exile at the start of the decree archive; one was able to return to Venezuela while the other died in the U.S. The limited data for Venezuela reflects

some of the challenges in this project, and is due to both limited public archives and inadequate turnover among recent presidents. Hugo Chavez had already been in office for a decade when the decree archive began, and later died in office, contributing no post-tenure events of his own.

I have left Venezuela in the dataset to make as much use as possible of the experiences of all five Andean countries. For the main hypothesis tests, the post-tenure event variables in the Venezuela observations are all zero. But Venezuela contributes to the analyses of decree categories and their relationships to other explanatory factors throughout this project, to help ensure a thorough analysis of the range of conditions in which leaders exercise their authority.

Regarding sanctions and rewards by specific actors, Figure 18 shows that the three countries with embattled former presidents included multiple sanctions by a successor government administration (e.g. arrest or prosecution), multiple sanctions by the judiciary (e.g. granting a hearing, delivering a conviction), and at least one sanction by the legislature (e.g. a legislative commission inquiry into unconstitutional acts), although the latter were less frequent than sanctions by the executive or judiciary. Peru and Ecuador experienced the most reprieves in the course of legal battles against former leaders, while Bolivia had one extradition request refused by a foreign country (which eventually acquiesced).

To conclude this discussion, sparse event data like those used here lend themselves to questions about how we should think about the sample size or number of relevant observations. As noted in Chapter 4, the initial data contain over 70,000 decrees, about 8,500 of which were classified as enabling the government executive. These decrees are aggregated by country-month, for a total of 1366 country-month rows in the final dataset, which serve as the main unit of analysis. These country-months contain variation from five countries and 29 different incumbent presidents, reacting to the potential fates of 45 predecessors who had served during the democratic period of study and were still alive during the years covered by available decrees. For the explanatory variables, a number like 19 total reprieve events might seem small, but I am able to observe how different leaders' reactions to those events varied over multiple months as other explanatory factors varied too. These different dimensions of variation offer detailed information about complex relationships as they unfolded across fine-grained intervals of time.

5.5 Statistical models and parameters for hypothesis tests

In order to combine the new dependent and independent variables into hypothesis tests, there are a number of additional decisions we need to make about model specifications and parameters, for which there are many reasonable options that could produce very different results. If these tests are constructed openly and transparently, the availability of multiple variable measures and other model components can be an advantage for demonstrating robust results. When researchers choose only a few models to present in analytic results, however, they face concerns that the results may have been selected to match prior expectations or achieve statistical significance, and that those results may not be reproducible (Nuzzo 2015).

Institutions – especially large and powerful central government institutions – are much more difficult to manipulate than individual behavior, so that institutional studies are usually observational rather than experimental. The number of cases and events that we are able to assemble may also be relatively small, since each country only has one president at a time, and the prosecution of former presidents requires considerable collective action and resources, making these events relatively infrequent. Given the limitations of modest-N observational studies, estimates from statistical models may vary widely with the choice of parameters and covariates, placing greater importance on robustness checks to ensure that a hypothesized relationship is clearly and consistently reflected in the available data and that other possible explanations have been thoroughly accounted for.

Modern computing power can help us address these concerns, by enabling us to test as many combinations of covariates and model parameters as possible, and to record and present the results in comprehensive ways. Doing so should produce more reliable indicators of whether the identified relationships hold across a relevant set of alternative specifications, or whether the relationships are more elusive and dependent on arbitrary parameters, which would be more suggestive of spurious correlations. Summarizing the distribution of regression estimates across a set of model options or parameters could also tell us more about the underlying relationship than we can learn from individual regression estimates alone.

I have chosen to set up my hypothesis tests using standard approaches to linear regression for cross-national time series data, including some combination of lagged explanatory variables, country and time fixed effects, and robust clustered standard errors (specific factors used are discussed below). Filling in this established framework, I have taken a comprehensive approach to the remaining choices that a researcher would need to make, to construct specific models for analysis. Since the number of possible combinations of control variables and other model parameters is very large, I randomly sample permutations, to estimate the distribution of results across as many relevant configurations as feasible.

For each independent variable, I construct 1000 regression models, randomly sampling the various components. First, I sample control variables to include in the regression with the independent variable, selecting the number of control variables from a uniform distribution in the range from zero to all of them. I include as potential control variables the factors used in the analysis in Chapter 4, which were derived from other explanations in the literature on presidential power and executive orders. Table 18 summarizes the control variables included, covering multiple measures of political opposition and gridlock, exogenous economic shocks, natural disasters, and remaining time left in the incumbent's term.

I have sought to only include plausibly exogenous control variables in this analysis, which is why I do not include alternative measures of legislative gridlock like the rate of passage of executive-sponsored bills, indicators of domestic economic performance or violent conflict, or public opinion polls. Some variables in Table 18 are quite similar, but the different specifications should help ensure that each alternative explanations are thoroughly accounted for.

Table 18. Control variables included in main regression models, by random selection

<i>Variable</i>	<i>Description</i>	<i>Source</i>
<i>gov seats in legislature</i>	Governing party's share of seats in the legislature	Database of Political Institutions (DPI), WB/IADB
<i>opp seats in legislature</i>	Largest opposition party's share of seats in the legislature	DPI
<i>party fractionalization</i>	The probability that two deputies picked at random from the legislature will be of different parties	DPI
<i>opposition party fractionalization</i>	The probability that two deputies picked at random from the legislative opposition will be of different parties	DPI
<i>export price change (1 month)</i>	Top three export commodities' mean price change from previous month	Database of Primary Commodity Prices, IMF
<i>export price change (3 month)</i>	Top three export commodities' mean price change over three months	IMF
<i>export price change (6 month)</i>	Top three export commodities' mean price change over six months	IMF
<i>export price change (year-on-year)</i>	Top three export commodities' mean price change over past year	IMF
<i>major earthquake deaths</i>	Count of fatalities from major earthquakes	Global Significant Earthquake Database, NCEI/NOAA
<i>major earthquake injuries</i>	Count of injuries from major earthquakes	NOAA
<i>major earthquake damage (mil USD)</i>	Property damage in millions USD from major earthquakes	NOAA
<i>major earthquake score, sum of codes</i>	Sum of codes for earthquakes, deaths, injuries, and property damage	NOAA
<i>months left in office</i>	Months left in the incumbent president's mandated term in office	Coded from Wikipedia entries specifying presidential terms
<i>eligible for reelection</i>	Whether the president is constitutionally eligible for reelection	Coded from each country's constitutional history

Next, I randomly select the variable on which to cluster standard errors, using either the presidential term or the overall country, since observations are consecutive in time within the same country and some unobserved factors may be correlated even across presidential terms. I also randomly select whether to include month fixed effects in addition to country and year fixed effects, in case there is any seasonality to the relevant trends. I include at least country and year

fixed effects in all models, to account for the fact that the data consist of cross-national time series panels. Naturally, a large portion of the variation across all observations is explained by the country in which the leader is issuing decrees, since each country has different constitutional provisions for decree-making authority, as well as a range of other subtler variations in the political culture of decision making. Trends in the use of decrees have also evolved over time.

Finally, I select a period over which to the explanatory variables, in months prior to the country-month in which the decrees were issued, and average the counts of post-tenure fate events over that window. I use a single moving average for each sanction, reprieve, or reward variable, instead of constructing separate variables for each lagged month (i.e. a separate variable for post-tenure events at $t-1$, $t-2$, etc.), since we would expect these lagged variables to show extensions of the same relationship. For each regression model, I randomly select a window of n months from a uniform distribution from 1 to 12 (i.e. a full year, as used for lagged variables in country-year studies). I use the average number of post-tenure fate events in each category (i.e. sanctions, reprieves, and rewards, all targeting former leaders) over the n months prior to each observation period t , as the explanatory variables for the given model.

As mentioned above, I chose lag periods from 1 to 12 months due to the expedient nature of decrees, which may be issued very rapidly after a leader's expectations change, but also the infrequency with which post-tenure events are observed and those expectations may be updated. In this project's dataset, the time between any two post-tenure events occurring to former leaders within the same country ranged from less than one month to (in a few outlier cases) several years. About 80% of all post-tenure events in the dataset were succeeded by another sanction, reprieve, or reward of a former leader in the same country within one year, suggesting that in most cases, 12 months should be enough time for incumbents to receive new information about the political environment and potential risks they will face after departing office.

For each of the 1000 regression models assembled and run for a given independent variable, I record the independent variable's estimated coefficient, standard error and p-value. In the results presented below, I report average coefficients and p-values across all 1000 regression models for each independent variable. I also plot histograms of the coefficients and p-values, to show the distribution of results across the different models run. This approach is both more thorough and more concise than presenting individual estimates for a limited number of selected models. Given the observational nature of the data, the condensed tests of robustness to alternative covariates and model specifications can help eliminate the possibility of spurious correlation. If any explanatory variable produces consistently significant results across the vast majority of regression models, and if all other reasonable explanations have been included as potential control variables, the aggregated results should constitute strong evidence of an empirical relationship which is unlikely to have been driven by some other cause.

These model permutations provide an alternative to existing methods for helping researchers select variables to include in regression models. A prominent option is to use least absolute shrinkage and selection operator, or Lasso, regression. Lasso imposes a constraint on the absolute value of the model coefficients, which forces some coefficients to equal zero,

producing a simpler model in which only those variables with non-zero coefficients would be included (Tibshirani 1996). Lasso may be an appropriate means of selecting control variables, but would be less useful for selecting other model parameters, such as fixed effects, clustering factor, and lag periods. It also assumes that one model is the correct one, and does not provide a means to analyze how regression estimates vary with different model parameters. An alternative that does allow for multiple regression models is Bayesian model averaging, which requires calculating model weights based on their likelihood, assuming some are closer to the truth than others. My approach is similar in principle, but simpler and more straightforward to implement, with a goal of showing the *full distribution* of estimates across potential regression models.

For computational simplicity and substantive interpretability, I only include one post-tenure explanatory variable in each model, along with the randomly sampled control variables and fixed effects. Post-tenure sanctions and rewards are not competing explanations, in which our goal would be to test which explains more of successors' decree-making decisions when both are present. By testing one type of post-tenure event at a time, we can see how leaders respond to each type of event on its own, in contrast to that event's absence, but without incorporating other concurrent events into the same model. In country-months in which the count of former leaders sanctioned is zero, there might be other things happening to former leaders, but the distribution of those other events is not assumed. We simply observe how leaders react to a sanction event, in comparison to how they behave on average when there is no sanction event, across the general distribution of other potential events that are not included in the given model.

To justify this decision, in the first set of tests in Figures 18 and 19 below, I show results for the individual explanatory variable models, as well as comparable results when including all three main variables for sanctions, reprieves, and rewards in the same model. The results are almost the same, suggesting very little correlation between the different post-tenure events within the same country-month observations. Sanctions, reprieves, and rewards are not independent, in terms of which leaders are subject to one or the other at some point after departing office. But the timing of a sanction does not appear to be influenced by the timing of another reprieve or reward. Since the results do not change when including more than one explanatory variable at a time, and the interpretation is most straightforward when including only one per model, I proceed to use individual explanatory variable models for the remaining tests.

5.6 Results for hypothesis tests, baseline post-tenure event categories

Tables 19 presents an overview of the results for regressing power-consolidating decrees on the main categories of post-tenure sanction, reprieve, and reward. For a baseline summary, the first column shows estimates from bivariate regressions of power-consolidating decrees on each post-tenure event variable individually, with no control variables included. For the baseline bivariate regressions, I include country and year fixed effects, cluster standard errors at the presidential term, and average post-tenure events over a full year lag window. The second

column summarizes the results of the more robust model permutations, including randomly sampled control variables, optional month fixed effects and country-level clustered standard errors, and lag windows varying from one to twelve months. The coefficients and standard errors in the second column are averaged over 1000 models selected for each independent variable.

Table 19: Regression estimates for main hypothesis tests

Dependent variable: power-consolidating decrees, independent variables: post-tenure fates

Post-Tenure Fate Variable	Estimate 1 (bivariate)	Estimate 2 (sampled covariates)
<i>sanctions</i>	1.397 (3.474)	0.417 (1.700)
<i>reprieves</i>	12.018 *** (3.556)	6.283 ** (2.387)
<i>rewards</i>	-0.035 (3.289)	0.786 (1.540)
<i>sanctions by legislature</i>	14.130 (9.710)	4.377 (5.163)
<i>sanctions by government</i>	0.874 (6.718)	0.006 (3.974)
<i>sanctions by judiciary</i>	1.145 (5.240)	1.294 (3.410)
<i>sanctions by foreign</i>	-7.010 (9.122)	-4.069 (5.088)
<i>reprieves by legislature</i>	6.652 (9.266)	-6.689 (4.920)
<i>reprieves by government</i>	44.583 *** (12.819)	22.120 ** (8.382)
<i>reprieves by judiciary</i>	19.496 *** (6.403)	11.320 ** (4.339)
<i>reprieves by foreign</i>	16.657 ** (8.481)	6.826 (3.959)
<i>rewards by public</i>	5.734 (4.521)	4.520 (3.232)
<i>rewards by private</i>	-15.909 * (8.288)	-9.352 (4.727)
<i>rewards by party</i>	6.419 (4.172)	3.296 (2.584)
<i>rewards by foreign</i>	-3.107 (5.659)	0.231 (2.613)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

All of the post-tenure variables in these tests are counts of one-time events recorded in the month after they occurred or began. The outcome variables are also raw counts of power-consolidating decrees issued in each month. The coefficients can be interpreted as the effect of one post-tenure sanction (or reprieve or reward, respectively) occurring to a former president, in terms of how many more (or fewer) power-consolidating decrees incumbents issue in the following one to twelve months, as opposed to how many they would have issued had the former leader not faced that post-tenure event.

At first glance, of the three high-level categories of post-tenure events in the top section of Tables 19, only former leaders' reprieves have a statistically significant average p-value across all randomly sampled models. Looking at the categories broken down by the instigating actor, we see some more significant relationships for reprieves, and at least one relationship marginally significant for rewards, but still none for sanctions. Since the second column only shows average estimates across many regression models, it does not tell us how many individual models produced significant results, or how consistent the coefficients were across models. In the following subsections, I drill down into each set of tests in greater detail.

To explore the distributions behind the average values reported in Table 19, the next figures include histograms of coefficients and p-values across the sampled models for a given independent variable. For comparison, I have stacked the same figures for several related or contrasting independent variables at a time. In each row of figures, when the distribution of coefficients (center) and/or p-values (right) is more spread out, this indicates that the estimates varied more with the inclusion or exclusion of different control variables, fixed effects, and/or clustering factors. Independent variables with more widely varying regression estimates might simply be more correlated with the potential control factors, or their relationship to the dependent variable might be more spurious or indirect. When the coefficients and p-values are more tightly grouped, especially when the coefficients are large and all positive or all negative, and when the p-values are concentrated close to zero, the relationship to the dependent variable appears to be strong and robust to different model parameters.

Figure 19 below shows the distribution of estimates for the high-level categories of *sanction*, *reprieve* and *reward*, including only one post-tenure variable in each regression model. Figure 20 on the same page shows the same tests including all three variables in the same model (along with the same randomly sampled options for control variables, fixed effects, clustered standard errors, and lag periods). The results are nearly identical across these two sets of tests, indicating that there is very little correlation between sanctions, reprieves, and rewards in the dataset. There seems to be little reason why a sanction should (or should not) occur at the exact same time as a reprieve or reward, and the sparseness of the data probably contributes to the lack of observed correlation over country-month observations.

Figure 19. Summary of estimates, separate models for sanctions, reprieves, and rewards
Dependent variable: decrees enabling executive, 1000 models for each independent variable

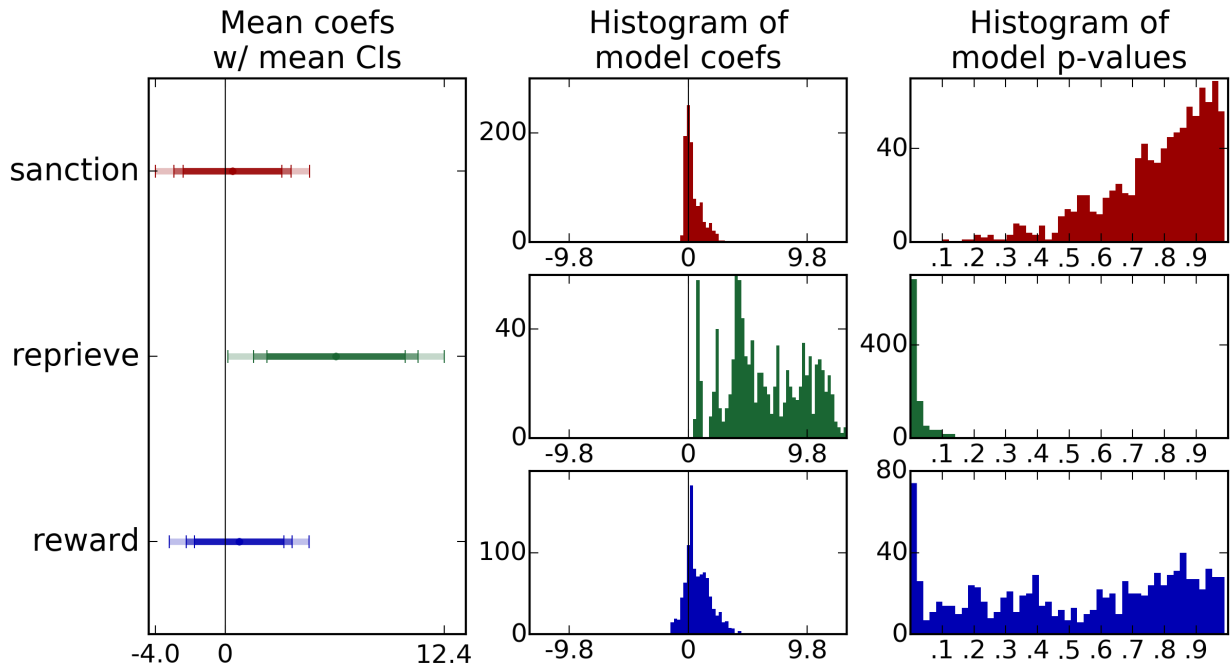
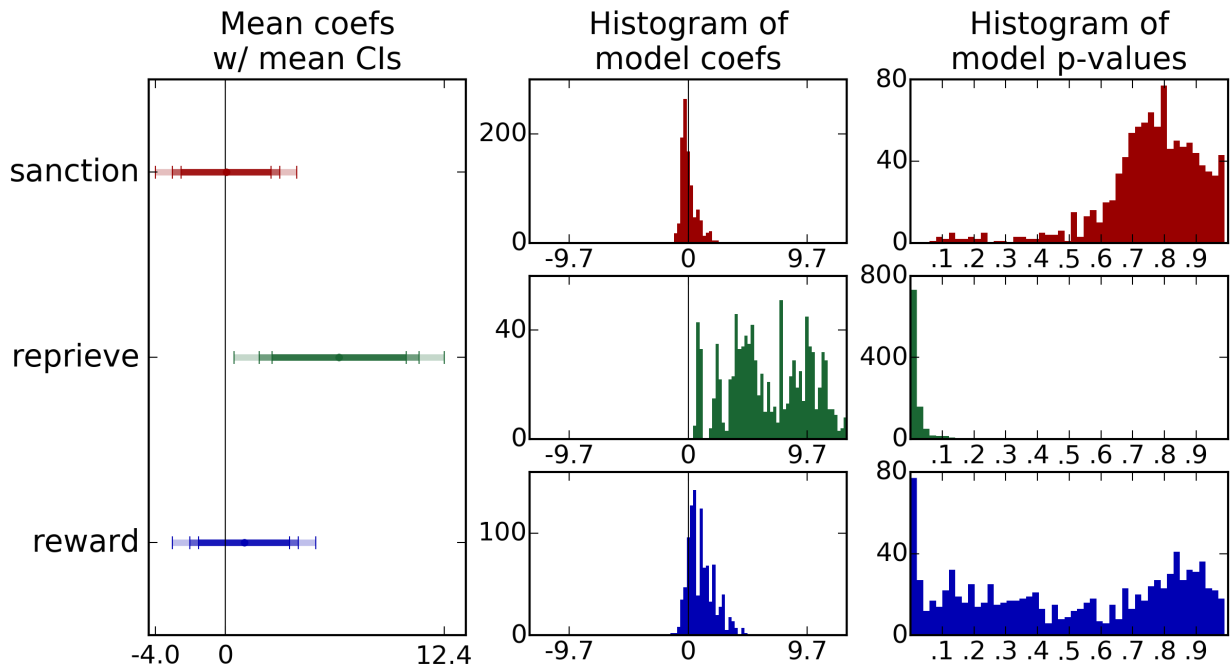


Figure 20. Summary of estimates, combined model for sanctions, reprieves, and rewards
Dependent variable: decrees enabling executive, 1000 models for each independent variable



There are notable differences in the distribution of estimates across these three event types. For post-tenure sanctions, the coefficients are clustered at zero, and the p-values are grouped at the higher end of the spectrum, indicating no significant effect. For reprieves, however, the coefficients are all positive and much larger on average, and the p-values are stacked tightly toward zero – almost all fall below 0.05. These results support the hypothesis that reprieves of former leaders encourage their successors to consolidate power.

On average, one additional reprieve of a former president is associated with about six additional power-consolidating decrees issued by a successor president over the subsequent one to twelve months (the lag periods sampled across regression models). The coefficients do range in size, and there might be differences in magnitude across different types of reprieves, but all models show highly significant relationships to successors' consolidation of power. As discussed in Chapter 2, reprieves – even as individual events – indicate at least some uncertainty or back-and-forth between sanctions and reversals of fate, which appears to encourage incumbents to strengthen their offices against rivals and protect themselves against potentially politicized struggles over their own future fate.

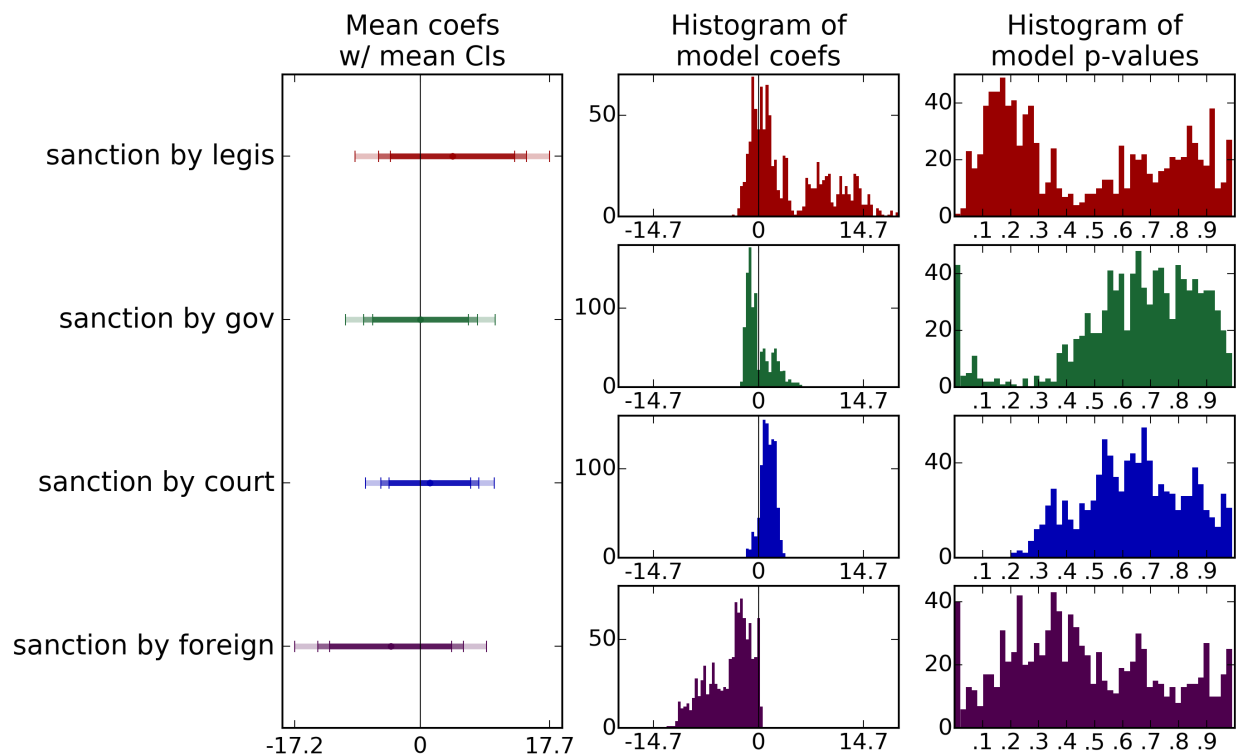
For rewards, there are coefficients with both positive and (some) negative values, and the p-values are widely spread out across the full spectrum, indicating no clear or robustly significant relationship to power-consolidating decrees. However, unlike with sanctions, there is a considerable spike – about 10% of statistical models sampled – that have p-values below 0.05. In Chapter 2, I expected the relationship between former leaders' rewards and successors' power-consolidating decrees to change depending on whether the reward was in the public or private sector, which might be responsible for the lack of a clear and consistent relationship for the overall rewards category. I turn next to events in the same three post-tenure fate categories, broken down by instigating actor.

5.7 Results for hypothesis tests, post-tenure events by instigating actor

Figure 21 shows the same representation of results, when regressing power-consolidating decrees on sanctions broken down by instigating actor. None of the categories of sanctions by different actors show significant relationships to subsequent leaders' actions; the p-values in are all spread out across the full range from 0 to 1. The coefficients on sanctions by domestic actors (the legislature, government, or judiciary) are generally clustered around zero (i.e. no effect). The coefficients on legislative decrees do range into larger positive values, which would be consistent with the hypothesis that sanctions by more partisan or politically motivated legislative bodies would encourage subsequent leaders to consolidate power, in order to protect themselves from their own political opponents. However, only a minority of models suggest a noticeably positive relationship, and with no statistical significance across sampled models.

Figure 21. Summary of regression models for sanctions by actor category

Dependent variable: decrees enabling executive, 1000 models for each independent variable



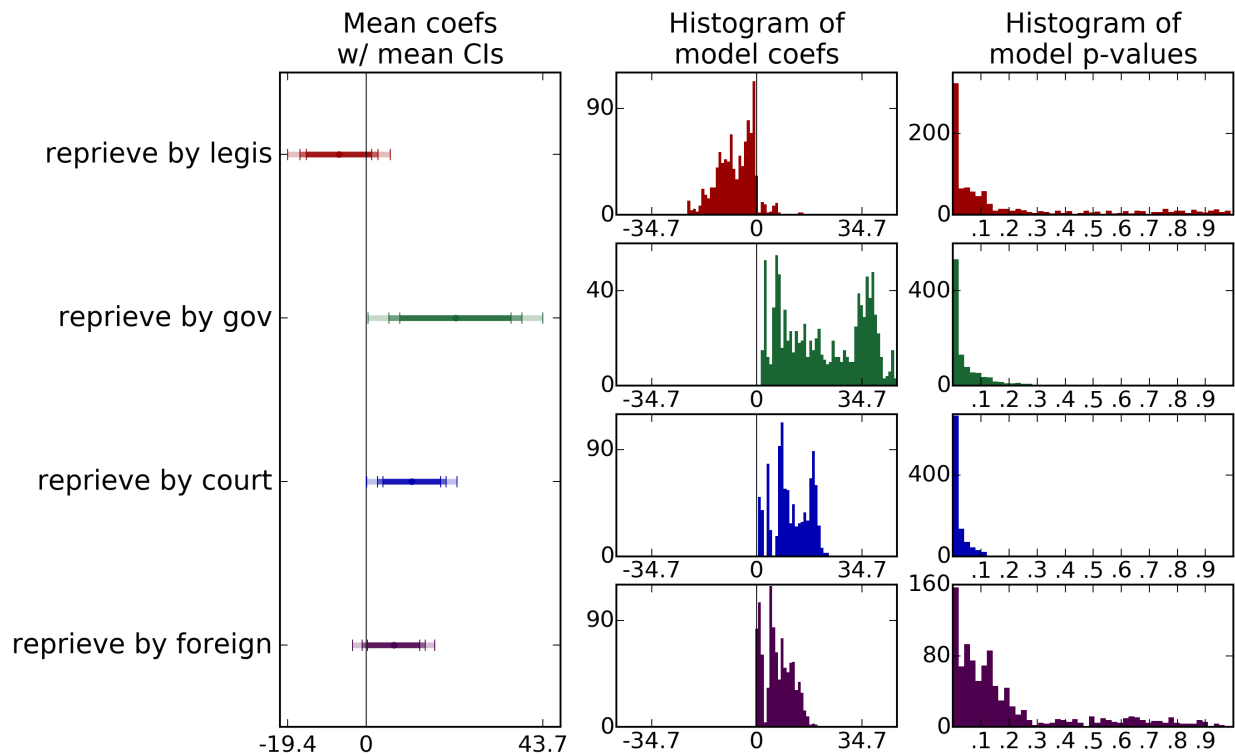
The coefficients on foreign sanctions are almost all negative and extend to large values in the opposite direction. The average coefficient is about -4, suggesting that one additional sanction of a former leader by a foreign government would be associated with a successor issuing about four fewer power-consolidating decrees over the next one to twelve months, than the successor would have done in the absence of any foreign sanction of a predecessor. This result would be consistent with the hypothesis that foreign sanctions are more likely to deter subsequent consolidations of power.

In the cases included in this project, sanctions by foreign leaders are almost entirely comprised of extraditions back to the home country, since all criminal cases against former leaders in the dataset were carried out in domestic courts. As hypothesized in Chapter 2, sanctions by foreign actors that defer to domestic authorities and reduce the possibility of escaping justice abroad, might discourage the consolidation and abuse of power by subsequent leaders. For instance, U.S. refusals to extradite leaders like Bolivia's Gonzalo Sánchez de Lozada have been widely seen as enabling corrupt politicians to use their personal connections to escape justice, while Chile's extradition of Alberto Fujimori back to Peru was less controversial and seen as supporting the home country's pursuit of justice. While the coefficients on foreign sanctions are all negative (consistent with my expectations), as with sanctions by other actors, the vast majority of sampled models are not statistically significant.

Figure 22 shows the distribution of regression results for reprieves by different actors. Here, all categories show much stronger relationships than for sanctions. Again, I expected reprieves to have the strongest (positive) relationship to power-consolidating decrees among the basic post-tenure fate categories. Reprieves by government actors (such as release from prison or dropped charges), by the courts (such as dismissal of a case for insufficient evidence), and by foreign governments (such as a denial of extradition) support the hypothesis. One reprieve of a former leader by a government agency is associated on average with over twenty additional decrees enabling executive offices, issued one to twelve months later. A reprieve by a court is associated with about a dozen additional power-consolidating decrees on average, while a reprieve by a foreign actor is associated with closer to 7 more power-consolidating decrees.

Figure 22. Summary of regression models for reprieves by actor category

Dependent variable: decrees enabling executive, 1000 models for each independent variable



The p-values for reprieves by government actors and courts are tightly stacked in the lowest bin, indicating highly significant relationships. This suggests that when the authorities traditionally responsible for prosecuting and convicting individuals for criminal acts choose to reverse those decisions (or deny each others' sanction attempts), subsequent leaders are more likely to consolidate power, to be able to protect themselves in contentious or uncertain legal battles of their own. The p-values for reprieves by foreign actors are a bit more spread out,

indicating a less robust relationship, although about half still fall below 0.1. This is consistent with the suggestive but insignificant relationship observed for foreign sanctions, where the coefficients had the opposite sign. When foreign governments do not comply with sanctions, but instead protect embattled former leaders, granting them asylum or refusing extradition, this may signal to subsequent leaders that they can escape justice for their acts, as long as they use their authority to secure powerful allies while they can.

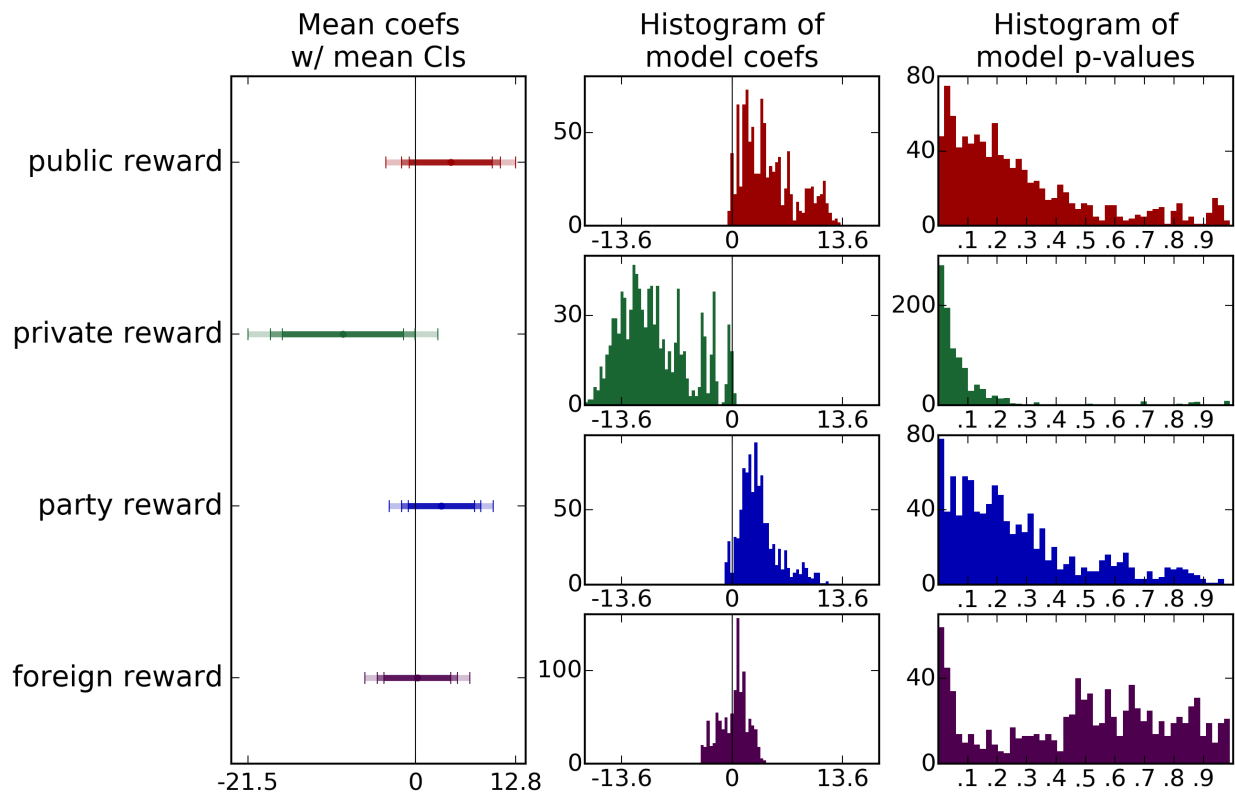
Reprieves by the legislature are the only category with negative coefficients in Figure 22, suggesting that reprieves by the legislature actually discourage successor presidents from consolidating power. The relationship is not quite significant on average across all sampled regression models, but almost half of the p-values fall below 0.05, with a thin but long tail over the rest of the spectrum. I did not anticipate this relationship in Chapter 2, and I put less weight on this finding, since sanctions and reprieves by the legislature are much less common in the dataset than sanctions or reprieves by the government or courts. However, there are reasons why we might find the negative relationship intuitive as well. The legislature is not responsible for legally prosecuting individuals for criminal acts, and can only initiate and terminate their own political investigations. If legislative sanctions appear to be the most politically motivated, then a decision to withdraw or terminate a legislative inquiry into misconduct (i.e. a reprieve) might be seen as alleviating a situation of politicized vengeance against a former president, which might make subsequent leaders *less* concerned about politicization of their own fate.

Wrapping up post-tenure fate events by instigating agent, Figure 23 shows the distribution of results for rewards by different actors. The strongest results appear for private rewards. The p-values are grouped closest to zero, with about half falling below 0.05 and the majority falling below 0.1. The coefficients on private rewards are all negative and most are large. One former president earning a prestigious position or honor in the private sector or civil society, is associated with an average of around ten fewer power-consolidating decrees issued by his/her successor over the subsequent one to twelve months. These results support the hypothesis that when former leaders are able to secure other opportunities outside the state, subsequent leaders should be encouraged to exercise restraint, avoid clinging to power and instead focus on building a positive legacy.

While none of the other categories of rewards show significant results across the majority of sampled models, the p-values are grouped lower than they were for the sanction categories in Figure 20, and the coefficients show some expected trends. In particular, the coefficients for rewards in the public sector and in political parties are almost all positive. This is consistent with the hypothesis that when former leaders retain other positions in government, incumbents may seek to expand their personal control while they remain in the presidency, in order to help them secure and wield influence from other jobs within the state later on. The coefficients suggest that a public or party reward of a former leader might be associated with an increase of three to five power-consolidating decrees on average issued by a successor president, although again these results are not significant across most models.

**Figure 23. Summary of regression models for rewards by actor category
(Recording reward events only in the month in which they began)**

Dependent variable: decrees enabling executive, 1000 models for each independent variable



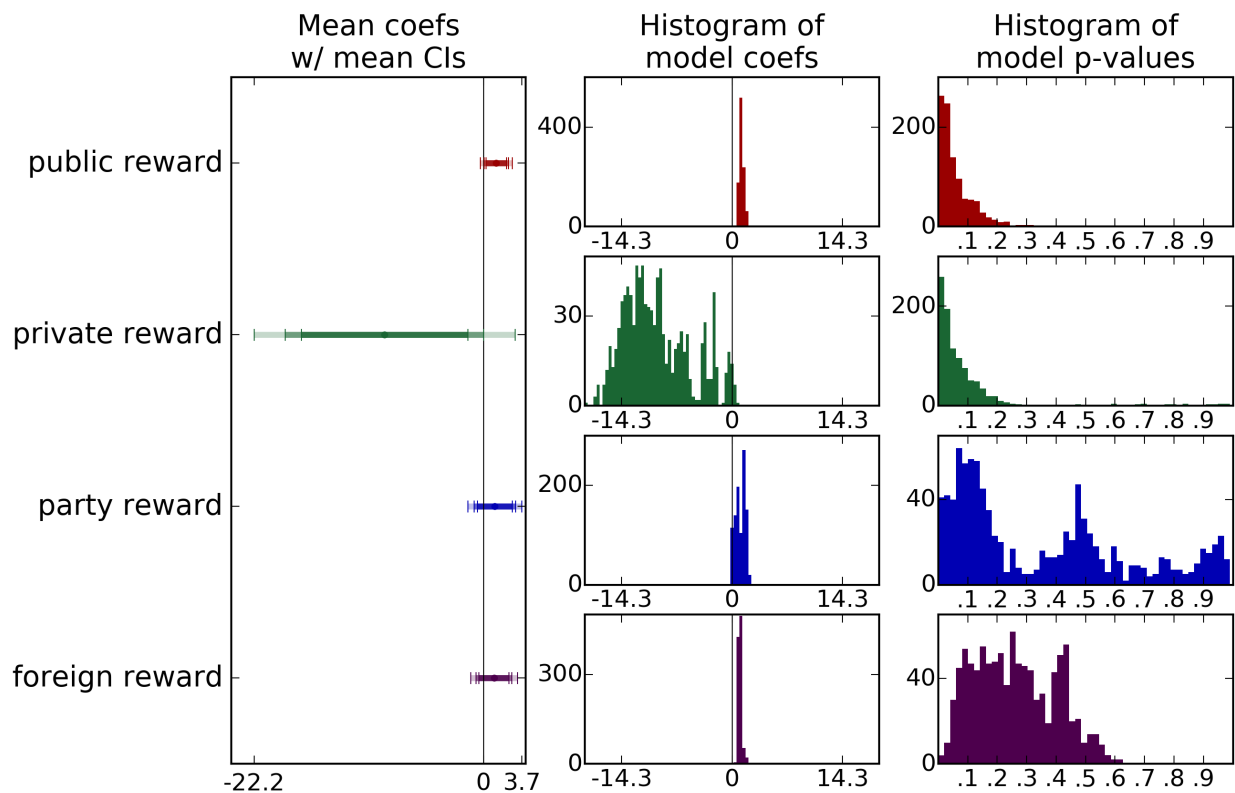
In the explanation of the coding scheme and resulting post-tenure event data, I mentioned that all events in the baseline dataset are recorded as one-time occurrences in the month in which they were enacted or began, but that rewards sometimes persist over longer periods of time. In Figure 24, I show the same analysis as in Figure 23, except that I count post-tenure rewards in every month in which they remained in effect. For instance, Colombia's Andrés Pastrana ended his presidency in 2002 and served as Colombian Ambassador to the U.S. from October 2005 to July 2006. In the baseline dataset, I recorded this public reward (an official government position) only in the country-month observation for November 2005 (the month after the appointment began). In the analysis in Figure 23, I include a 1 in the public rewards category for every month between November 2005 and August 2006 (one month after the appointment concluded).

The results are very similar to those in Figure 23, except that the three positive sets of coefficients have gotten smaller and tighter, and the p-values on public rewards have become more significant. The p-values for public rewards now look similar to those for private rewards, with about half of the p-values falling below 0.05 and most falling below 0.1. The coefficients are all positive (if small), suggesting that if a former leader is currently serving in another public office in a given month, this is associated with an increase of about one additional power-

consolidating decree issued by the successor president. The coefficients may be smaller than in Figure 23 because a leader’s reaction to a predecessor’s reward diminishes over time, once the news has set in and the incumbent has enacted the additional power-consolidation desired. Overall, these results support the hypotheses that former leaders retaining positions within the state encourages successors to consolidate power, while former leaders retiring to non-state positions instead encourages successors to exercise restraint.

Figure 24. Summary of regression models for rewards by actor category
(Recording reward events in all months in which the reward was in effect)

Dependent variable: decrees enabling executive, 1000 models for each independent variable



The coefficients on foreign rewards are also now all positive, although the p-values are still spread out and generally insignificant. Rewards by foreign actors are mainly granted by private organizations, such as leadership or advisory positions in foreign corporations and universities. The lack of a significant relationship might reflect a conflict between incumbents’ reactions to private sector rewards – since foreign rewards might reduce leaders’ desire to cling to power within the state – and incumbents’ reactions to their predecessors going into exile – which might signal that leaders can escape justice and encourage them to abuse power.

5.8 Summary of analysis for general categories of post-tenure events

In general, the estimates in the figures above offer some interesting and useful results, especially for reprieves and rewards. Reprieves show the strongest support for the proposed theory, indicating that a reversal or setback in a case against a former leader is significantly associated with successor presidents issuing more power-consolidating decrees. Interestingly, reprieves by the legislature show the opposite relationship, although there are relatively few of those events in the dataset. While unexpected, that result is also intuitive, if we consider that legislative investigations into wrongdoing may already appear politicized (when first initiated as a sanction), and that a legislature's decision to drop such an investigation might indicate that a campaign for politicized vengeance has abated.

Rewards of former leaders showed somewhat weaker or less robust relationships to successors' consolidation of power, but the coefficients are generally consistent with the proposed hypotheses, and the greater challenge pinning down a clear relationship may relate to the fact that rewards often play out over longer periods of time. Event data can be constructed in different ways; in the main analysis, I sought to use one-time event counts consistently across the different types of post-tenure fates, since most sanctions and reprieves only occur at a single point in time. The alternative analysis of post-tenure rewards in Figure 24, in which each reward is recorded in all country-months in which it remained in effect, adds stronger support for the proposed hypotheses. In particular, former leaders remaining in the public sector appear to be associated with increased power consolidation by successors, while former leaders retiring to private positions or rewards are associated with successors issuing less power-consolidating decrees (potentially to focus more on building positive legacies).

However, we still have not arrived at any clear relationships for sanctions against former presidents, and the prosecution of former leaders is the main subject of debate at the heart of this project. The results in this chapter suggest why that debate may be so strong. As discussed in Chapter 2, proponents argue that punishing former leaders for abuses of power will deter successors from doing the same, yet we do not see this relationship in any of the categories of post-tenure sanction events, except possibly with regard to foreign sanctions. Other observers express concerns that prosecutions of embattled leaders – which are often politicized – may instead encourage incumbents to cling to power to protect themselves. Again, we do not see this relationship clearly reflected in any of the post-tenure sanction categories so far. If both reactions are occurring in different cases, cancelling each other out in the overall dataset, we have not yet disentangled them to figure out which occurs when.

However, there is more that can be done with the data collected so far. We have not yet addressed the more latent factors at the heart of the proposed theory: the objectivity or predictability of former leaders' sanctions may play a role in their consequences. In the second part of the empirical analysis, in Chapter 6, I dig further into post-tenure sanctions, introducing a series of additional steps to break down different aspects of sanction events or their context, to get closer to the underlying expectations at the heart of my theory.

Chapter 6: The Role of Context, Consensus, and Predictability in Post-Tenure Sanctions

As stated in the introduction and theory chapters, I am not only interested in how many former leaders were sanctioned, but in the ways in which those sanctions appear to be connected to the accused former leaders' actual decisions and actions. I am especially interested in whether a former leader's fate appears to have been objectively determined based on his/her actions in office, or reflects politicized retribution by rivals regardless of the former leader's innocence or guilt. It would be very difficult to measure the fairness of particular criminal cases in any indisputable way. Instead, I have sought to identify observable signals of how much consensus or contention there was surrounding each former leader's fate, and how well actual post-tenure events match predicted probabilities of sanction or reward, which might influence successors' expectations about how certain or manipulable their own fates will be.

6.1 Data on more specific sanction events within the legal process

As discussed in Chapter 2, punishing former leaders for abuses of power often involves a lengthy legal process that may play out for many months or even years. Different stages in that process might send different signals to subsequent leaders, about the fairness or objectivity of the process, and the certainty of the accused leader's fate. When collecting data on former leaders' fates, I included additional variables for specific types of sanctions that fall toward the beginning and end of the legal process. In particular, I recorded events in which former leaders were detained or incarcerated, and distinguished pretrial detention from a prison sentence associated with a criminal conviction, since pretrial detention might seem more arbitrary and dependent on whoever controls the relevant law enforcement agencies. I also recorded events in which former leaders were formally tried and convicted of criminal charges, which usually occurs much later, event years after initial arrests are made and charges are filed.

To get more leverage out of these process-based events, I analyze them both as one-time event counts (i.e. recording detention or trial only in the month in which they began), and as multi-month event counts recorded in every month in which they were in effect (i.e. recording the same events in all country-months for the duration of a former leader's detention or trial). I also include two variables for former leaders fleeing into exile, which might occur in reaction to accusations or warnings of imminent prosecution, since former leaders are less likely to escape abroad after they have already been arrested, detained, and tried. Exile is also a type of sanction that can be recorded either at the start, or for its duration. Including these variables provides more consistency with studies that used the Archigos dataset, which codes former leaders' fates based on whether they survived or were killed, imprisoned, or exiled.

In my initial exploration of presidents’ biographies, I found it difficult to determine whether some former leaders were technically evading justice in exile or simply seeking better opportunities elsewhere. Many former leaders of developing countries have gone on to fill prestigious roles in other societies, such as consulting for foreign corporations or obtaining fellowships at foreign universities, while rumors circulate back home that if they were to return, they would face prosecution for corruption or other abuses of power. Some eventually do face charges and requests for extradition, even if they did not request asylum protections before such charges came through. I therefore include all leaders who were living abroad at some point after departing office, other than for a brief visit (e.g. for medical care) or serving in an official capacity for their home government (e.g. as ambassadors).

6.2 Results for post-tenure sanctions at different stages of legal process

For the analysis in this chapter, I use the same process used in Chapter 5, randomly sampling 1000 regression models for each independent variable, and including the same control variables, fixed effects, clustering factors, and lag periods as in Sections 5.5 to 5.7. Table 20 shows the average coefficients and p-values for the new regression tests, using the process-oriented sanction events that typically occur toward the beginning and end of criminal proceedings against a former leader. In the upper half of the table, I use one-time event counts, recording exile, detention, and trials only in the month in which they began. In the second half of the table, I use multi-month event counts, recording exiles, detentions, and trials in every month in which they were in effect.

Table 20: Average regression estimates for decrees enabling government executive offices
Mean coefficients and p-values from 1000 regression models for specific types of sanctions

Variable (one-time event counts)	Mean coefficient	Mean p-value
<i>former leaders exiled/abroad (start month)</i>	4.061	0.389
“ <i>pretrial detention (start month)</i>	10.520	0.020 **
“ <i>on trial (start month)</i>	4.518	0.403
“ <i>convicted (one-time event)</i>	-1.846	0.805
Variable (multi-month event counts)	Mean coefficient	Mean p-value
<i>former leaders exiled/abroad (all months)</i>	-0.318	0.673
“ <i>pretrial detention (all months)</i>	2.173	0.009 ***
“ <i>on trial (all months)</i>	-1.973	0.035 **
“ <i>on trial (months since start)</i>	-0.119	0.024 **

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

In the top half of Table 20, for one-time event counts, the early-stage events have positive coefficients and final convictions have a negative average coefficient, although only pretrial detentions are significant. In the lower half of the table, for multi-month duration event counts, pretrial detentions have become even more significant (again with a positive coefficient) and former leaders on trial have also become significant (with a negative coefficient). To drill into what these numbers mean in greater detail, Figures 24 and 25 below show the full distribution of regression results across the randomly sampled model specifications. Figure 24 shows the results for one-time event counts, while Figure 25 shows the results for multi-month event counts.

The rows in both figures show these different types of sanctions in the order in which they usually occur, with leaders potentially escaping into exile before arrest or trial, then (for those who have not escaped) arrest and pretrial detention, which may eventually lead to a trial and conviction. (There are often many other steps in between, such as an investigation, filing of criminal charges, preliminary hearings, and various legal motions to bring about a formal trial, but those events are less consistently reported and harder to assign to one stage of the legal process.) In Figure 24, from the top to the bottom row, the coefficients generally move from more positive to more negative values. This is consistent with the hypothesis that earlier steps toward punishing former leaders tend to appear more arbitrary and politically motivated, while later steps entail higher burdens of proof and more collective consensus, suggesting more objectivity and certainty of the former leader's fate.

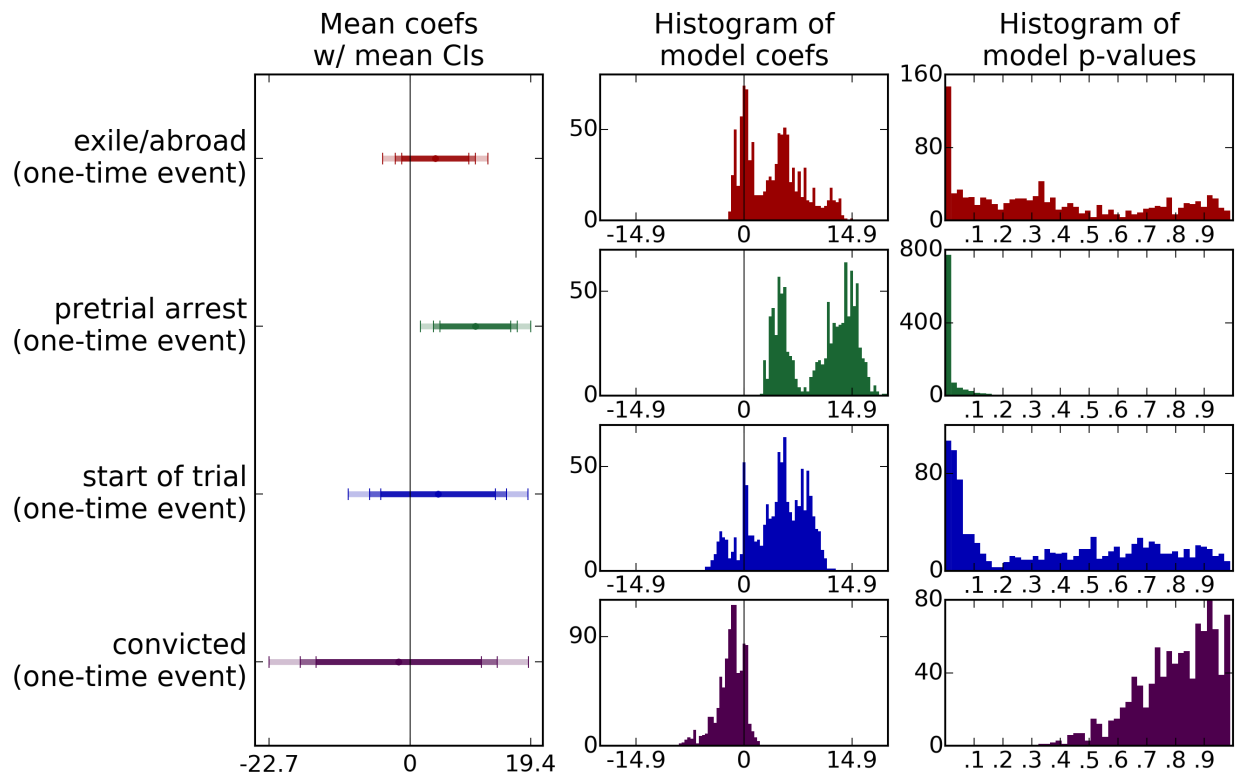
In Figure 25, only the pretrial arrests are highly and robustly significant across model specifications, although two of the other variables show some significant models. For pretrial arrests, the coefficients are all positive and the p-values are stacked tightly in the lowest bin, with almost all p-values falling below 0.05. On average, a single pretrial arrest of a former president is associated with about ten additional power-consolidating decrees issued by the successor president in the following one to twelve months. While the average p-value is not significant for the other process-based sanction events, the distribution of p-values for exile shows that the largest concentration of p-values does fall in the bin closest to zero, although with a large tail spread across the remaining range. The coefficients are almost all positive, which is consistent with the hypothesis that former leaders fleeing into exile may encourage successors to consolidate and abuse power, if they believe they can escape justice. The lack of robust results across all models may reflect greater noise in this post-tenure fate variable, since I included all leaders living abroad, due to the difficulty of identifying which were actually fleeing justice.

The coefficients for the start of a formal criminal trial are spread across both negative and positive values, but on average are weakly positive. This suggests that even after a lengthy legal process, the start of a criminal trial against a predecessor does not discourage incumbents from consolidating power. However, the trials themselves may go on for long periods of time – over a year for Fujimori's longest criminal trial – and successors might not perceive that the accused will ultimately be convicted until closer to the end. For convictions themselves, the coefficients are almost all negative, although the p-values are clustered at the top of the range, indicating no significant relationship. Upon inspection, there were only six convictions of former presidents in

the dataset during years in which I also have data on successors' decrees, four of which were issued against Fujimori, which may be too little data to make useful conclusions.

Figure 25. Summary of regression models for sanction events at stages of legal process (Recording events only in the month in which they began)

Dependent variable: decrees enabling executive, 1000 models for each independent variable



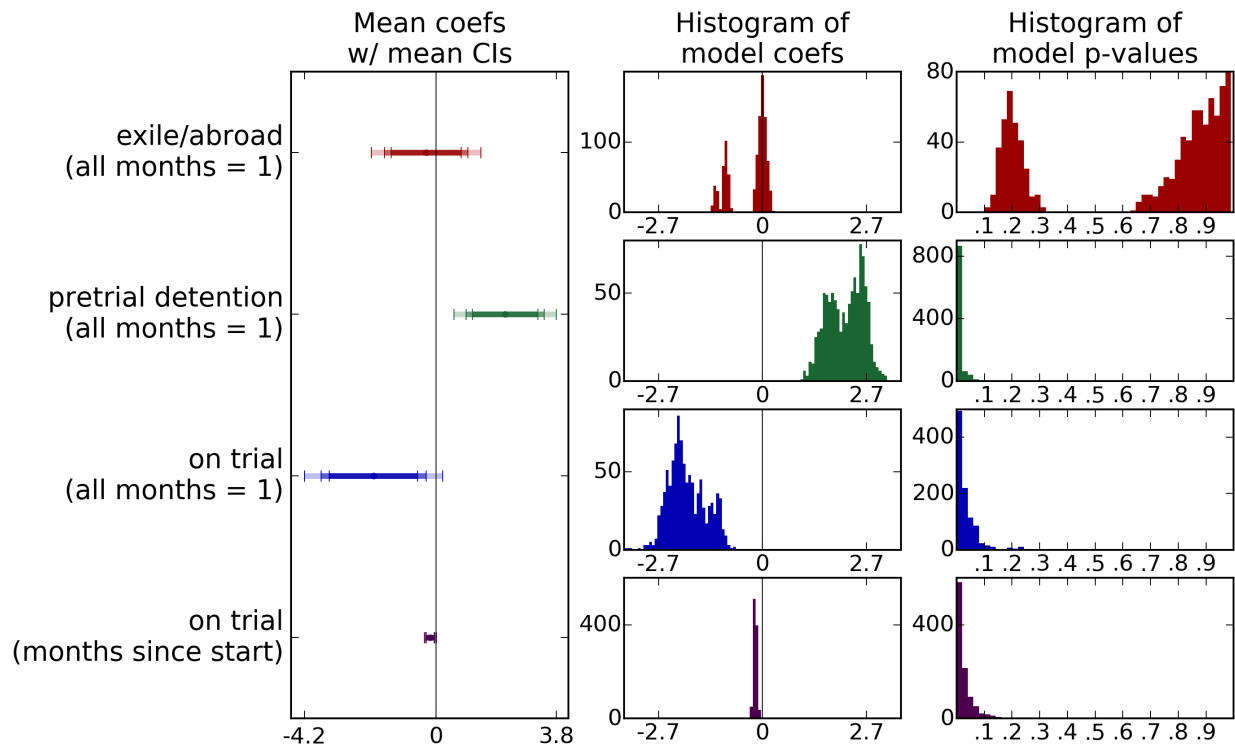
However, we might be able to get useful variation and statistical power out of the months during which a former leader was on trial, not only at the start of the trial but for its duration, since incumbents might become increasingly convinced that they will be held accountable as they observe a predecessor's trial play out in full and lead toward conviction. Figure 26 below shows the same variables recorded in every month in which they were in effect (except convictions, which only occur at one point in time). The relationships for former leaders in exile have become indistinguishable from zero, but the others have become more significant.

The coefficients on pretrial detention are still positive and grouped well away from zero, and all p-values fall below 0.05. For a month in which a former president was sitting in pretrial detention, successors issued an average of just over two additional power-consolidating decrees over the subsequent one to twelve months, than they did in a comparable month in which a former president was not in jail. The magnitude of the estimates is smaller than in Figure 25, because leaders' reactions are spread out over many months for each detention, as we saw with

the smaller coefficients in the analysis of multi-month rewards. However, the average duration of a former leader’s pre-trial detention in the dataset is about six months, meaning that each pre-trial detention in aggregate is associated with about 13 more power-consolidating decrees issued by a successor throughout the period of detention and up to one year afterwards.

**Figure 26 Summary of regression models for sanction events at stages of legal process
(Recording events in all months in which the sanction was in effect)**

Dependent variable: decrees enabling executive, 1000 models for each independent variable



The coefficients on former leaders’ trials (third row) have now become strongly negative, and the p-values are stacked tightly toward zero, with most falling below 0.05, indicating robust significance across model specifications. On average, in a month in which a former president was formally on trial for abuse of power, successors issued about two *fewer* power-consolidating decrees than in a comparable month in which no predecessor was on trial. Since this relationship was not present at the start of the trial, and the coefficients were even weakly positive, the relationship appears to have become more negative over the course of the trial. That is, as conviction becomes more imminent, successors respond by exercising restraint in their own actions. To test for this relationship over the course of the trials, I include a final row in which the explanatory variable is the number of months since the start of a former leader’s trial. The estimates are again all negative and highly significant. The magnitude is smaller because the

variable is an incremental count of months on trial. Incumbents issue about one fewer power-consolidating decree for every five additional months that a predecessor has been on trial.

Qualitative inspection of the data supports the interpretation that leaders become more deterred as trials progress, since all formal trials in the dataset led to convictions. This is not necessarily always the case; one former president in Venezuela did face the start of a trial before it was dismissed by the court, but that trial occurred before the period in which I was able to collect successors' decrees. In general, the fact that all trials in the dataset led to convictions and that the relationship to power-consolidating decrees becomes increasingly negative the longer a trial has gone on, suggest that incumbents began to curb their own power consolidation efforts later in these trials, as it became more apparent that each was headed toward conviction. This is consistent with the hypothesis that convictions at the conclusion of a lengthy legal process entail a high burden of proof and consensus as to the former leader's guilt, which may deter successors from amassing excessive power themselves, in order to avoid a similar fate.

While this last set of tests made use of more country-months in which a detention or trial was ongoing, the results still draw from only a few former presidents' criminal trials and convictions, so they are at best suggestive of potential broader trends. Ideally, we would like to draw conclusions from a larger number of post-tenure sanctions taken against more leaders in the dataset, including sanctions throughout the legal process, if few former leaders' cases ever reached a criminal conviction during the period of study. However, it might be difficult to tell whether a former leader's case is headed to trial, when looking at individual post-tenure events along the way. If we expect leaders to react to weaker or more politicized cases against their predecessors before getting to a clear verdict of guilt or innocence, we need a better measure of the veracity or objectivity of sanction events at any point in the legal process.

6.3 Measuring the probability or predictability of post-tenure fates

As discussed in Chapter 2, the aspect of former leaders' fates that I expect to have the greatest influence on subsequent leaders' actions is the extent to which the former leaders' punishments or rewards appear objective or predictable based on what they actually did in office. This is a fairly abstract idea that observers might judge in different ways. So far, I have identified some post-tenure events that seem on their own to appear more or less objective. For instance, reprieves represent at least some back-and-forth between attempted sanctions and reversals, suggesting uncertainty over whether to punish the accused. Early arrests and pretrial detention may also be ordered by partisan actors on weak legal grounds, while trials and convictions usually entail a higher burden of proof and consensus as to the former leader's guilt.

In this section, I introduce an additional measure of the nature of former leaders' fates, to more directly estimate the underlying concept of interest: the objectivity or predictability of former leaders' punishments and rewards. Consider several presidents of the same country, such as Peru's Alan García, Alberto Fujimori, and Valentín Paniagua (who served as interim president

after Fujimori's removal). García and Fujimori faced at least some attempted sanctions after departing office (although Fujimori was more consistently prosecuted and eventually convicted). García also achieved later rewards (returning to the presidency a second time), while Paniagua received honorary academic positions and related awards. On the surface, García's fate seems more mixed and uncertain than the other two, since he was both partially sanctioned (but never convicted) and also rewarded. But how does García's fate relate to the actions he took in office? Did he take steps like Fujimori to consolidate power, unilaterally impose his own personal control over the state, and abuse those powers financially and/or with oppressive force? Or did García's policy decisions look more like Paniagua's, in which case his sanctions are more surprising (and may appear more politically motivated) than his rewards?

What I seek to measure is how closely former leaders' actual punishments or rewards match expectations about what *should* have happened to them, given the decrees they enacted while in office. This can be seen as a two-step process. First, we need to assess what should have happened to these former leaders, or at least what reasonable observers might have expected to happen, based on what each leader actually did in office, if post-tenure fates are objective and consistent across leaders. We can treat this as a prediction problem: we see what a former leader did in office and try to predict whether they would be sanctioned, reprieved, or rewarded, given how similar actions by other leaders were punished or rewarded at other times. Since I have data on the decrees that many leaders enacted while in office, as well as data on those same leaders' post-tenure fates later on, we can train a supervised machine learning classifier to predict which fate best matches each leader's record of decrees.

The second step is then to see how accurately the classifier predicts each leader's real post-tenure fate. If the predictions are correct for some leaders, we might say that those leaders' fates were unsurprising or consistent with expectations, which might signal to successors that their own fates will be objective and predictable. If the predictions were *incorrect* for other leaders, we might say that those leaders' fates were more surprising or inconsistent with expectations, which might signal to successors that their own fates may be unpredictable or arbitrarily determined, perhaps driven more by political motivations than by innocence or guilt. In other words, if we calculate leaders' predicted fates and then distinguish real post-tenure events that match the predictions from post-tenure events that do not, we can test whether incumbents react differently to unsurprising sanctions against their predecessors, versus how they react to surprising (or seemingly undeserved) sanctions.

For the first step, I use machine learning to calculate the predicted probability of each former leader being sanctioned, reprieved, or rewarded after they departed office, based on the decrees that same leader issued while in office (sometime earlier). I use supervised classification, as introduced in Chapter 3 for document classification. Here, the goal is to predict the correct fate (i.e. sanction, reprieve, or reward) for each former president in the dataset. I include all leaders for whom I have data on both the decrees they issued while in office, and the sanctions and/or rewards they faced afterwards.

The data used to train the fate prediction model is as follows. I construct one observation for each post-tenure event in the dataset (e.g. one row for Fujimori's arrest, another for his extradition, and one for each of his convictions). For input features, I use the number of decrees issued by the same leader during his/her tenure, in each of the project's action and target categories. In other words, the columns in this matrix are counts of decrees enabling executive offices, decrees enabling private organizations, decrees regulating public offices, etc., summed over all months in which Fujimori was president. I also include a feature for the leader's country. For the output labels, I assign the general category of *sanction*, *reprieve*, or *reward* for the given post-tenure event. For simplicity, I include no element of time. The same decree counts are used to predict all post-tenure events that occurred to the same leader. This means that if a leader faced multiple sanctions and rewards, the data will include multiple identical rows of decree counts, but some will be assigned the label for sanction and others the label for reward.

I then use a random forest classifier to calculate the probability of each post-tenure event. To ensure that the resulting scores represent *predicted* probabilities based on each leader's record of decrees, and not each leader's actual (observed) fate, we need to calculate the predicted fates using a "held out" test set. I use the same method of k-fold cross validation that was used to calculate the accuracy of the document classification models in Chapter 3. I iteratively hold out 1/10th of the post-tenure event observations, train the random forest classifier on the remaining 9/10^{ths}, then save the predicted probabilities of sanction, reprieve, and reward from the held-out 10th. For leaders with multiple post-tenure events in the dataset, I average the predicted probabilities of sanction, reprieve, and reward across all of that leader's rows. (These predicted probabilities should be the same for a given leader, in expectation, since all observations for the same leader's post-tenure events use the same input features.)

To complete the example from Peru, since Fujimori faced multiple post-tenure sanctions, there will be multiple identical rows in the training data, containing the numbers of decrees in each category that he issued throughout his tenure (i.e. before he left office to face the sanctions). Each of these identical rows will be assigned the output label for *sanction*. I then shuffle these rows with the remaining data reflecting other leaders' decree counts and post-tenure fates. Whenever one of Fujimori's rows is assigned to the held-out test set, I train the classifier on the other rows in the dataset, then record the classifier's predicted label probabilities for Fujimori's row. The probabilities of sanction, reprieve, and reward should sum to one.

Table 21 shows the average predicted probabilities for certain types of sanctions and instigating actors. When leaders in the dataset were actually sanctioned by the legislature, they were only predicted as 46.4% likely to be sanctioned, on average, based on the decrees they enacted while in office. In contrast, when leaders were actually sanctioned by the government, those events were predicted as 73.6% likely on average, and leaders sanctioned by the courts were predicted as 76.6% likely to be sanctioned on average. In other words, sanctions by the legislature look more arbitrary and less appropriate to former leaders' actual decisions in office, while sanctions by the government were much more predictable, and sanctions by the courts even more so. Former leaders who reached a formal trial and were convicted of crimes were

predicted as over 90% likely to face sanctions at some point after departing office, indicating that the trials and convictions in the dataset targeted individuals who appear to have strongly deserved to be punished for their acts in office, at least based on the standards set by other former leaders' fates. These predicted probabilities do not amount to moral judgments of innocence or guilt, they simply reflect how consistent and predictable one former leader's fate was based on who else was punished or rewarded for similar actions in office.

Table 21. Predicted probabilities of actual sanctions in dataset, by actor or stage

<i>Type of sanction event</i>	<i>Mean predicted probability of observed events</i>
Former leader sanctioned by legislature	46.4%
Former leader sanctioned by government	73.6%
Former leader sanctioned by courts	76.6%
Former leader on trial	91.2%
Former leader convicted	92.5%

The next step is to incorporate these predicted fates – and how well they match reality – into the main analysis of subsequent leaders' actions. What I am interested in is whether incumbents react to sanctions against their predecessors differently, depending on the predicted probability of those sanctions. In other words, do incumbents react one way to highly anticipated sanctions that seem appropriate to the accused predecessor's record, but react another way to surprising sanctions that do not seem to be consistent with the former leader's actions? We might think of these as heterogeneous effects of sanctions at different levels of predicted probability. One way to test for heterogeneous effects would be to interact the variable for monthly counts of observed sanction events, and the predicted probability of each of those sanctions (i.e. the predicted probability that the leader being sanctioned in each case should have been).

Another way to approach this question is to simply split the post-tenure sanction events in the dataset into those that had a higher probability of occurring, and those that had a lower probability of occurring. To produce an even comparison between these groups, I split the post-tenure sanctions in the dataset into q quantiles, with roughly even numbers of sanctions (n/q) in each bin, where n = the total number of sanction events and q = the number of quantile bins chosen. For a given number of quantiles q , I take the original variable for all sanctions in the dataset, and split it into one variable that just includes the n/q sanctions with the lowest predicted probabilities, another variable with the n/q sanctions that had the next highest probabilities, etc. up to a final variable with the n/q sanctions that had the highest predicted probability.

For example, consider several of the former presidents mentioned in the motivating cases in Chapter 1. The leader with the lowest predicted probability of sanction was Ecuador's Lucio Gutiérrez, at 23.6%. Gutiérrez was removed from office early in April 2005 and fled into exile, soon after which an order was issued for his arrest, but the charges were dropped a year later.

The leader with the highest predicted probability of sanction in the dataset is Peru’s Fujimori, at 95.0%, which may be unsurprising since this is the leader who was convicted four times for abusing power and is currently serving multiple sentences in prison. Somewhat in between is Ecuador’s Jamil Mahuad, at 80.9% predicted probability of sanction. Mahuad’s case was one of the most contentious during Rafael Correa’s rise to power; his charges were dropped in 2006 but his case was reopened in 2007, after considerable political lobbying, including by Correa. Mahuad was eventually convicted of corruption in 2014.

The mean probability of sanction across all leaders in the dataset is 29.5%, along with 10.4% predicted probability of reprieve and 60.1% probability of reward. Most leaders who were actually sanctioned in the dataset have a higher predicted probability of sanction (by design, since the classifier learned from observed patterns). Meanwhile, some leaders with low predicted probability of being sanctioned never were. If we only look at leaders who were actually sanctioned, then, the mean predicted probability of sanction across observed sanction events in the dataset (rather than across all leaders) is 73.1%. These numbers suggest that the few sanctions against Gutiérrez were surprising, since his expected probability of being sanctioned was very low, while Mahuad’s sanctions were somewhat expected, and Fujimori’s sanctions were highly predicted (and would have been surprising if they had not occurred).

Table 22 below shows what the new variables look like. The first three rows show a sample country-month observation in which one of the three leaders mentioned above did face a sanction, while the fourth row shows a country-month in which no former leader was sanctioned.

Table 22. Example values for new sanction variables incorporating predicted probability

<i>Leader</i>	<i>Prob of sanction</i>	<i>Event</i>	<i>Country month</i>	<i>Sanction count (orig)</i>	<i>Sanction x (prob – mean)</i>	<i>Sanctions in 1/3 Q by prob</i>	<i>Sanctions in 2/3 Q by prob</i>	<i>Sanctions in 3/3 Q by prob</i>
Lucio Gutiérrez	0.236	arrest	Ecuador 06/2005	1	-0.496	1	0	0
Jamil Mahuad	0.809	case reopen	Ecuador 08/2007	1	0.078	0	1	0
Alberto Fujimori	0.950	convict	Peru 12/2007	1	0.218	0	0	1
NA	NA	no event	Peru 12/2016	0	0	0	0	0

mean predicted probability across observed sanction events = 0.731

The first column (shaded lightest gray) shows the original variable for all sanction events, recorded by country-month in the dataset. In the second column (shaded medium gray), I interact

the original monthly sanctions variable with the de-measured predicted probability of each sanction. In the examples shown, this is just the predicted probability of sanction for the leader sanctioned that month, minus the mean predicted probability across observed sanctions (73.1%), multiplied by 1 for that month's event count. Unexpected sanctions have negative values, expected sanctions have positive values.

In the last three columns (shaded dark gray), I have split the sanction events into terciles by predicted probability. The green columns contain event counts as in the original variable (i.e. ones, not predicted probabilities). The ones from the original sanction variable in the first column have simply been separated into terciles by predicted probability in the last three columns, so that Gutiérrez's arrest order falls into the lowest probability tercile, the reopening of Mahuad's case falls into the middle tercile, and Fujimori's conviction falls into the highest probability tercile.

These two sets of variables serve as alternative ways of combining the observed sanction events with their predicted probabilities, in order to test for heterogeneous relationships to power-consolidating decrees. For the interaction in the second column, there is only one new explanatory variable, and I will be testing for whether there is a linear relationship between the probability of each sanction event and successors' subsequent use of decrees. For the events divided into quantiles in the last three columns, I will run separate regressions of decrees on the sanction events in each quantile, testing for distinct relationships between lower probability sanctions and decrees, and between higher probability sanctions and decrees, which may not follow a linear pattern across all predicted probabilities.

6.4 Results for post-tenure sanctions by predicted probability

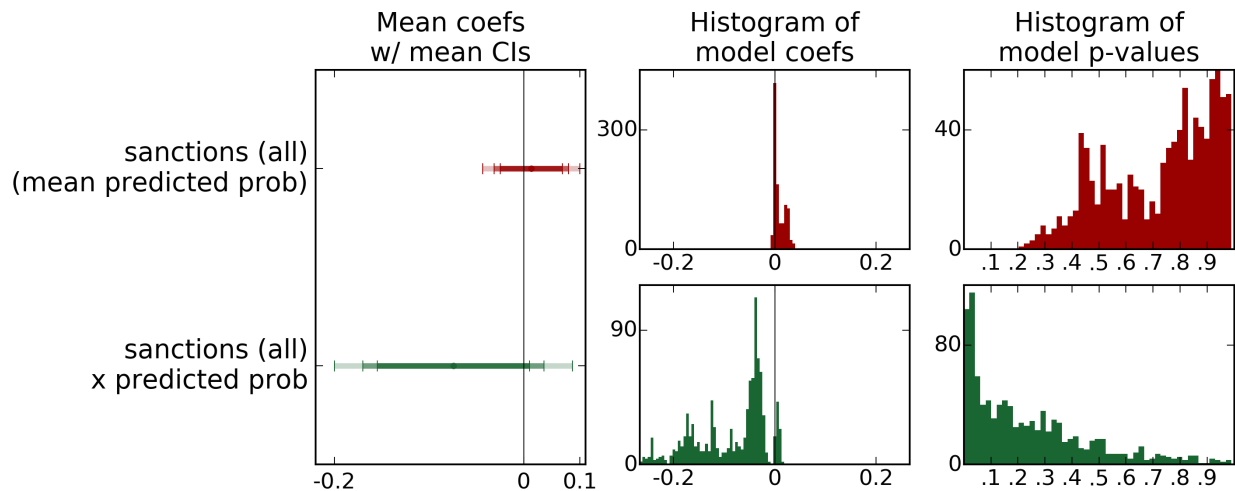
Figure 27 summarizes the results for regressing power-consolidating decrees on the interaction between post-tenure sanction events and those events' predicted probability. Since I have characterized this test as an interaction, I include both the original count of sanction events by country-month (the first column in Table 22) and the interaction term containing the predicted probabilities of each event (the second column in Table 22). I construct the regression models in the same way as all previous hypothesis tests, randomly sampling control variables, fixed effects, clustering factor, and lag period, 1000 times. In Figure 27, I show the distribution of coefficients and p-values for the original sanction variable on its own (the top row) and for the interaction term that contains the predicted probabilities of the observed sanction events in each country-month (the second row), both of which were included in the same 1000 regression models.

The top row essentially represents the relationship between power-consolidating decrees and sanctions that were observed, but are represented by zeros in the interaction term, meaning they had a predicted probability right at the mean. We observe no clear or significant relationship in the top row, suggesting that sanction events with average probability do not sway successors one way or the other. The second row represents the relationship between power-consolidating decrees and the predicted probability of observed sanctions on either side of the mean. There

does appear to be a fairly strong and somewhat robust relationship there. The coefficients are almost entirely negative, and the p-values are clustered toward the lower range of the distribution, although not tightly enough to call this a highly significant relationship. About a quarter of the regression models do have p-values below 0.05, and over a third are below 0.1. On average across all models, the relationship is almost significant at the 90% level.

Figure 27. Summary of estimates for sanctions interacted with predicted probability

Dependent variable: decrees enabling executive, 1000 models for each independent variable



The interaction with predicted probabilities makes the magnitude of the relationship a bit difficult to interpret, so again it helps to consider examples of different leaders in the dataset. Fujimori’s predicted probability of sanction was just 14.1% higher than Mahuad’s, at 95% and 80.9%, respectively. With an average coefficient of -0.01217 on the interaction term, a sanction against Fujimori in one country-month is associated with about $14.1 \times -0.01217 = -0.17$ fewer power-consolidating decree, issued by a successor government, than would have been issued if Mahuad had been sanctioned instead of Fujimori. Meanwhile, Mahuad’s predicted probability of sanction was much higher than Gutiérrez’s, at 80.9% and 23.6%, respectively, or 57.3% difference. A sanction against Mahuad in one country-month is associated with about $57.3 \times -0.01217 = -0.70$, or over two-thirds of a power-consolidating decree fewer issued by a successor government (over the next one to twelve months), than would have been issued if a much more surprising sanction against Gutiérrez had occurred that month instead.

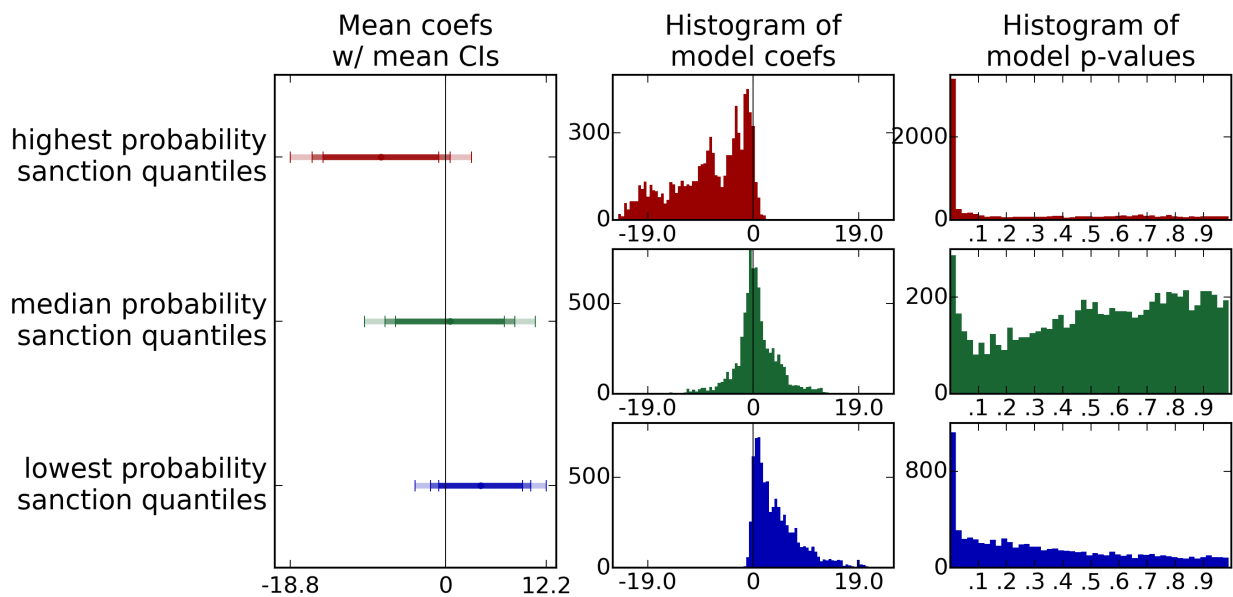
Figure 28 shows similar results, this time separating sanctions into quantiles by predicted probability, while leaving the variables as raw event counts (which may make the interpretation clearer). Again, I select a number of quantiles q , to divide the sanctions into q bins by predicted probability. I have tried to avoid selecting an arbitrary partition of the data, which might occur when using only one number of quantiles. Since I have 49 sanction events in the dataset, I

include quantiles from $q = 3$ to $q = 10$, which means splitting the data into about 5 to 15 sanction events in each bin. All regressions still make use of the full 1366 country-months in the dataset, but the explanatory variables have become sparser, with many more zeros, since I am only testing for leaders' reactions to specific sanctions within certain ranges of predicted probability.

We then need some way to aggregate and make sense of the results, across iterations in which sanction events were split into a different number of bins. In the figure below, I do not show results for all quantiles. Instead, I summarize estimates only for quantiles at certain points along the spectrum of predicted probabilities. In particular, I expect the strongest relationships to appear at the two extremes, i.e. sanctions of leaders who strongly deserved to be sanctioned, and sanctions of leaders that were highly unexpected because they did *not* deserve to be.

Figure 28. Summary of estimates for sanctions by predicted probability quantile

Dependent variable: decrees enabling executive, 1000 models for each independent variable
Quantiles: $q \in \{3, \dots, 10\}$; lowest = quantile 1, median = quantile $q / 2$, highest = quantile q



To show aggregate results at these points on the spectrum, I combine the results for the top quantile in every partition of the data (across values of q), into one row of Figure 28. I also combine the results for all median quantiles into a second row, and results for all bottom quantiles into a third row. In other words, I regress power-consolidating decrees on the top tercile of sanction events (i.e. the 15 highest probability sanctions) as the explanatory variable, then the top quartile of sanctions, on up to the top decile (i.e. the 5 highest probability sanctions). Those results appear in the top row of Figure 28. The top quantile models do not include any of the lower probability events; we are simply comparing how leaders react to a highly expected sanction of a predecessor, versus how they act when no sanction occurred at all. The distribution

of estimates in the top row places greater weight on the very highest probability sanctions, since they are always included in the top quantile (even when q is larger and the bins are smaller). The second row places the greatest weight on the median-probability sanctions, and the third row places the greatest weight on the very lowest-probability sanctions, again comparing country-months with one of those sanctions to country-months with no sanctions at all.

The results in Figure 28 are consistent with those in Figure 27 and provide visually compelling evidence of the main hypothesis. Consistently across different sizes of quantiles, the highest probability sanctions in the top quantile are associated with fewer power-consolidating decrees, while the lowest probability sanctions in the bottom quantile are associated with more power-consolidating decrees. The coefficients for the top quantiles are almost all negative and the p-values are stacked tightly in the lowest bin. There is a thin tail over the remainder of the spectrum, which makes the average p-value fall just outside the 95% confidence interval, but the full distribution of results suggests that there is a strong negative relationship between power-consolidating decrees and sanctions against former leaders who were highly expected to be sanctioned, which is robust across most model specifications. One sanction of a former leader, who is among the top most predicted to be sanctioned based on his/her previously issued decrees, is associated with about eight fewer power-consolidating decrees, on average, issued by a successor in the next one to twelve months.

For sanctions in the median quantile of predicted probabilities, there is no clear relationship that is robust across quantile bin sizes or model parameters. The coefficients have both negative and positive values, and the p-values are more spread out across the full range. These results are consistent with the upper row of Figure 27, which showed estimates right at zero for the relationship between power-consolidating decrees and sanctions at the mean predicted probability among observed sanctions of former leaders.

For sanctions in the lowest quantile of predicted probabilities, the coefficients are now almost all positive, and the p-values are again stacked in the lowest bin. There is a somewhat fatter tail over the remaining range than there was for the top quantile of highly predicted sanctions, which makes the average p-value not quite significant over all model specifications for the lowest probability quantile. There may be more noise in the lower predicted probabilities of sanction than in the higher predicted probabilities, since the higher probabilities correspond to the decree records of leaders who actually faced more sanctions in the dataset.

That said, the distribution of estimates at the bottom of Figure 28 suggests a consistently positive, fairly large and sometimes significant, if not very robust, relationship between unexpected sanctions of former leaders and successors' power-consolidating decrees. One sanction of a former leader, who is among the leaders least expected to be sanctioned based on his/her previous decrees, is associated with about four additional power-consolidating decrees on average, issued by a successor in the next one to twelve months. This relationship is significant at the 95% level in close to a quarter of all models run, although limited data and noise in the predicted probabilities may be weakening the relationship across other models. The stark

contrast between the estimates in the highest, middle, and lowest quantiles of sanctions by predicted probability, provides support overall for the main hypothesis in this project.

6.5 Robustness checks using alternative predictions of sanction probability

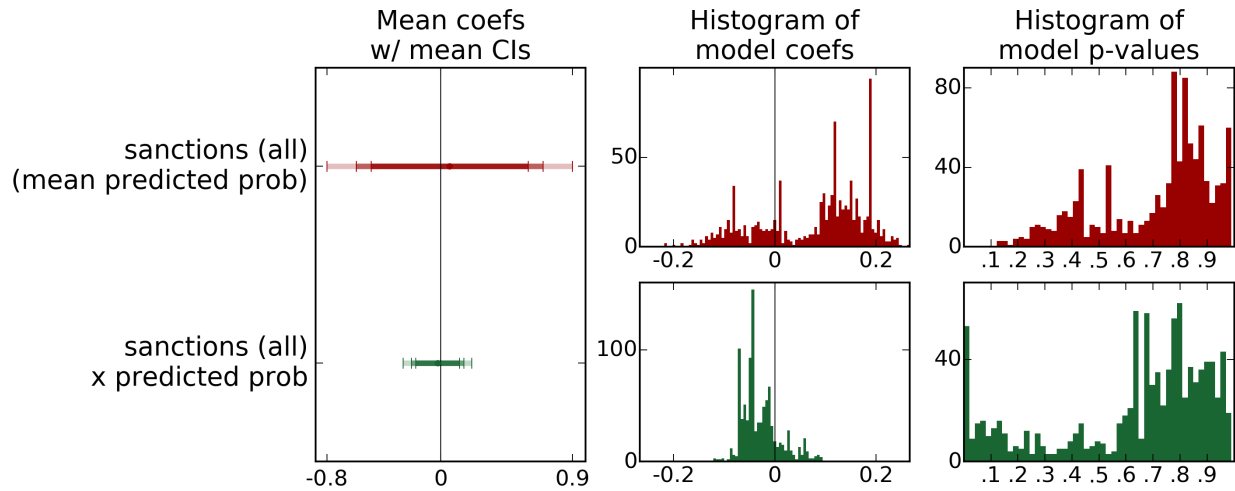
The process used to predict the probability of a sanction, reprieve, or reward does take into account other post-tenure events that occurred to the same leader, if those other events are randomly shuffled into the training set (as opposed to the test set) during k-fold cross validation. The benefit of calculating the predicted probabilities in this way is that they capture *both* consistency of fates across leaders who issued similar decrees, and the consistency of the same leader's own fate. This offers an alternative measure of the contention implied by reprieves alone, but taking into account potential disagreement between sanctions and rewards of the same leader as well. We might be able to distinguish these two types of consistency a bit more, however, by calculating each leader's predicted probability of sanction or reward based solely on the fates of *other* former leaders who issued similar decrees.

However, if we are most concerned about a realistic estimate of the incumbent's own expectations, we might also consider a third alternative: calculating predicted probabilities of each sanction event chronologically, based on how former leaders who issued similar decrees were treated *before* each event in question. The first time Fujimori faced sanction, the incumbent (his immediate successor Valentín Paniagua) might have only been able to compare that event to previous sanctions against other former leaders. As Fujimori's case progressed, however, Paniagua might have begun to assess the consistency of Fujimori's own post-tenure events. Each time Fujimori faced new charges or a new trial, the expectation that he deserved or would likely be punished might have grown, making each additional sanction against him less surprising.

As robustness checks, I recalculate the predicted probabilities of sanction in these two additional ways. First, instead of shuffling the data and using k-fold cross validation, I simply take former leader in turn, reserve that leader's own post-tenure fate events, train the classifier on the post-tenure fates of all other former leaders, then calculate the reserved leader's predicted probability of sanction based solely on his/her previously issued decrees. Second, I construct the training set cumulatively over time, training the classifier on all prior post-tenure events that occurred to any former leader before a given month, then calculating the predicted probability of each sanction event that occurred in the new month, based on the targeted leader's prior decrees. Figures 29 through 32 show results from these alternative predicted probabilities.

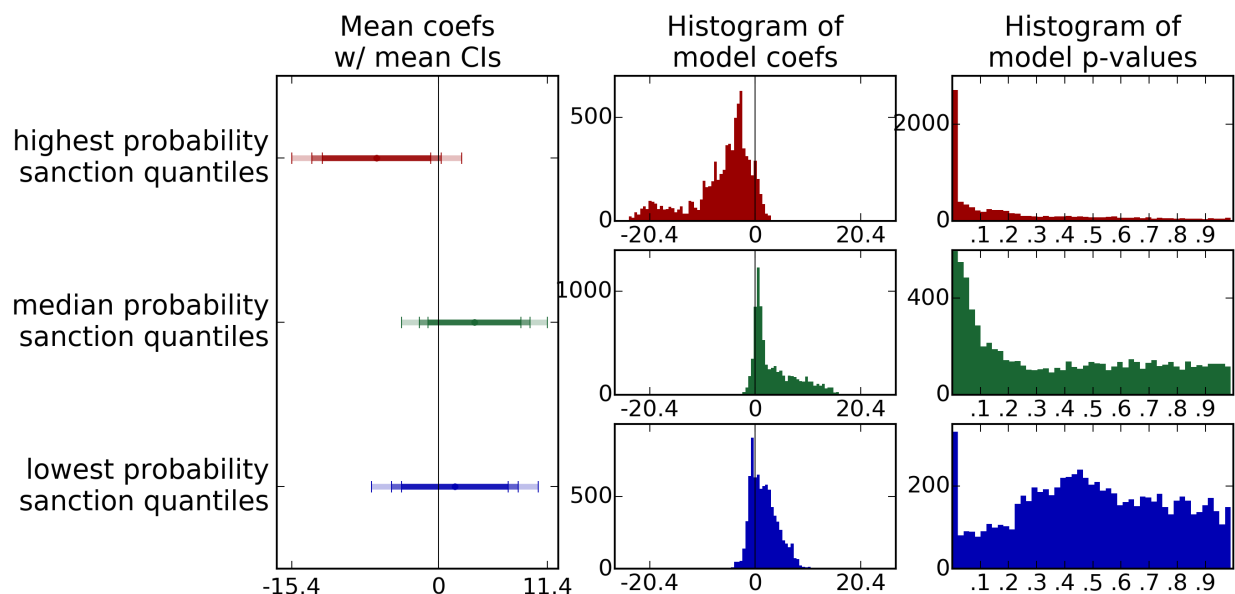
Figure 29 shows estimates for the interaction between monthly sanction counts and their probability when predicted solely based on the fates of other leaders who issued similar decrees (but not the same leader's own prior sanctions or rewards). Most coefficients on the interaction term are negative, but the relationship is insignificant, when fitting a linear regression to observed sanctions across the full spectrum of predicted probabilities.

Figure 29. Sanctions interacted with predicted probability, predictions stratified by leader
Dependent variable: decrees enabling executive, 1000 models for each independent variable



However, when splitting the same probabilities into quantiles in Figure 30 below, there is a modestly significant relationship for the highest probability quantiles. The coefficients are negative in the top row, as expected, and generally positive in the second and third rows. While there is some concentration of p-values in the lowest bins, the estimates for median and lowest probability sanctions are not robustly significant across all models.

Figure 30. Sanctions by predicted probability quantile, predictions stratified by leader
Dependent variable: decrees enabling executive, 1000 models for each independent variable
Quantiles: $q \in \{3, \dots, 10\}$; lowest = quantile 1, median = quantile $q / 2$, highest = quantile q



We might infer that without using inconsistencies within the same leader's fate (i.e. when leaders were both punished and rewarded, which shows a high degree of contention about what they deserved), it has become more difficult to distinguish the most surprising sanctions from sanctions that simply had an average or ambiguous likelihood. Yet the negative relationship in the top row suggests that the predicted probabilities *do* still capture meaningful trends in terms of who was *most* expected to be sanctioned based on their use of decrees. There are some decisions that leaders make in office which *are* expected to be punished, in such a way that actually punishing those leaders does deter successors from consolidating power themselves.

Figure 31 shows the interaction results again when using probabilities predicted in chronological order, based on other events that have already occurred. As mentioned above, this is probably the most realistic simulation of incumbents' actual observations. Here there is a large and highly robust negative relationship to successors' power-consolidating decrees, reinforcing the main conclusions. As successive sanctions, reprieves, or rewards against former leaders begin to appear less consistent with events that already occurred to the same leader or to leaders who issued similar decrees, incumbents tend to issue more power-consolidating decrees.

Figure 31. Sanctions interacted with probability, predicted from prior fates chronologically

Dependent variable: decrees enabling executive, 1000 models for each independent variable

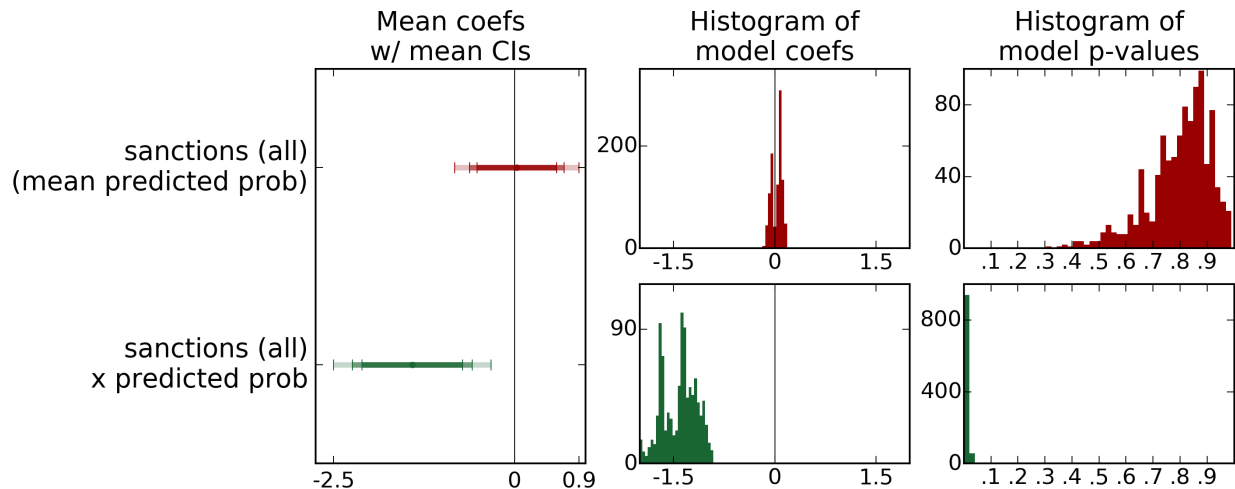
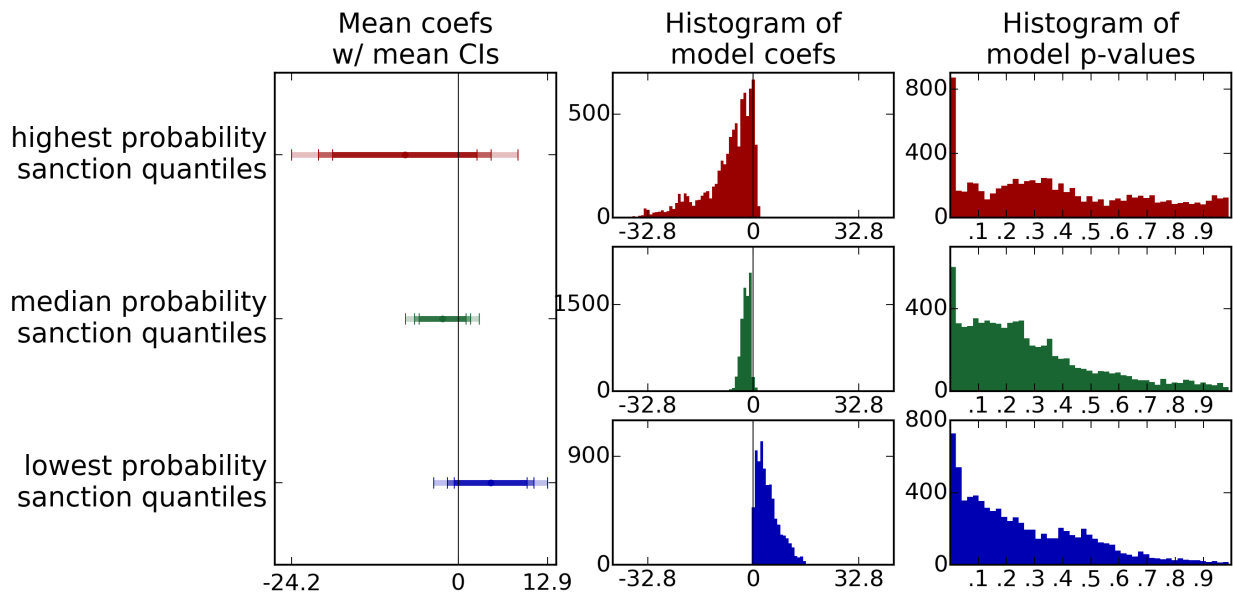


Figure 32 shows the same quantile regression as used above, partitioning sanctions by probabilities predicted chronologically. The estimates here are not very robust across models, but in each row, the largest concentration of p-values is in the lowest bin (over 10% fall below 0.05 in both the top and bottom rows), and the coefficients move from negative values for high-probability sanctions to positive values for low-probability sanctions, as expected. Again, these results support the main hypotheses and the conclusions drawn from the previous analysis.

Figure 32. Sanctions by probability quantile, predicted from prior fates chronologically
Dependent variable: decrees enabling executive, 1000 models for each independent variable
Quantiles: $q \in \{3, \dots, 10\}$; lowest = quantile 1, median = quantile $q / 2$, highest = quantile q



When former leaders are punished for their prior actions in ways that are consistent how other similar leaders have been treated, successors appear to restrain their efforts to consolidate power. In those cases, the prosecution of former leaders does appear to have the intended result of encouraging accountable leadership, deterring future leaders from consolidating too much unchecked power and the potential to abuse it. However, when former leaders are sanctioned in ways that are highly inconsistent with established standards about what types of behavior should be punished (suggesting that the punishment was instead arbitrary or politicized), successors tend to consolidate power more, potentially to protect themselves from similar threats.

The answer to achieving more constrained and accountable authority appears to be neither more punishment nor less, but more deliberate punishment when appropriate – with signals of impartiality, such as avoiding excessive pretrial detention and ensuring that authorities pursue strong cases that lead to criminal trial and conviction – while refraining from seeking punishment when leaders’ actions do not clearly warrant it. For ease of interpretation, I have often referred to these sanctions as *deserved* or not, based on the accused leaders’ previous actions in office. As mentioned in Section 6.3, I have not sought to impose any moral judgment as to whether former leaders are guilty of crimes. The key distinction is whether former leaders’ fates are *predictable*, so that if successors exercise restraint, they are confident that they can avoid similar punishments.

6.6 Relationship between near-term pressures and post-tenure fates

As a final point of analysis, we may have enough data to begin to ask how government leaders' expectations of future punishment or reward interact with near-term pressures to respond to crises today. This last set of analyses addresses the fact that leaders are probably motivated by multiple interests and often have to weigh the trade-offs between competing goals. In Chapter 4, I introduced measures of exogenous price shocks, in an effort to analyze how government leaders use executive decrees when facing urgent crises that might require them to get more things done right away, but that might also increase the risks of punishment later on, if they are blamed for economic failures. The tests in Chapter 4 revealed a negative relationship between export price changes and total decrees, but no clear relationship for a specific decree category.

It might be the case that leaders respond to crises differently, depending on how great a threat they perceive to their own future fate. For instance, some leaders who are not very worried about being blamed for the crisis – or not very worried about the consequences of being blamed – focus much more on simply enacting external programs and services to help industries recover and alleviate hardship to affected populations. However, leaders who are more worried about what might happen to them as a consequence of making tough choices under bad circumstances, might be more inclined to take advantage of the crisis to expand their power in more lasting ways. Such leaders might anticipate that rivals will also take advantage of the crisis to challenge them personally, so that they had better consolidate power to block rivals from doing so.

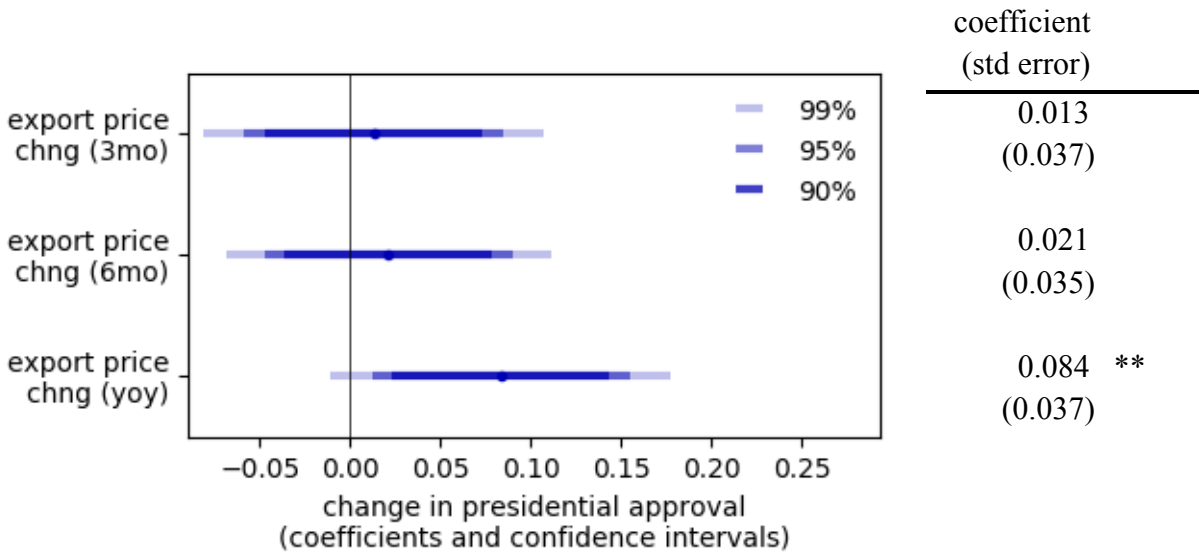
In order to bring the exogenous price shocks into the main analysis in this chapter, a relevant first question to ask is whether global commodity price shocks are really blamed on domestic leaders, who have no ability to significantly affect those prices. It seems likely that leaders will face pressure to *respond* effectively to the crisis either way, but will they also face greater accusations of mismanagement simply because of the economic downturn? Since the price shocks are exogenous, we can test whether they have an observable effect on presidential approval ratings, to see if these price shocks represent an exogenous source of variation in opposition to the president and the consequences that might come with it.

For this check, I regress presidential approval on exogenous price shocks, similar to the bivariate regressions on control variables in Chapter 4. I use presidential approval ratings from the Executive Approval Project (Carlin et al 2016), which provides monthly time-series data compiled from hundreds of polling firms and thousands of individual polls. As in the analysis in Chapter 4, I include country and year fixed effects, and cluster standard errors by presidential term. Since the price shocks are strongly exogenous, and there are countless other explanations for presidential approval, it would be unrealistic to include other control variables here.

Figure 33 shows the bivariate regression estimates. While all coefficients are positive, only the year-on-year price change is statistically significant. An additional one percent increase in the average price of each country's top three primary commodity exports, over the average price one year before, is associated with close to one tenth of a percent increase in the president's approval ratings collected in major public opinion polls. This suggests that even exogenous

economic shocks do factor into the public’s perceptions of presidential performance, although with some lag, since it takes time for remote price changes to filter through domestic revenue into employment and other economic conditions that might be felt by the population.

Figure 33. Bivariate regression of presidential approval on export price shocks



6.7 Results for interaction between sanctions and economic shocks

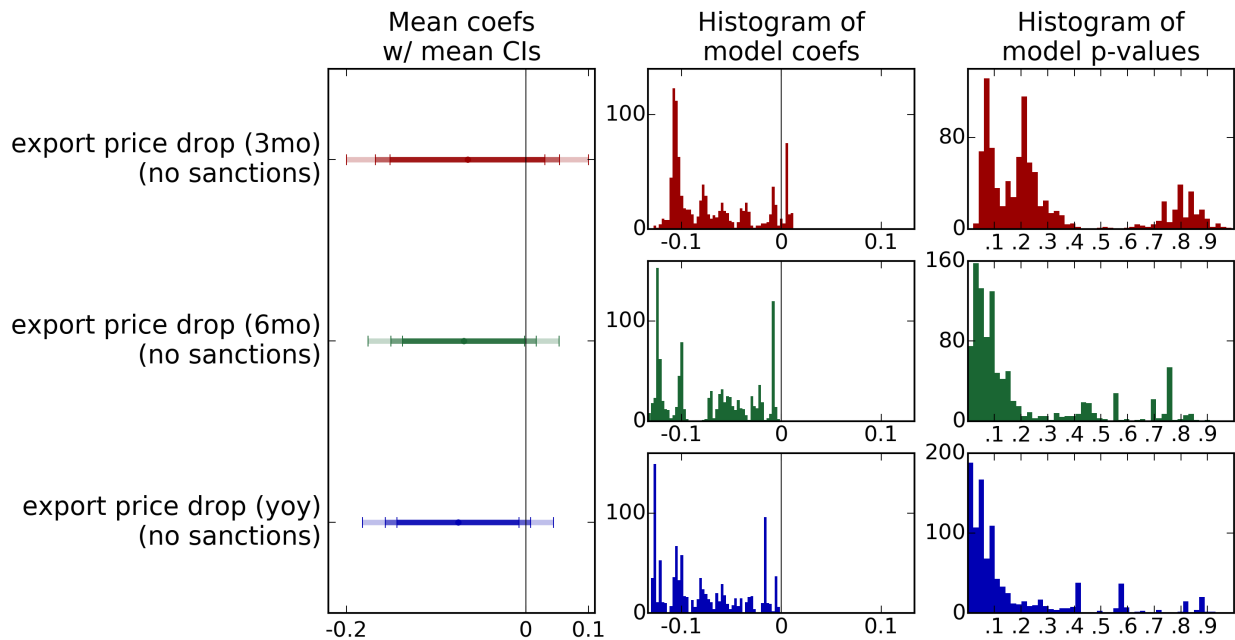
The next step is to test for a relationship between power-consolidating decrees and the interaction between price shocks and former leaders’ fates. In each of the following tests, I incorporate one economic price change variable, one post-tenure fate variable, and the product of the two. I then run 1000 models randomly sampling the other potential control variables, fixed effects, clustered standard errors, and lag periods, using the same options as in all previous hypothesis tests. For the economic shocks, I now use export price *drops*, only including negative export price changes and flipping the sign. This enables us to interpret these tests as capturing leaders responses’ to economic downturns that might pose risks to their future fate.

For post-tenure fates, I use the original variable for all sanctions by country-month. The objective in this section is not to investigate how leaders respond to near-term pressures when they see former leaders being rewarded, nor how leaders behave when the economy is doing well, but how they respond when they fear blame and its potential consequences. Using the variable for all post-tenure sanctions maximizes the available data included in these tests. The goal here is to learn something more about how leaders react to the possibility of sanction, i.e. how they assess threats to themselves in different contexts, in addition to the lessons learned about leaders’ behavior toward more specific aspects of sanctions earlier in this chapter.

Figures 34 and 35 below show the distribution of estimates for the individual price drop and sanction terms, and Figure 36 shows the estimates for the interaction terms, all drawn from the same randomly sampled regression models. In all three figures, I have stacked plots for each duration of export price drops: one row each for three months, six months, and one year. Since one term for sanctions and one term for export price drops were included in each model along with their interaction, the individual estimates show the relationship to power-consolidating decrees when the other variable is zero. In Figure 34, the coefficients on each of the economic price drop variables show the relationship to power-consolidating decrees when there were no sanctions of former leaders during the lag period leading up to the month of observation.

Figure 34. Summary of estimates for export price drop terms in interaction models

Dependent variable: decrees enabling executive, 1000 models for each independent variable



The coefficients on export price drops alone are all negative, and p-values are clustered toward the lower end of the spectrum, achieving moderate significance on average when the export price drop has lasted a year. In Chapter 4, there was no significant relationship between export price changes and power-consolidating decrees without the inclusion of the interaction term. The significant estimates here suggest that when there have been *no* recent sanctions against former leaders, incumbents actually issue fewer power-consolidating decrees during economic downturns. They issue up to one tenth of a decree fewer per one percent additional drop in export prices over the previous year. If incumbents are not worried about being blamed for the crisis and suffering politicized vengeance if they lose power, they may be more willing to

collaborate during emergencies, especially if other parties “rally ‘round the flag” as well, increasing executives’ ability to work through the legislative process instead of going it alone.

In Figure 35, the coefficients on the sanction terms represent the relationship between post-tenure sanctions and power-consolidating decrees in each of the interaction models, when there was no drop in export prices at three months (top), six months (middle), or one year (bottom), respectively. In all three cases, we see no clear or significant relationship, with coefficients spread across positive and negative values, and p-values clustered in the upper range. This suggests that when the economy is doing well – i.e. when incumbent leaders do not face near-term pressures that might make them fear imminent removal – they demonstrate no clear reaction to the sanction of their predecessors. They might not perceive that former leaders’ fates apply to themselves, if they feel secure in office, or they might perceive that they still have plenty of time to weaken rivals and protect themselves later on.

Figure 35. Summary of estimates for sanction terms in interaction models

Dependent variable: decrees enabling executive, 1000 models for each independent variable

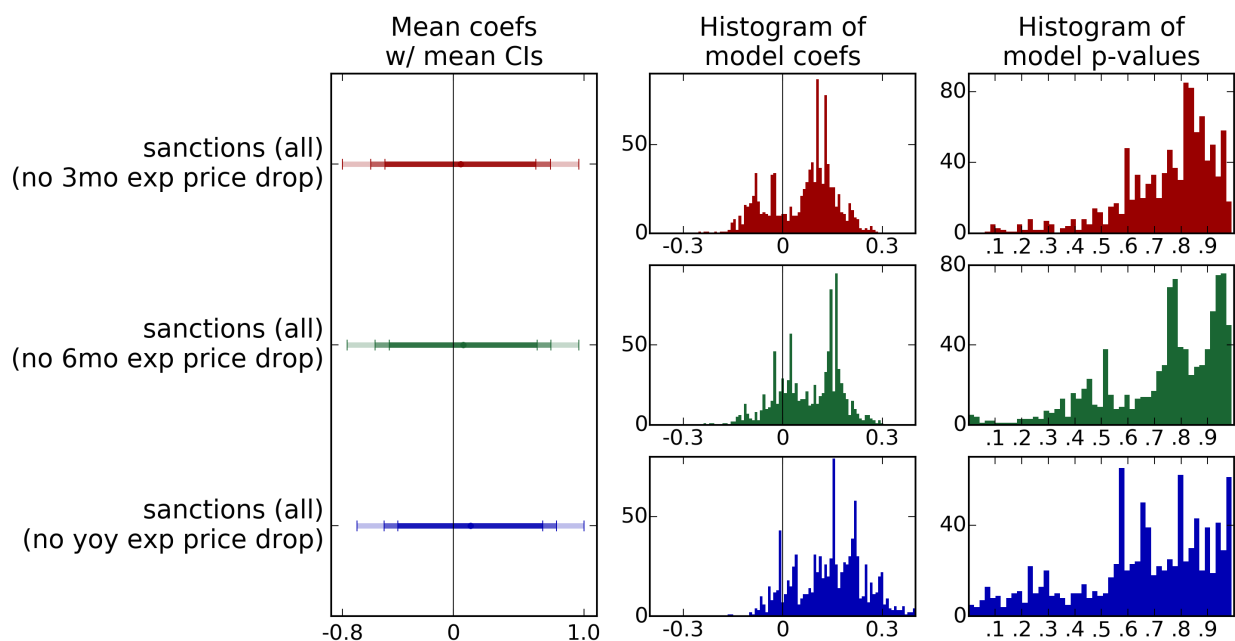
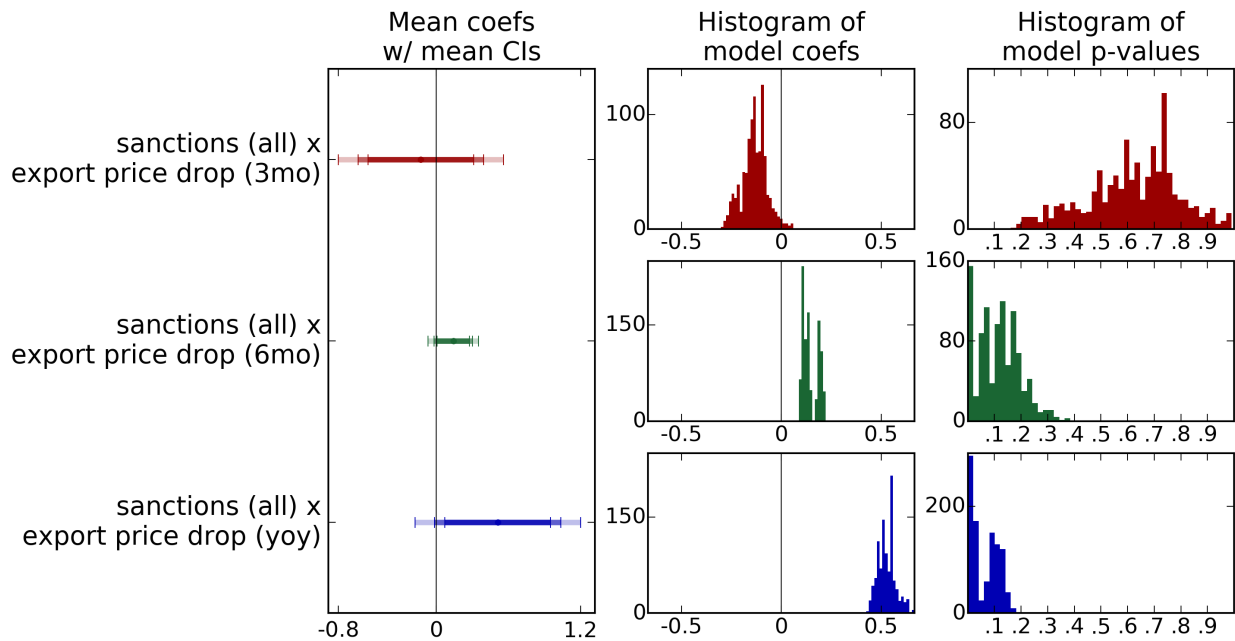


Figure 36 shows the regression estimates for the interaction terms, again summarized across model specifications. In the top row, there is no significant relationship; almost all coefficients are negative, but the p-values are clustered toward the upper range. As the period of economic downturn lengthens, however, the coefficients become more positive and the p-values become more significant. A one percent export price drop over the past six months, accompanied by a sanction against a former leader, is associated with about one-tenth of an additional power-consolidating decree. A one percent export price drop over the past *year*, accompanied by a

sanction against a former leader, is associated with an additional half of a power-consolidating decree. At one year, this relationship also becomes robustly significant across most models, with about half of the p-values stacked below 0.5 and the rest tightly grouped at 0.1.

Figure 36. Summary of estimates for sanction-price shock interaction terms

Dependent variable: decrees enabling executive, 1000 models for each independent variable



In sum, the estimates from the interaction models suggest that government leaders react to economic crises differently, depending on whether their predecessors have recently faced punishment for prior acts in office. Or, in parallel, government leaders react to their predecessors' fates differently, depending on whether they are facing good economic conditions or bad, especially when economic shocks have persisted over longer periods of time. This suggests that leaders' reactions to others' fates depends on the near-term pressures they face in their own jobs. Given how frequently Latin American presidents have been removed before the end of their term, in the face of popular opposition and criticism of their performance (not only during violent conflict, but also due to economic crises or other governance failures), sustained economic downturns might make incumbents fear that their own removal is near, and might compel them to cling to power if they fear that they will face punishment afterwards.

6.8 Summary of additional analyses of post-tenure sanctions

This chapter presents additional analyses of former leaders' sanctions, distinguished by different stages in the legal process, by predicted probability based on the accused leaders' prior decrees, and by the context in which they occur. The results lend weight to the central argument of my theory, that leaders not only care about how many of their predecessors have been punished, but how objective or predictable those punishments were, so that the incumbents can decide how best to avoid a similar fate. The results also provide additional lessons about the timing of sanctions and the competing pressures that government leaders may face, especially during difficult times, when the desire to hold leaders accountable may be strongest.

As discussed in Chapter 2, I originally expected convictions to lead to greater restraint by subsequent leaders, since convictions should indicate greater consensus and objectivity than the filing of charges or an initial arrest warrant. Limited data makes this relationship difficult to analyze. But by including all months in which a former president was formally on trial for criminal acts, I was better able to identify a relationship between the events leading up to a conviction, and reductions in successors' consolidation of power. Verdicts at the conclusion of high-profile trials may be anticipated before they finally arrive. The highly significant and robust negative relationship between former leaders' criminal trials and subsequent leaders' efforts to consolidate power suggests that leaders can be held accountable and even deterred from amassing excessive unilateral authority, when rival authorities follow through and formally punish former leaders for crimes, and when there is enough evidence and consensus as to the accused former leaders' guilt to achieve a conviction.

In Sections 6.3 and 6.4, the analysis of whether leaders' fates appeared consistent with their own previous actions in office, provided an alternative way of measuring the objectivity or predictability of former leaders' fates. By first predicting the probability that each leader would face sanction or reward at some point after departing office, then splitting the observed sanctions into new variables based on whether they were expected (with high predicted probability) or unexpected (with low predicted probability), I sought to imitate the logical steps that incumbents might use to draw lessons from their predecessors' experiences. Leaders who have the ability to make decisions that affect powerful government institutions, probably care about more than just how many of their predecessors faced sanctions, regardless of the circumstances. They may want to know which predecessors' cases are more relevant to their own. And they may want to know how their own actions can affect their future fate, and what course of action is most likely to help them avoid sanctions against themselves.

The analysis of expected and unexpected sanctions supports the central hypotheses in this project, that when former leaders face sanctions that appear arbitrary, unpredictable, and potentially manipulable, successors are more likely to seek to consolidate power. When former leaders are instead sanctioned only when they deserve to be, or at least when their past behavior suggests that they are likely to be, incumbents appear to issue *fewer* power-consolidating decrees. The results from section 6.4 are consistent with those in section 6.2, since the sanctions

by judicial actors toward the end of the legal process – especially trials and convictions – received the highest predicted probabilities. Both sets of analyses support the theory, and together provide a more complete picture that makes use of multiple measures of the underlying phenomena, to overcome data limitations and noise in the estimation of latent concepts.

Finally, the analysis of interactions between near-term pressures and long-term fates helps flush out this project's policy implications. As mentioned in the discussion of potential interaction effects in Chapter 2, understanding how leaders react to crises may be useful for predicting changes to government power, but is less useful for preventing those developments. Government leaders will sometimes need to face emergencies or other governing challenges, and make difficult decisions that may be unpopular. Discovering that leaders tend to issue more decrees during crises would not be very surprising, nor would it help practitioners learn how to constrain unilateral executive action and maintain effective checks on power. However, the evidence from the interaction effects is more useful, showing that some leaders refrain from consolidating power even in the face of economic downturns, when those leaders are less concerned about facing future sanctions even if they do lose their jobs as a result of the crisis.

When former leaders deserve to be punished for abuses of power, the first two sections suggest that those leaders should be sanctioned to the fullest extent of the law. The interaction tests, however, suggest that it might still be prudent to reserve major sanctioning events for stable economic and political conditions, while refraining from issuing criminal charges or holding a high-profile trial during major crises. This last conclusion may appear to prioritize stability over justice. But if justice can still be served while timing new developments in former leaders' legal cases to coincide with points at which sitting presidents feel more secure in office and can afford to operate with longer time horizons, prudent timing might help ensure that sanctions produce the right deterrent effect on subsequent actions.

However, if concerns about economic downturns or similar near-term challenges become an excuse to avoid prosecuting leaders for too long, when those leaders need to be held accountable and are expected to face punishment, the absence of expected sanctions could embolden subsequent leaders by suggesting they too can act with impunity and escape justice. Together, these results suggest ways of assessing the potential consequences and trade-offs involved in prosecuting former presidents under different circumstances, including what crimes the leaders committed and the consensus surrounding their guilt, as well as what other conditions are present that might affect how successors will interpret the risks to their own fate.

Chapter 7. Conclusion

In this dissertation, I have investigated the relationship between the prosecution of former leaders for abuses of power and subsequent leaders' efforts to consolidate their own power through executive decrees. This project was motivated by several debates among scholars and policy practitioners, regarding the causes and consequences of concentrated or unchecked executive authority, and the effectiveness of holding leaders accountable for their actions after they depart office. I have argued that what matters to government leaders is not necessarily how many of their predecessors have been punished or rewarded, but how objective former leaders' punishments have been, in relation to how they actually used their power.

The hypotheses I proposed involved complex and nuanced aspects of leaders' actions and subsequent punishments, requiring new measures of both the dependent and independent variables. I also faced common empirical challenges involved in conducting an observational study of powerful government institutions, with the limited number of cases and events that could be assembled for each country, given resource constraints and the infrequent nature of presidential turnover and prosecution. All of these challenges required innovative uses of data and measurement, which I sought to tackle by applying computational tools and methods to each of the substantive tasks undertaken. In doing so, I have presented several theoretical and methodological contributions to the study of comparative political institutions and public policy, which are in many ways intertwined.

In this concluding chapter, I summarize those contributions and discuss their broader applications and implications. The theoretical contributions may be useful to other researchers formulating substantive theories and hypotheses, but should also provide real-world lessons that are useful for practical application. In section 7.1, I discuss policy implications for developing governments, practitioners involved in the prosecution of former leaders and struggles for transitional justice, and foreign actors concerned about how their influence may affect these processes. The methodological contributions are most useful if they can be directly adopted by other researchers, and I am in the process of developing general-purpose versions of the code used in this dissertation, to make those innovations available to others. In section 7.2, I discuss broader applications of the tools and methods used in this dissertation, as well as more general aspects of the research process and approach that I have taken, which might improve the transparency and reproducibility of other social science research.

7.1 Theoretical contributions and policy implications

I began with the development of a new theory, building on ideas from the literature on presidential power and executive orders, as well as more recent debates from the literature on the prosecution of former presidents. Established theories assume that presidents generally seek

more power in order to accomplish their policy agenda. That is, leaders may use more decrees simply to get more things done, in the face of urgent demands or gridlock in other parts of government. I proposed that leaders might also sometimes seek more power to protect their own interests and weaken rivals who might use state authority to challenge them. In particular, leaders' efforts to increase their institutional power and protect their own interests might be influenced by their expectations of future punishment or reward.

The empirical analysis presented in Chapters 5 and 6 supports my hypotheses that leaders react differently to their predecessors' fates, depending on the types of sanctions imposed and the objectivity or predictability of those punishments. Expected sanctions against former leaders that are consistent with the accused leaders' prior actions, and on which there is enough consensus about the former leaders' guilt to proceed – with few reprieves – to trial and conviction, are most strongly associated with reductions in successors' power-consolidating decrees.

However, unexpected sanctions of former leaders, and those involving more immediate consequences like pretrial detention, as well as frequent reversals through reprieves, are instead strongly and robustly associated with successors *increasing* their use of power-consolidating decrees. In other words, objective and conclusive sanctions appear to produce the most effective accountability and discourage subsequent leaders from consolidating (and potentially abusing) power. Arbitrary sanctions that are not consistent with the accused leaders' prior actions and that may appear more politicized or manipulable, instead encourage the kind of leadership decisions that may perpetuate abuses of power.

These results have a number of policy implications, especially for developing countries seeking to hold government leaders accountable while discouraging power struggles or excessive concentrations of authority. Presidents appear to respond well when predecessors are formally tried and convicted of abusing power, and when those predecessors' actions are consistent with the verdict. Leaders of government branches and opposition groups might be able to send even more direct signals to incumbents that justice will be objective and appropriate, by establishing rules and processes that guarantee political leaders a thorough investigation and fair trial for any claims of wrongdoing, even if the charges are politically contentious.

Another conclusion, drawn from the interaction between sanctions and economic downturns, refers to the timing of legal proceedings undertaken against former leaders. At the end of Chapter 6, I suggested that even when former leaders do deserve to be punished for abuses of power, it might be prudent to avoid initiating new developments in former leaders' cases when incumbents are at their most vulnerable, such as when facing crises that threaten the incumbents' own survival in office. This conclusion may also be relevant to debates about the prosecution of former government leaders in the context of post-conflict transitions.

As mentioned in Chapter 2, some scholars and practitioners have warned that the threat of prosecution may compel embattled leaders to cling to power rather than deter them from additional abuses. This study's results suggest that challengers and mediators might be wise to avoid threats of prosecution while pressing for or negotiating incumbents' departure from office. However, it might be unwise to grant outgoing leaders immunity from future prosecution, which

could send signals to subsequent leaders that they can act with impunity and escape justice later on. But embattled leaders might respond better if cases against their predecessors do not appear to be moving forward too swiftly, while the incumbents are most concerned about their own fate. Signs that predecessors' sanctions are being pursued cautiously and patiently might help incumbents avoid reacting to crises with short-sighted efforts to fortify their positions, especially at times when the incumbents may be facing rising opposition or calls for their own removal.

The analysis of different types of sanctions at different stages in the legal process has related implications for what types of guarantees might encourage government leaders' restraint. Signaling objectivity might include refraining from imposing pretrial detention – which appears to encourage successors' power consolidation – or other major consequences as a result of initial legislative inquiries or executive action, before a former leader's case has reached the courts. That said, there may be valid reasons for arresting accused political leaders who are a flight risk, since many former presidents have obtained asylum and even rewards in foreign countries, while facing criminal charges at home. Ensuring more stringent criteria for pretrial arrest and detention might signal that justice will be fair, while ensuring that former leaders who do deserve to be tried domestically for crimes will be around to face their sentences.

The implications of former leaders' escape into exile are complicated, and in this study, I did not find a clear relationship between former leaders living abroad and subsequent leaders' decisions to consolidate power. However, the results for foreign reprieves indicated that when foreign governments protect former leaders from justice by granting them asylum and refusing to extradite them, incumbents are more likely to consolidate their own power. There was also suggestive evidence that sanctions by foreign governments choosing to comply with extradition requests tend to deter subsequent leaders' power consolidation.

These lessons might be important for foreign actors seeking to encourage responsible leadership in other countries, or at least seeking to avoid playing a detrimental role in developing countries' efforts to establish accountable governance. If foreign governments support and protect individual political leaders they favor, over domestic rules and processes about leadership selection and justice, those protections may encourage subsequent leaders to amass and even abuse power, believing they can escape justice if they secure powerful allies first. If foreign actors reserve their own judgment and defer to internal legal proceedings, when domestic investigations produce sufficient evidence to request a former leader's extradition to stand trial, compliance may strengthen political accountability and reduce subsequent abuses of power.

At the center of my argument is the proposition that leaders are not simply trying to predict their future, they are trying to change it. They want to know how their actions may affect their fate, and determine the best course of action to improve their fate, given the political context they face and how determined their opponents may be to eliminate them. The same might be said of any social actors, trying to assess the likely costs and rewards of the choices they face. The actors need not be political leaders or elites; the central ideas in this project have many parallels to current debates about law enforcement and national security as well. Relevant debates include the extent to which heavy-handed punitive measures actually reduce crime, or

instead provoke a backlash among targeted populations while hindering many law-abiding citizens from pursuing productive opportunities.

Studies of criminal justice and counter-terrorism might benefit from measuring the objectivity or consistency of punishments with the prior behavior of targeted individuals. There may be many systemic factors that make it difficult to determine whether convicted or killed individuals were innocent of crimes (Gross and O'Brien 2008; Koppl and Sacks 2012). But that might not be strictly necessary to determining whether the fates of targeted populations are *predictable*. The predicted probabilities of sanction used in this study might offer a more neutral measure, not of guilt or innocence, but of how consistent individuals' fates are with the fates of others who engaged in similar activities. Punishments and rewards that are at least objective and predictable, on the basis of actions or behaviors that individuals can change (rather than demographic or geographic factors that they cannot), might lead to more effective deterrence and restraint, rather than pushing individuals to resort to other means to secure or protect their fates.

7.2 Methodological contributions and broader applications

This dissertation presents several innovative uses of computational tools and methods to overcome common challenges in data collection, measurement, and statistical analysis. One of the most difficult aspects of social science research is how to turn theoretical concepts into observable units of data for analysis. This step in the research process does not always receive the attention it deserves; some critics have noted that social science studies often jump from theoretical propositions to data measures selected for analysis, to hypothesis tests, without adequately validating that the chosen measures represent the underlying phenomenon of interest (Adcock and Collier 2001). In hypothesis testing, studies may also present only a limited number of alternative model specifications as robustness checks, which represent a tiny fraction of the possible combinations of variables and parameters that could be tested.

I have sought to construct a thorough and deliberate research process from start to finish, comparing and validating options at each step in the transformation of ideas into data and back into lessons drawn from that data. I constructed new measures of the main concepts using relatively consistent digital sources of information, especially official public records, and encoded relatively complex categories of actions and entities involved. I validated the resulting data in two ways, to ensure both accurate categorization in relation to human coding, and consistency with regard to expected relationships from previous literature. I also took a liberal approach to incorporating those data into hypothesis tests, seeking to cover as many relevant combinations of covariates and model parameters as possible.

In order to accomplish these tasks, I have made use of several types of emerging computational methods, from web scraping public archives to processing natural language, to using machine learning for classification and prediction of the main variable measures, to random sampling, large-scale repetition and summarization of statistical tests. In each of these

steps, I have sought to leverage existing open-source tools wherever applicable, while also developing my own algorithms and code where necessary, to accomplish the tasks I have undertaken in both principled and practical ways. I have discussed the pros and cons of the various options that I investigated and tested, many of which might be useful for a wide variety of other applications in political studies and public policy research.

In particular, I discussed the growing advantages of studying government action directly through public records, which are increasingly available in digital archives, and which report formally enacted decisions in consistent legal terminology. There are many other policy areas in which public records may be underutilized, but where digital archives are becoming accessible and comprehensive enough to facilitate new forms of investigation. In order to be able to use large volumes of government documents, which might contain sparse information about specific decisions or actions of interest, some form of automated text analysis or natural language processing is also becoming increasingly essential. I compared several methods for classifying documents into categories of actions and target entities. Similar methods could be used to extract many other types of events and their participants from text records, including who gains or loses from legislation and policy decisions that do not involve institutional change.

There is still much work to be done to develop computational methods for extracting events from documents, especially in less supervised or less resource-intensive ways. Some projects may require more complex event details, including the source and destination of a transfer, the instrument or mechanism of implementation, and/or the location or duration of the action. Many of the same preprocessing steps and linguistic components used in this study could be used to extract more complex event information as well. I wrote the code for the rule-based pattern matching system in this project, but parsed the text with existing open-source tools, and leveraged other existing resources like the WordNet lexicon. I am developing a general-purpose version of the information extraction code used in this project, which might be used by other researchers to extract other types of events from document sources.

To ensure that the project's data were consistently coded and good reflections of the phenomena of interest, I presented two forms of validation, one akin to "criterion validity" (in the sense that the data accurately capture a model human-coded sample), and the other akin to "construct validity" (demonstrating correlations to other real-world phenomena that are theoretically expected to be associated with the concept of interest). The accuracy tests showed that reasonable high scores – over 85% on average – can be achieved even when automatically encoding documents into multiple action and target categories across multiple countries' decrees at the same time. I achieved the best scores by combining several different types of classifiers into an ensemble, making use of multiple types of input features, and adding some intuitive rules.

For the second form of validation, I used control variables from the literature on presidential power and executive orders, to show that the different categories of decrees from Chapter 3 capture expected relationships to established explanations for unilateral executive action. I used those tests to help eliminate alternative explanations for leaders' use of decrees, by distinguishing different categories of decrees that leaders use to get more things done in the face

of gridlock, from the decrees they use to expand their own power against rival authorities. Validating measures from multiple angles, to fully contextualize a set of decisions or events in terms of previous explanations and distinguish which components represent the current theory, might help to better connect related studies. Different explanations for political leaders' behavior do not need to compete or be mutually exclusive. Instead, different motivations might play greater roles under different circumstances, and distinguishing indications of different strategies might help us dissect leaders' actions in ways that add up to greater knowledge collectively.

Finally, the approach I took to the final hypothesis tests might be useful to other researchers as well. Linear regression is a very common and useful approach to modeling hypothesized relationships, but it is a fairly blunt instrument that probably only very loosely approximates most relationships between social phenomena. Regression models rely on many assumptions and parameters to produce this approximation, and regression estimates can vary widely with slight changes in the composition of the model, even if there is a fairly direct and meaningful empirical relationship that is well captured in the underlying data. The instability of regression models has contributed to growing concerns about researchers' ability to select models for analysis that support their hypotheses over others that would not, or "p-hack" their models until they show significant results (Nuzzo 2015).

If it becomes more common practice to permute covariates and parameters in many more combinations, and to report the results of those tests in terms of aggregate distributions rather than individual model estimates, we might take another step toward transparency and reproducibility in social scientific analysis. The process of randomly permuting model parameters is quite simple; the more important step to encourage this practice might be the presentation of digestible results. In Chapters 5 and 6, I presented the results of my hypothesis tests in new ways, plotting distributions of coefficients and p-values across model specifications, to condense a large number of tests into a readable format that conveys more information about the empirical relationships than individual model estimates alone.

My approach to randomly permuting covariates and parameters, and presenting the results in terms of aggregate distributions rather than single-model estimates, might be seen as an addition to emerging solutions for transparency and reproducibility in social science research. The more common approach is to tie one's hands by defining a "pre-analysis plan," before the researcher has run any hypothesis tests, which indicates exactly which model specifications will be used. However, it may be difficult to pre-specify enough specific models ahead of time to cover all relevant tests, especially when some hypotheses may develop based on findings from others, in ways that are not arbitrary model tweaks but important and substantively distinct alternatives (Olken 2015). The vast space of potential models also suggests that pre-analysis plans will still contain many fairly arbitrary decisions, which might reasonably have been specified another way, and which therefore might not lead to more reproducible results.

My approach addresses many of the same concerns and reflects similar intentions to pre-analysis plans, except that I use modern computing power to go the other way. Pre-analysis plans permit only a few models to be run, so that the results reported do not reflect a selective window

into a much larger hidden process. I instead propose that researchers can and should run as many tests as seem reasonable and feasible given the available data and hypothesized relationships, they should simply report *all of the results*. In order to do so, we need to move beyond reporting regression coefficients, standard errors and p-values for a single regression model at a time. I hope that my approach will encourage other researchers to aggregate and summarize estimates over multiple regression models, and I am making a general-purpose version of the code I developed for this task available for other researchers to use as well.

In sum, I have sought to present a well developed and connected sequence of steps from theoretical concept formation through measurement to analytic results, in order to isolate the mechanisms at the heart of the theory, and produce conceptually valid and reproducible results. This study addresses highly contentious issues about why government leaders seek more power, what they intend to do with it, and how they should be held accountable for its use. The allocation of authority at the highest levels of government cannot be experimentally manipulated, in most settings, meaning that these questions must be tackled through observation. Studying presidents' actions and fates also requires collecting data from multiple countries over multiple years, since each country only has one president at a time, capable of issuing decrees that are authoritative, publicly accessible, and can be reasonably attributed to one main political actor.

Within these constraints, I have sought to use digital sources of information and emerging computational methods to collect and dissect as much granular data as possible, out of a relatively modest number of countries, leaders, and events over a limited period of time. The results would naturally be stronger with the inclusion of more countries, leaders, and time, but I have sought to provide a framework for analysis that could be replicated by others, along with results that are thorough, transparent, and robust, in the hope that they will be reproducible in related work. There may be many other ways to measure the central concepts in this study, such as the objectivity or predictability of leaders' fates. The conclusions would also be strengthened by studies of other types of actors, beyond government heads, to see if they react in similar ways to threats of future punishment, depending on whether that punishment appears to fit the individuals' true actions, so that other actors can adjust their behavior accordingly.

I plan to continue developing the methods used in this study, and to expand the substantive applications to related policy areas, as well as to different regions and political contexts. I welcome collaboration with other researchers along similar lines, and hope that many of the contributions from this project will be useful to others.

References

- Daron Acemoglu and James A. Robinson. 2001. "A Theory of Political Transitions." *The American Economic Review* 91(4): 938-963.
- Emily Achtenberg and Rebel Currents. 2015. "Our Brand Is Impunity: Why is the U.S. Harboring Bolivia's Most Wanted Fugitive?" *North American Congress on Latin America*, *nacla.org*, October 29, 2015. <https://nacla.org/blog/2015/10/30/our-brand-impunity-why-us-harboring-bolivia%E2%80%99s-most-wanted-fugitive>
- Andrés Mejía Acosta and John Polga-Hecimovich. 2011. "Coalition Erosion and Presidential Instability in Ecuador." *Latin American Politics and Society* 53(2): 87-111.
- Ed Adamczyk. 2016. "Peru's ex-spy chief sentenced in forced disappearances, burning of bodies." *United Press International (UPI)*, September 28, 2016. https://www.upi.com/Top_News/World-News/2016/09/28/Perus-ex-spy-chief-sentenced-in-forced-disappearances-burning-of-bodies/5041475074884/
- Robert Adcock and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3): 529-546.
- Natalie Ahn. 2017. "Inducing Event Types and Roles in Reverse: Using Function to Discover Theme." *Proceedings of the workshop on Events and Stories in the News, Association for Computational Linguistics (ACL)*.
- Scott L. Althaus, Nathaniel Swigger, Svitlana Chernykh, David J. Henry, Sergio C. Wals, and Christopher Tiwald. 2011. "Assumed Transmission in Political Science: A Call for Bringing Description Back In." *Journal of Politics* 73 (4):1065-1080.
- Mercedes Alvaro. 2014. "Former Ecuador President Convicted of Embezzlement." *The Wall Street Journal*, May 30, 2014. <https://www.wsj.com/articles/former-ecuador-president-convicted-of-embezzlement-1401477777>
- Steven Ambrus. 1996. "Colombia Congress Halts Impeachment." *Los Angeles Times*, June 13, 1996. http://articles.latimes.com/1996-06-13/news/mn-14453_1_colombia-congress
- Barry Ames. 1987. *Political Survival: Politicians and Public Policy in Latin America*. University of California Press.
- Santiago Anria. 2017. "Evo Morales wants to change the law so he can remain president. Is Bolivia's democracy in danger?" *The Washington Post: Monkey Cage*, November 28, 2017. https://www.washingtonpost.com/news/monkey-cage/wp/2017/11/28/bolivias-president-thinks-hes-irreplaceable-what-does-this-mean-for-democracy-there/?utm_term=.0813dc4a7e2c
- Ronald P. Archer and Matthew Soberg Shugart. 1997. "The Unrealized Potential of Presidential Dominance in Colombia." In Scott Mainwaring and Matthew Soberg Shugart, Eds. *Presidentialism and Democracy in Latin America*. Cambridge University Press.
- Mabel Azcui. 2004. "El Parlamento de Bolivia autoriza un juicio político contra el ex presidente Sánchez de Lozada." *El País*, October 15, 2004.

- https://elpais.com/diario/2004/10/15/internacional/1097791227_850215.html
- Santiago Basabe-Serrano, Simón Pachano, and Andrés Mejía Acosta. 2010. "La democracia inconclusa: Derechos fundamentales, instituciones políticas y rendimientos gubernamentales en Ecuador (1979-2008)." *Revista de Ciencia Política* 30(1): 65-85.
- Alexander Baturo. 2010. "The Stakes of Losing Office, Term Limits and Democracy." *British Journal of Political Science* 40(3): 635-662.
- Eva Bellin. 2012. "Reconsidering the Robustness of Authoritarianism in the Middle East: Lessons from the Arab Spring." *Comparative Politics* 44(2): 127-149.
- Ali Berzon, Sarah Gordon, Aaron Selverston and Marlena Telvick. 2005. "Peru – The Curse of Inca Gold: Montesinos's Web." *PBS Frontline World*.
<http://www.pbs.org/frontlineworld/stories/peru404/pmontesinos.html>
- Carlo Biagioli, Enrico Francesconi, Andrea Passerini, and Claudia Soria. 2005. "Automatic semantics extraction in law documents." *Proceedings of the Tenth International Conference on Artificial Intelligence and Law*.
- Carles Boix and Milan W. Svobik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships." *The Journal of Politics* 75(2): 300-316.
- Doug Bond, Joe Bond, Churl Oh, J. Craig Jenkins, and Charles Lewis Taylor. 2003. "Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development." *Journal of Peace Research* 40(6):733-745.
- Francis Bond and Kyonghee Paik. 2012. "A survey of wordnets and their licenses." *Proceedings of the 6th Global WordNet Conference (GWC)*.
- Sally Bowen. 2000. *The Fujimori File: Peru and Its President 1990-2000*. Lima: Montiro Peru.
- Arthur Brice. 2009. "Bolivia vote on constitution could help president." *CNN* January 25, 2009.
<http://www.cnn.com/2009/WORLD/americas/01/23/bolivia.referendum/index.html>
- Ian Bremmer. 2017. "The 5 Countries That Illustrate the Decline of the Latin American Left." *Time*, April 3, 2017. <http://time.com/4719076/ecuador-venezuela-latin-america-left-wing/>
- Sibylla Brodzinsky. 2016a. "Venezuela president declares economic emergency as inflation hits 141%." *The Guardian* January 15, 2016.
<https://www.theguardian.com/world/2016/jan/16/venezuela-president-declares-economic-emergency-as-inflation-hits-141>
- Sibylla Brodzinsky. 2016b. "Colombia's peace deal rejection returns Álvaro Uribe to political limelight." *The Guardian*, October 7, 2016.
<https://www.theguardian.com/world/2016/oct/07/peace-deal-rejection-returns-alvaro-uribe-to-political-limelight>
- James Brooke. 1994. "Venezuela's Tight Grip Has Investors Gasping." *The New York Times*, August 19, 1994. <https://www.nytimes.com/1994/08/19/business/venezuela-s-tight-grip-has-investors-gasping.html>
- Stefanie Brüninghaus and Kevin D. Ashley. 2001. "Improving the representation of legal case texts with information extraction methods." *Proceedings of the 8th International Conference*

on Artificial Intelligence and Law.

- Bruce Bueno de Mesquita, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival*. MIT Press.
- Jo-Marie Burt. 2009. "Guilty as Charged: The Trial of Former Peruvian President Alberto Fujimori for Human Rights Violations." *International Journal of Transitional Justice* 3(3): 384-405.
- Nick Caistor. 2000. "Who is Vladimiro Montesinos?" *BBC News*, November 22, 2000. <http://news.bbc.co.uk/2/hi/americas/992770.stm>
- Hubert Cam. 1991. "Former president of Peru impeached." *United Press International*, October 3, 1991. <https://www.upi.com/Archives/1991/10/03/Former-president-of-Peru-impeached/1845686462400/>
- Ian Campbell. 2000. "The strongman's return." *The Guardian*, August 8, 2000. <https://www.theguardian.com/theguardian/2000/aug/09/guardiananalysispage>
- Claire Cardie and John Wilkerson. 2008. "Text Annotation for Political Science Research." *Journal of Information Technology & Politics* 5(1): 1-6.
- John M. Carey. 2003. "The Reelection Debate in Latin America." *Latin American Politics and Society* 45(1): 119-133.
- John M. Carey and John Polga-Hecimovich. 2005. "Primary Elections and Candidate Strength in Latin America." *Journal of Politics* 68(3): 530-543.
- John M. Carey and Matthew Soberg Shugart. 1998. *Executive Decree Authority*. Cambridge University Press.
- Ryan E. Carlin, Gregory J. Love, Timothy Hellwig, Cecilia Martinez-Gallardo, and Matthew M. Singer. 2016. "Executive Approval Database 1.0." www.executiveapproval.org.
- Rory Carroll. 2008. "A family affair." *The Guardian*, December 5, 2008. <https://www.theguardian.com/world/2008/dec/06/chavez-family-venezuela>
- Tommaso Caselli, Valentina Bartalesi Lenzi, Rachele Sprugnoli, Emanuele Pianta, and Irina Prodanof. 2011. "Annotating Events, Temporal Expressions and Relations in Italian: the It-TimeML Experience for the Ita-TimeBank." *Proceedings of the Fifth Law Workshop (Law V) of the Association for Computational Linguistics*.
- Andrew Cawthorne and Alexandra Ulmer. 2016. "Venezuela protests against Maduro escalate, dozens injured." *Reuters*, October 25, 2016. <https://www.reuters.com/article/us-venezuela-politics/venezuela-protests-against-maduro-escalate-dozens-injured-idUSKCN12Q0B6>
- Teresa Cespedes and Terry Wade. 2009. "Peru's Fujimori gets 25 years prison for massacres." *Reuters*, April 7, 2009. <https://www.reuters.com/article/us-peru-fujimori/perus-fujimori-gets-25-years-prison-for-massacres-idUSTRE5363RH20090407>
- Nathanael Chambers. 2013. "Event schema induction with a probabilistic entity-driven model." *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*.
- Nathanael Chambers and Dan Jurafsky. 2011. "Template-based information extraction without the templates." *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*.

- Tin Tin Cheng, Jeffrey Leonard Cua, Mark Davies Tan, Kenneth Gerard Yao, and Rachel Edita Roxas. 2009. "Information extraction from legal documents." *Proceedings of the Eighth International Symposium on Natural Language Processing (SNLP)*.
- Jackie Chi Kit Cheung, Hoifung Poon, and Lucy Vanderwende. 2013. "Probabilistic frame induction." *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.
- Giacomo Chiozza and H.E. Goemans. 2011. *Leaders and International Conflict*. Cambridge University Press.
- Laura Chiticariu, Yunyao Li, and Frederick R. Reiss. 2013. "Rule-based information extraction is dead! Long live rule-based information extraction systems!" *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Catherine M. Conaghan. 2005 *Fujimori's Peru: Deception in the Public Sphere*. University of Pittsburgh Press.
- Catherine M. Conaghan. 2008. "Ecuador: Correa's Plebiscitary Presidency." *Journal of Democracy* 19(2): 46-60.
- Catherine M. Conaghan. 2012. "Prosecuting Presidents: The Politics within Ecuador's Corruption Cases." *Journal of Latin American Studies* 44(4): 649-678.
- Phillip J. Cooper. 2002. *By Order of the President: The Use and Abuse of Executive Direct Action*. University Press of Kansas.
- Brian F. Crisp. 1998a. "Presidential Decree Authority in Venezuela." In John M. Carey and Matthew Soberg Shugart, Eds. *Executive Decree Authority*. Cambridge University Press.
- Brian F. Crisp. 1998b. "Institutional Design and Compromised Legitimacy." In Damaris Canache and Michael R. Kulisheck, Eds. *Reinventing Legitimacy: Democracy and Political Change in Venezuela*. Westport, CT: Greenwood Press.
- Cesi Cruz, Philip Keefer, and Carlos Scartascini. 2016. "The Database of Political Institutions 2015 (DPI2015)." Inter-American Development Bank, IDB-DB-121 Edition.
- Eduardo Dargent. 2011. "Agents or Actors?: Assessing the Autonomy of Economic Technocrats in Colombia and Peru." *Comparative Politics* 43(3): 313-332.
- Emilia Diaz-Struck and Juan Forero. 2013. "Venezuelan president Maduro given power to rule by decree." *The Washington Post*, November 19, 2013.
https://www.washingtonpost.com/world/the_americas/venezuelan-president-maduro-given-power-to-rule-by-decree/2013/11/19/af304c3c-516b-11e3-9ee6-2580086d8254_story.html?utm_term=.d8975a72775b
- Pilar Domingo. 2010. "Judicialization of politics or politicization of the judiciary? Recent trends in Latin America." *Democratization* 11(1): 104-126.
- Patrick Duddy. 2015. "Political Crisis in Venezuela." Contingency Planning Memorandum, *Council on Foreign Relations (CFR) Center for Preventive Action*. March 30, 2015.
<https://www.cfr.org/report/political-crisis-venezuela>
- Evan Ellis. 2017. "Ecuador under President Moreno: Rethinking the U.S. relationship." *Global Americans*, May 26, 2017. <https://theglobalamericans.org/2017/05/ecuador-president->

- moreno-rethinking-u-s-relationship-latin-americas-left/
- Steve Ellner. 2011. "The Distinguishing Features of Latin America's New Left in Power: The Chávez, Morales, and Correa Governments." *Latin American Perspectives* 39(1): 96-114.
- Brian Ellsworth and Julia Symmes Cobb. 2017. "Venezuela-U.S. relations at lowest point ever: Maduro." *Reuters*, August 22, 2017. <https://www.reuters.com/article/us-venezuela-politics/venezuela-u-s-relations-at-lowest-point-ever-maduro-idUSKCN1B225P>
- Anthony Faiola. 2007. "Ex-President of Bolivia Faces Suit in U.S." *The Washington Post*, September 26, 2007. <http://www.washingtonpost.com/wp-dyn/content/article/2007/09/25/AR2007092502252.html>
- Linda Farthing. 2016. "One Step Forward for Bolivia's Gas War Victims." *North American Congress on Latin America, nacla.org*, July 13, 2016. <https://nacla.org/news/2016/07/13/one-step-forward-bolivia%E2%80%99s-gas-war-victims>
- Gustavo A. Flores-Macías. 2010. "Statist vs. Pro-Market: Explaining Leftist Governments' Economic Policies in Latin America." *Comparative Politics* 42(4): 413-433.
- Juan Forero. 2010. "Venezuelan legislature grants Chavez decree powers." *The Washington Post*, December 19, 2010. <http://www.washingtonpost.com/wp-dyn/content/article/2010/12/18/AR2010121802933.html>
- Jonathan Franklin. 2007. "Chile orders Fujimori back to Peru." *The Guardian*, September 21, 2007. <https://www.theguardian.com/world/2007/sep/21/chile.jonathanfranklin>
- Kenneth Freed. 1993. "Venezuelan Senate Orders President Tried: Latin America: By unanimous vote, Perez is suspended while Supreme Court weighs corruption charges against him." *Los Angeles Times*, May 22, 1993. http://articles.latimes.com/1993-05-22/news/mn-38532_1_supreme-court
- Natalie Gallón, Stefano Pozzebon and Euan McKirdy. 2017. "Venezuela says it will split from OAS as unrest continues." *CNN*, April 27, 2017. <https://www.cnn.com/2017/04/27/americas/venezuela-leaving-oas/index.html>
- Eduardo A. Gamarra. 1997. "Hybrid Presidentialism and Democratization: The Case of Bolivia." In Scott Mainwaring and Matthew Soberg Shugart, Eds. *Presidentialism and Democracy in Latin America*. Cambridge University Press.
- Jennifer Gandhi and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40(11): 1279-1301.
- Mercedes García Montero. 2001. "La Década de Fujimori: Ascenso, Mantenimiento y Caída de un Líder Antipolítico." *América Latina Hoy: Revista de Ciencias Sociales* 29: 49-86.
- Todd F. Gaziano. 2001. "The Use and Abuse of Executive Orders and Other Presidential Directives." *Texas Review of Law & Politics* 5(2): 267-316.
- Barbara Geddes, Joseph Wright, and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions: A New Data Set." *Perspectives on Politics* 12(2): 313-331.
- Sean M. Gerrish and David M. Blei. 2012. "How They Vote: Issue-Adjusted Models of Legislative Behavior." *Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS)*.

- Daniel Gildea and Daniel Jurafksy. 2002. "Automatic Labeling of Semantic Roles." *Computational Linguistics* 28(3): 245-288.
- Jim Giles. 2005. "Internet encyclopaedias go head to head." *Nature* 438(7070): 900-1.
- Tom Ginsburg, James Melton, and Zachary Elkins. 2011. "On the Evasion of Executive Term Limits." *William and Mary Law Review* 52(6): 1807-1872.
- H.E. Goemans, Kristian Skrede Gledistch and Giacomo Chiozza. 2009. "Introducing Archigos: A Data Set of Political Leaders." *Journal of Peace Research* 46(2): 269-283.
- Laurie Goering. 1997. "Ecuadorans Say 'El Loco' Took The Money And Ran." *Chicago Tribune*, April 2, 1997. http://articles.chicagotribune.com/1997-04-02/news/9704020115_1_face-corruption-charges-fabian-alarcon-illegal-coup
- Efrain Gonzales de Olarte. 1993. "Economic Stabilization and Structural Adjustment under Fujimori." *Journal of Interamerican Studies and World Affairs* 35(2): 51-80.
- A. Gonzalez-Agirre, E. Laparra, and G. Rigau. 2012. "Multilingual central repository version 3.0." *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC)*.
- Joshua Goodman. 2006. "Colombia's Ambassador to U.S. Resigns." *The Washington Post*, July 12, 2006. <http://www.washingtonpost.com/wp-dyn/content/article/2006/07/12/AR2006071200103.html>
- Gustavo Gorriti. 1994. "The Betrayal of Peru's Democracy: Montesinos as Fujimori's Svengali." *Covert Action Quarterly*, Summer 1994.
- Gustavo Gorriti. 2015. "Espías." *El País*, March 25, 2015. <https://idl-reporteros.pe/espias/>
- Jeffrey Green. 2009. *The Eyes of the People: Democracy in an Age of Spectatorship*. Oxford University Press.
- Glenn Greenwald. 2012. "America's refusal to extradite Bolivia's ex-president to face genocide charges." *The Guardian*, September 9, 2012. <https://www.theguardian.com/commentisfree/2012/sep/09/america-refusal-extradite-bolivia>
- Justin Grimmer. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18:1-35.
- Justin Grimmer and Brandon M. Stewart. 2013. "Text as data: The promise and pitfalls of automated content analysis methods for political texts." *Political Analysis* 21(3):267-97.
- Samuel R. Gross and Barbara O'Brien. 2008. "Frequency and Predictors of False Conviction: Why We Know So Little, and New Data on Capital Cases." *Journal of Empirical Legal Studies* 5(4): 927-962.
- Debarati Guha-Sapir. 2018. *EM-DAT: The Emergency Events Database*. Université catholique de Louvain (UCL) – CRED. Brussels, Belgium. Accessed May 21, 2018. <http://www.emdat.be>
- Andrew Gumbel. 2005. "Ecuador revolts at the corrupt power behind President." *The Independent*, April 19, 2005. <http://www.independent.co.uk/news/world/americas/ecuador-revolts-at-the-corrupt-power-behind-president-495460.html>
- Kirk A. Hawkins. 2010. "Populism in Venezuela: the rise of Chavismo." *Third World Quarterly*

- 24(6): 1137-1160.
- Kirk A. Hawkins. 2016. "Responding to radical populism: Chavismo in Venezuela." *Democratization* 23(2): 242-262.
- Victor J. Hinojosa and Aníbal Pérez-Liñán. 2003 "Presidential Impeachment and the Politics of Survival: The Case of Colombia." In Jody C. Baumgartner and Naoko Kada, Eds. *Checking Executive Power: Presidential Impeachment in Comparative Perspective*. Westport, CT: Praeger Publishers.
- Kathryn Hochstetler and David Samuels. 2011. "Crisis and Rapid Reequilibration: The Consequences of Presidential Challenge and Failure in Latin America." *Comparative Politics* 43(2): 127-145.
- Daniel Hopkins and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54(1):229-47.
- William Howell and David Lewis. 2002. "Agencies by presidential design." *Journal of Politics* 64(4):1095-114.
- William G. Howell. 2003. *Power without persuasion: The politics of direct presidential action*. Princeton University Press.
- Human Rights Watch. 2012. "Tightening the Grip: Concentration and Abuse of Power in Chávez's Venezuela." *Human Rights Watch Report*, July 17, 2012.
<https://www.hrw.org/report/2012/07/17/tightening-grip/concentration-and-abuse-power-chavez-venezuela>
- Karl Magnus Johansson and Jonas Tallberg. 2010. "Explaining Chief Executive Empowerment: EU Summitry and Domestic Institutional Change." *West European Politics* 33(2): 208-236.
- Peter Jackson, Khalid Al-Kofahi, Alex Tyrrell, and Arun Vachher. 2003. "Information extraction from case law and retrieval of prior cases." *Artificial Intelligence* 150(1-2):239-290.
- Dan Keane. 2006. "Bolivian Uprising Still an Open Wound." *The Washington Post*, October 26, 2005. <http://www.washingtonpost.com/wp-dyn/content/article/2006/10/26/AR2006102600205.html>
- Hans Mathias Kepplinger. 2002. "Mediatization of Politics: Theory and Data." *Journal of Communication* 52(4):972-986.
- Jin-Dong Kim, Tomoko Ohta and Jun'ichi Tsujii. 2008. "Corpus annotation for mining biomedical events from literature." *BMC Bioinformatics* 9(10).
- Young Hun Kim and Donna Bahry. 2008. "Interrupted Presidencies in Third Wave Democracies." *The Journal of Politics* 70(3): 807-22.
- Young Hun Kim. 2014. "Impeachment and presidential politics in new democracies." *Democratization* 21(3): 519-53.
- Gary King and Will Lowe. 2003. "An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design." *International Organization* 57(03):617-642.
- Dan Klein and Christopher D. Manning. 2003. "Accurate unlexicalized parsing." *Proceedings of the 41st Meeting of the Association for Computational Linguistics (ACL)*.

- Roger Koppl and Megahn Sacks. 2012. "The Criminal Justice System Creates Incentives for False Convictions." *Criminal Justice Ethics* 32(2): 126-162.
- Jessica Kwong. 2017. "U.S. is the 'real threat to world security and peace,' says Bolivian president." *Newsweek*, November 15, 2017. <http://www.newsweek.com/us-real-threat-world-security-and-peace-says-bolivian-president-712996>
- Charles Lane. 1992. "The 'Self-Coup' that Rocked Peru." *Newsweek*, April 19, 1992. <http://www.newsweek.com/self-coup-rocked-peru-197280>
- Steven Levitsky and Maxwell A. Cameron. 2003. "Democracy Without Parties? Political Parties and Regime Change in Fujimori's Peru." *Latin American Politics and Society* 45(3): 1-33.
- Serge F. Kovaleski. 1999. "Venezuelan Vote Gives President New Powers." *The Washington Post*, December 16, 1999. <http://www.washingtonpost.com/wp-srv/WPcap/1999-12/16/111r-121699-idx.html>
- Margarita López Maya. 2014. "Venezuela: The Political Crisis of Post-Chavismo." *Social Justice* 40(4): 68-87.
- Kintto Lucas. 2003. "Ecuador: Ex-President Finds Refuge from Corruption Charges." *Inter Press Service*, July 30, 2003. <http://www.ipsnews.net/2003/07/ecuador-ex-president-finds-refuge-from-corruption-charges/>
- Beatriz Magaloni and Ruth Kricheli. 2010. "Political Order and One-Party Rule." *Annual Review of Political Science* 13: 123-143.
- Gideon Maltz. 2007. "The Case for Presidential Term Limits." *Journal of Democracy* 18(1): 128-142.
- Francesco Manetto and Sally Palomino. 2017. "La fiscalía de Colombia confirma pagos de Odebrecht a las campañas de Santos y de su rival." *El País*, July 14, 2017. https://elpais.com/internacional/2017/07/12/colombia/1499886477_264659.html
- Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. "The Stanford CoreNLP natural language processing toolkit." *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*.
- Marie-Catherine de Marneffe, Bill MacCartney and Christopher D. Manning. 2006. "Generating Typed Dependency Parses from Phrase Structure Parses." *5th International Conference on Language Resources and Evaluation (LREC 2006)*.
- Monty G. Marshall, Ted Robert Gurr, and Keith Jagers. 2017. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016." Dataset Users' Manual, Center for Systemic Peace, July 25, 2017. <http://www.systemicpeace.org/inscr/p4manualv2016.pdf>
- Monty G. Marshall and Donna Ramsey Marshall. 2017. "Coup D'etat Events, 1946-2016: Codebook." Center for Systemic Peace, May 23, 2017. <http://www.systemicpeace.org/inscr/CSPCoupsCodebook2016.pdf>
- William P. Marshall. 2008. "Eleven Reasons Why Presidential Power Inevitably Expands and Why It Matters." *Boston University Law Review* 88(2): 505-522.

- Leiv Marsteintredet and Einar Berntzen. 2008. "Reducing the Perils of Presidentialism in Latin America through Presidential Interruptions." *Comparative Politics* 41(1): 83-101.
- Gabriel X. Martinez. 2006. "The political economy of the Ecuadorian financial crisis." *Cambridge Journal of Economics* 30(4): 567-585.
- John D. Martz. 1999. "Political Parties and Candidate Selection in Venezuela and Colombia." *Political Science Quarterly* 114(4): 639-659.
- Kenneth Mayer. 1999. "Executive orders and presidential power." *Journal of Politics* 61(2): 445-66.
- Kenneth Mayer. 2001. *With the stroke of a pen: Executive orders and presidential power*. Princeton University Press.
- Kenneth Mayer and Kevin Price. 2002. "Unilateral presidential powers: Significant executive orders, 1949-99." *Presidential Studies Quarterly* 32(2):367-86.
- Rene Antonio Mayorga. 2006. "Outsiders and Neopopulism: The Road to Plebiscitary Democracy." In Scott Mainwaring, Ana Maria Bejarano, and Eduardo Pizarro Leongomez, Eds. *The Crisis of Democratic Representation in the Andes*. Stanford University Press, Chapter 5.
- Cynthia McClintock. 1999. "The Decimation of Peru's Sendero Luminoso." In Cynthia J. Arnson, Ed. *Comparative Peace Processes in Latin America*. Woodrow Wilson Center Press.
- Cynthia McClintock and Fabian Vallas. 2003. *The United States and Peru: Cooperation at a Cost*. Taylor & Francis Books, Inc.
- Jennifer L. McCoy. 2017. "Venezuela's controversial new Constituent Assembly, explained." *The Washington Post: Monkey Cage*, August 1, 2017.
https://www.washingtonpost.com/news/monkey-cage/wp/2017/08/01/venezuelas-dubious-new-constituent-assembly-explained/?utm_term=.4a6e03bc9c70
- Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. 2010. "Google Ngram Viewer". Culturomics Users Guide accompanying article in *Science*, published online 12/16/2010. <http://culturomics.rice.edu/resources/google-ngram-viewer/>
- Teresa M. Miguel-Stearns. 2011. "The Digital Legal Landscape in South America: Government Transparency and Access to Information." *Yale Law School – Lilian Goldman Law Library, Yale University*. June 16, 2011.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1865891
- Slava Mikhaylov, Michael Laver, and Kenneth R. Benoit. 2012. "Coder reliability and misclassification in the human coding of party manifestos." *Political Analysis* 20(1):78-91.
- Michael E. Miller. 2010. "Álvaro Uribe, Ex-Prez of Colombia, Named Visiting Scholar at Georgetown Despite Protests." *Miami New Times*, September 29, 2010.
<http://www.miaminewtimes.com/news/andaacute-lvaro-uribe-ex-prez-of-colombia-named-visiting-scholar-at-georgetown-despite-protests-6527950>
- Yusuke Miyao, Rune Sætre, Kenji Sagae, Takuya Matsuzaki, and Jun'ichi Tsujii. 2008. "Task-

- oriented evaluation of syntactic parsers and their representations.” *Proceedings of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT)*.
- Terry M. Moe and William G. Howell. 1999. “The presidential power of unilateral action.” *Journal of Law, Economics, and Organization* 15:132–79.
- Mery Mogollon and Chris Kraul. 2017. “As Venezuelan death toll rises, President Nicolas Maduro bans protests before vote to rewrite the constitution.” *Los Angeles Times*, July 27, 2017. <http://www.latimes.com/world/mexico-americas/la-fg-venezuela-protests-20170727-story.html>
- Juan Antonio Morales and Jeffrey D. Sachs. 1987. “Bolivia’s Economic Crisis.” In Jeffrey D. Sachs, Ed. *Developing Country Debt and the World Economy*. University of Chicago Press.
- Andrew Morgan. 2009. “Bolivia ex-president goes on trial *in absentia* on genocide charges.” *JURIST: Paper Chase*, May 19, 2009. <http://www.jurist.org/paperchase/2009/05/bolivia-ex-president-goes-on-trial-in.php>
- Anthony Mughan. 2000. *Media and the Presidentialization of Parliamentary Elections*. New York, NY: Palgrave Macmillan.
- Yusuke Murakami. 2007. *Perú en la era del chino: la política no institucionalizada y el pueblo en busca de un salvador*. Lima, Peru: Instituto de Estudios Peruanos (IEP).
- Peter F. Nardulli, Michael Martin, Michael Slana, Sina Toosi, Joseph Bajjalieh. 2013. “The Coup D’etat Project”. White Paper, Cline Center for Democracy, September 2013. http://www.clinecenter.illinois.edu/publication/white/event/documents/Coup_Project.pdf
- Nathaniel C. Nash. 1991. “Peru’s President is Trying to Tame a Maelstrom of Inflation and Poverty.” *The New York Times*, July 7, 1991. <http://www.nytimes.com/1991/07/07/world/peru-s-president-is-trying-to-tame-a-maelstrom-of-inflation-and-poverty.html?pagewanted=all>
- William Neuman. 2014. “President of Bolivia Claims Victory in Election.” *The New York Times* October 12, 2014. <https://www.nytimes.com/2014/10/13/world/americas/bolivia-seen-likely-to-re-elect-morales-as-president.html>
- Kiem-Hieu Nguyen, Xavier Tannier, Olivier Ferret, and Romaric Besancon. 2015. “Generative event schema induction with entity disambiguation.” *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*.
- Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, and Kristina Miler. 2015. “Tea Party in the House: A Hierarchical Ideal Point Topic Model and Its Application to Republican Legislators in the 112th Congress.” *Proceedings of the Association for Computational Linguistics (ACL)*.
- Regina Nuzzo. 2015. “How scientists fool themselves – and how they can stop.” *Nature: News Feature*, 7 October 2015.
- Guillermo O’Donnell. 1994. “Delegative Democracy.” *Journal of Democracy* 5(1): 55-69.
- Chitu Okoli. 2009. “A Brief Review of Studies of Wikipedia in Peer-Reviewed Journals.” *Proceedings of the Third International Conference on Digital Society*.

- Benjamin A. Olken. 2015. "Promises and Perils of Pre-Analysis Plans." *Journal of Economic Perspectives* 29(3): 61-80.
- Tricia D. Olsen, Leigh A. Payne, and Andrew G. Reiter. 2010. "The Justice Balance: When Transitional Justice Improves Human Rights and Democracy." *Human Rights Quarterly* 32(4): 980-1007.
- David G. Ortiz, Daniel J. Myers, N. Eugene Walls, and Maria-Elena D. Diaz. 2005. "Where Do We Stand with Newspaper Data." *Mobilization: An International Journal* 10 (3):397-419.
- Mark Osiel. 2000. "Why Prosecute? Critics of Punishment for Mass Atrocity." *Human Rights Quarterly* 22(1): 118-47.
- Joshua Partlow and Lucien Chauvin. 2009. "Peru's Fujimori Gets 25 Years." *The Washington Post*, April 8, 2009. <http://www.washingtonpost.com/wp-dyn/content/article/2009/04/07/AR2009040701345.html>
- Mark Peceny and Michael Durnan. 2008. "The FARC's Best Friend: U.S. Antidrug Policies and the Deepening of Colombia's Civil War in the 1990s." *Latin American Politics and Society* 48(2): 95-116.
- Carlos Pereira, Timothy J. Power, and Lucio Rennó. 2005. "Under what conditions do presidents resort to decree power? Theory and evidence from the Brazilian case." *Journal of Politics* 67(1):178-200.
- Anibal Pérez-Liñán. 2007. *Presidential Impeachment and the New Political Instability in Latin America*. Cambridge University Press.
- James Petras and Michael M. Brescia. 2000. "The FARC Faces the Empire." *Latin American Perspectives* 114(27): 134-142.
- Thomas Poguntke and Paul Webb. 2005. *The Presidentialization of Politics in Democratic Societies*. Oxford University Press.
- Eduardo Posada-Carbó. 1998. *Colombia: The Politics of Reforming the State*. Palgrave Macmillan.
- Jonathan Powell and Clayton Thyne. 2011. "Global Instances of Coups from 1950-Present." *Journal of Peace Research* 48(2): 249-259.
- Oleh Protsyk. 2004. "Ruling with decrees: presidential decree making in Russia and Ukraine." *Europe-Asia Studies* 56(5):637-660.
- Stephen Purpura and Dustin Hillard. 2006. "Automated classification of congressional legislation." *Proceedings of the International Conference on Digital Government Research*.
- Alfonso W. Quiroz. 2008. *Corrupt Circles: A History of Unbound Graft in Peru*. Woodrow Wilson Center Press.
- Clionadh Raleigh, Andrew Linke, Havard Hegre and Joakim Karlsen. 2010. "Introducing ACLED: An Armed Conflict Location and Event Dataset." *Journal of Peace Research* 47(5):651-660.
- Jess Rapp. 2017. "Profile: Former President Alberto Fujimori." *Peru Reports*, October 17, 2017. <https://perureports.com/2017/10/17/alberto-fujimori-profile/>
- Napoleon C. Reyes and Jurg Gerber. 2011. "Above the Law? A Comparative Study of National

- Prosecutions of Heads of State.” *Critical Criminology* 19(1): 43-73.
- Kenneth M. Roberts. 1996. “Economic Crisis and the Demise of the Legal Left in Peru.” *Comparative Politics* 29(1): 69-92.
- Margaret E. Roberts, Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, David G. Rand. 2014. “Structural Topic Models for Open-Ended Survey Responses.” *American Journal of Political Science* 58(4):1064-1082.
- Terence Roehrig. 2002. *The Prosecution of Former Military Leaders in Newly Democratic Nations: The Cases of Argentina, Greece and South Korea*. McFarland.
- Santiago Rodríguez and Jesús Carretero. 1996. “A formal approach to Spanish morphology: The COES tools.” *Procesamiento del Lenguaje Natural* 19.
- Roy Rosenzweig. 2006. “Can History Be Open Source? Wikipedia and the Future of the Past.” *Journal of American History* 93(1): 117-146.
- Sebastian Rotella. 1996. “Secrecy and Scandal in Peru.” *The Los Angeles Times*, September 26, 1996. http://articles.latimes.com/1996-09-26/news/mn-47803_1_drug-scandals
- James Rothwell, Harriet Alexander and Anna Schaverien. 2017. “Venezuela protests: 200,000 march against President Maduro as riots and looting spread across country.” *The Telegraph*, May 21, 2017. <https://www.telegraph.co.uk/news/2017/05/21/venezuela-protests-200000-march-against-president-maduro-riots/>
- Jorge Rueda. 2014. “Jaime Lusinchi: Politician who was unable to deal with an economic crisis in Venezuela and was later charged with corruption.” *The Independent*, May 23, 2014.
- Andreas Schedler. 2012. “Judgment and Measurement in Political Science.” *Perspectives on Politics* 10(1): 21-36.
- Diana Jean Schemo. 1996. “Ex-President of Venezuela Sentenced to 28 Months in Graft Case.” *The New York Times*, May 31, 1996. <http://www.nytimes.com/1996/05/31/world/ex-president-of-venezuela-sentenced-to-28-months-in-graft-case.html>
- Gregory Schmidt. 1998. “Presidential Usurpation or Congressional Preference? The Evolution of Executive Decree Authority in Peru.” In John M. Carey and Matthew Soberg Shugart, Eds. *Executive Decree Authority*. Cambridge University Press.
- Philip A. Schrodtt. 2006. “Twenty years of the Kansas Event Data System project.” *The Political Methodologist, Newsletter of the Political Methodology Section, APSA* 14(1):2–6.
- Philip A. Schrodtt. 2012. “Precedents, Progress, and Prospects in Political Event Data.” *International Interactions* 38(4): 546-569.
- Philip A. Schrodtt. 2014. “TABARI: Text Analysis by Augmented Replacement Instructions.” Parus Analytical Systems, TABARI Manual, Version 0.8.4B3.
- Orçun Selçuk. 2016. “Strong presidents and weak institutions: populism in Turkey, Venezuela and Ecuador.” *Southeast European and Black Sea Studies*. 16(4): 571-589.
- Aaron Shapiro. 2017. “Reform predictive policing.” *Nature: Comment*. January 25, 2017. <http://www.nature.com/news/reform-predictive-policing-1.21338>
- Dan Slater and Erica Simmons. 2012. “Coping by Colluding: Political Uncertainty and

- Promiscuous Powersharing in Indonesia and Bolivia.” *Comparative Political Studies* 46(11): 1366-1393.
- Susan C. Stokes. 1997. “Democratic Accountability and Policy Change: Economic Policy in Fujimori’s Peru.” *Comparative Politics* 29(2): 209-226.
- Lei Sha, Sujian Li, Baobao Chang, and Zhifang Sui. 2016. “Joint learning templates and slots for event schema induction.” In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.
- Ronald D. Sylvia and Constantine P. Danopoulos. 2010. “The Chávez phenomenon: Political change in Venezuela.” *Third World Quarterly* 24(1): 63-76.
- Mohammad Ayatollahi Tabaar. 2013. “Assessing (In)security after the Arab Spring: The Case of Egypt.” *Political Science & Politics* 46(4): 727-735.
- Maureen Taft-Morales. 2009. “Peru: Current Conditions and U.S. Relations.” *Congressional Research Service*, July 21, 2009. <https://fas.org/sgp/crs/row/R40716.pdf>
- Martín Tanaka. 2006. “Systems and Dilemmas of Democratic Representation: Peru and Venezuela.” In Scott Mainwaring, Ana María Bejarano, and Eduardo Pizarro Leongómez, Eds. *The Crisis of Democratic Representation in the Andes*. Stanford University Press.
- Hristo Tanev, Jakub Piskorski, and Martin Atkinson. 2008. “Real-Time News Event Extraction for Global Crisis Monitoring.” *Proceedings of the 13th International Conference on Applications of Natural Language to Information Systems, NLDB*.
- Julio C. Teehankee. 2016. “Weak State, Strong Presidents: Situating the Duterte Presidency in Philippine Political Time.” *Journal of Developing Societies* 32(3): 293-321.
- Matt Thomas, Bo Pang, and Lillian Lee. 2006. “Get out the vote: determining support or opposition from congressional floor-debate transcripts.” *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Robert Tibshirani. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society*. 58(1): 267-288.
- Hector Tobar. 2003. “Bolivian President Exits Amid Uprising.” *Los Angeles Times*, October 18, 2003. <http://articles.latimes.com/2003/oct/18/world/fg-bolivia18>
- Maarten Truysens and Patrick Van Eecke. 2014. “Legal aspects of text mining.” *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC)*.
- Theodor Tudoroiu. 2007. “Rose, Orange, and Tulip: The failed post-Soviet revolutions.” *Communist and Post-Communist Studies* 40(3): 315-342.
- Rodrigo Uprimny. 2003. “The constitutional court and control of presidential extraordinary powers in Colombia.” *Democratization* 10(4): 46-69.
- Arturo Valenzuela. 2004. “Latin American Presidencies Interrupted.” *Journal of Democracy* 15(4): 5-19.
- Nicolas van de Walle. 2002. “Africa’s Range of Regimes.” *Journal of Democracy* 13(2): 66-80.
- Mario Vargas Llosa. 1995. *El pez en el agua: Memorias*. Planeta Publishing Corp.
- Ludmila Vinogradoff. 1994. “Rafael Caldera suspende las garantías constitucionales y decreta el control de cambio y precios en Venezuela.” *El País*, June 29, 1994.

- https://elpais.com/diario/1994/06/29/internacional/772840806_850215.html
- Lucan Way. 2008. "The Real Causes of the Color Revolutions." *Democracy and Human Rights* 19(3): 55-69.
- Jeffrey R. Webber. 2010. "Carlos Mesa, Evo Morales, and a Divided Bolivia (2003-2005)." *Latin American Perspectives* 37(3): 51-70.
- David L. Weiden. 2011. "Judicial Politicization, Ideology, and Activism at the High Courts of the United States, Canada, and Australia." *Political Research Quarterly* 64(2): 335-347.
- Nils Weidmann. 2015. "On the accuracy of media-based conflict event data." *Journal of Conflict Resolution* 59(6):1129-49.
- Bill Weinberg. 2011. "The Bolivian Genocide Case: Ex-regime figures convicted, US shelters top fugitive." *Indian Country Today*, September 15, 2011.
<https://indiancountrymedianetwork.com/news/the-bolivia-genocide-case-ex-regime-figures-convicted-us-shelters-top-fugitive/>
- Kurt Weyland. 2001. "Clarifying a contested concept: Populism in the study of Latin American politics." *Comparative Politics* 34(1):1-22.
- Claire Wright. 2014. "Executives and emergencies: presidential decrees of exception in Bolivia, Ecuador, and Peru." Instituto de Estudios Latinoamericanos (IELAT), Universidad de Alcalá (60).
- Alberto Zuazo. 1985. "Bolivia suffering from tin crisis." *United Press International*, October 29, 1985. <https://www.upi.com/Archives/1985/10/29/Bolivia-suffering-from-tin-crisis/5589499410000/>