

The Dictator as Discriminating Monopsonist;
and the Rally Effect in Wartime Ukraine

Calvin Garner

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Reading Committee:

James D. Long, Chair

Victor Menaldo

Scott Radnitz

Program Authorized to Offer Degree:

Political Science

©Copyright 2024

Calvin Garner

University of Washington

Abstract

The Dictator as Discriminating Monopsonist; and the Rally Effect in Wartime Ukraine.

Calvin Garner

Chair of the Supervisory Committee:

James D. Long

Department of Political Science

This dissertation engages with questions from the study of authoritarianism and international conflict. Most authoritarian regimes today are multi-party authoritarian regimes, often with a dominant party that shapes policy and controls the flow of rents, and one or more parties in “loyal” opposition. This leads to a puzzle: how does the presence of multiple parties benefit the authoritarian leader, and why would someone join a loyal opposition party instead of the dominant party? I present a new model of the authoritarian regime as a discriminating monopsonist. By providing elites expressive benefits and rents through the choice of political parties, authoritarian regimes economize on the purchase of regime support. I test this theory by analyzing party switching by Duma deputies in Russia in 2006-2007 and find preliminary support for the model’s expectations.

The second question examines the “rally ‘round the flag” effect in Ukraine following Russia’s 2014 invasion. It adds to the existing literature by asking how ethnic identity and conflict proximity relate to attitudes towards the home state and the adversary. Using a lab-based implicit association test (IAT) and survey, we examine whether implicit biases, reflexive preferences that are hard to manipulate, match explicitly stated preferences for either Ukraine or Russia. We find that, on average, ethnic Ukrainians and Russians in Ukraine are explicitly and implicitly pro-Ukraine, although we observe slightly lower levels of pro-Ukraine bias among ethnic Russians. This study was first published in *International Studies Quarterly* (Erlich & Garner, [2021](#)).

Contents

- Acknowledgements** **11**

- Dedication** **13**

- 1 Introduction** **15**

- 2 Chapter 2** **21**
 - 2.1 Introduction 21
 - 2.2 Literature Review 24
 - 2.2.1 Gaps in existing theory 25
 - 2.2.2 Empirical problems: systemic opposition in Russia 30
 - 2.3 A theory of rents and ideas 35
 - 2.3.1 Strategic objectives 38
 - 2.3.2 A choice theoretic model of party affiliation 40
 - 2.3.3 A game theoretic model of party affiliation 42
 - 2.3.4 Results 45
 - 2.3.5 Ideological heterogeneity and sustainable systemic opposition parties 46
 - 2.4 Testable implications 48
 - 2.5 Conclusion 49

- 3 Chapter 3** **53**

3.1	Introduction	53
3.2	Case and Setting	54
3.3	Inference strategy	55
3.4	Russia under Putin: a relevant case?	56
3.4.1	Before Putin: central authority weakens under Yeltsin	56
3.4.2	Authoritarian consolidation under Putin	59
3.4.3	United Russia as a dominant party	60
3.4.4	The relevance of ideology in Russian politics	62
3.5	Empirical test: who joins “A Just Russia”?	63
3.5.1	Theoretical expectations	63
3.5.2	Data	64
3.5.3	Statistical model	65
3.5.4	Model Interpretation	70
3.6	Potential for additional empirical tests	74
3.7	Conclusion	75
	Appendices	77
3.A	Additional regression analyses	77
3.B	2003 Candidates Who Repeat	80
4	Chapter 4	83
4.1	Introduction	83
4.2	Case selection	87
4.3	Theoretical framework	91
4.3.1	Implicit association and explicit attitudes	96
4.4	Research Design	98
4.4.1	Subject Recruitment	99
4.4.2	Measuring implicit and explicit attitudes	100

4.4.3	Operationalizing predictors and other control variables	103
4.5	Results	104
4.5.1	Regression analysis	105
4.5.2	Dissociation in implicit and explicit attitudes	107
4.5.3	Additional findings	109
4.5.4	Robustness	110
4.6	Conclusions	111
Appendices		115
4.A	Alternative grouping of cities for H3	115
4.A.1	Justification for Kyiv vs. Other Cities comparison	115
4.A.2	Expectations of conflict: Kyiv vs. Other Cities	117
4.B	Summary Statistics	118
4.C	Political Party Categorization	121
4.D	IAT procedures	121
4.E	Results from statistical models	123
4.E.1	Regression tables	123
4.E.2	Predicted bias	125
4.F	Analysis weighted by recruitment category	126
4.F.1	Regression tables	126
4.F.2	Predicted bias (weighted data)	127
4.F.3	Weighted IAT data	128
5	Conclusion	129
	References	135

List of Figures

2.1	Changes in the ratio of democratic, single-party, and multi-party authoritarian regimes	22
2.2	The ideas-rents preference frontier (left) and a three-dimensional ideas space (right)	37
2.3	Equilibrium level of political participation and rents. the points q' and r' . . .	42
2.4	The classical divide-the-dollar game with the possibility of successful revolt.	44
2.5	The rent-ratio necessary to sustain an opposition party as a function of ideological distance	48
3.1	The rise in global oil prices in the early 2000s gave the Russian state new resources, enabling it to construct and support a permanent dominant party organization, United Russia.	61
3.2	First Difference Estimate: KPRF	72
3.3	First Difference Estimate: LDPR	73
3.B.1	The figure above shows what percentage of candidates who ran for the Duma in 2003 on a party list for a given party and then ran again in 2007, 2011, and 2016 for any party. The party labels include each party's total seats won out of total seats nominated in parentheses. The parties that won any seats in the 2003 election from the party list are boxed.	80

3.B.2	The figure above shows what percentage of candidates who ran for the Duma in 2003 in a single member district for a given party and then ran again in 2007, 2011, and 2016 for any party. The party labels include each party's total seats won out of total seats nominated in parentheses. The parties that won any seats in the 2003 election from the party list are boxed.	81
4.2.1	Data from September 2014 reported in Kulyk (2016) show a substantial rally effect as measured by respondents' reported changes in attitudes towards subjects associated with Ukraine and Russia. The values in the figure are the difference between the percentage of respondents who said that their attitude towards a subject improved a lot or a little and the percentage who said their attitude worsened a lot of a little.	89
4.2.2	The four cities in which IATs were conducted	90
4.4.1	IAT screen example	102
4.5.1	Predicted levels of bias and first differences from main statistical models (see Appendix Table ??). Implicit bias is measured in d -score units, while the predicted values for explicit bias reflect the range of possible values in $[-6, 6]$	107
4.5.2	The majority of IAT participants expressed pro-Ukraine implicit and explicit preferences (upper right).	108
4.A.1	Russian propaganda map depicting half of Ukraine as part of "Novorossiia."	116
4.A.2	Differences in average expectation of conflict compared to Kyiv.	117
4.B.1	The density and means of responses to the explicit and implicit bias questions by ethnicity and city (both variables normalized to match the scale of the IAT measurement $[-1.6, 1.6]$ to facilitate comparison). Positive values indicate a bias toward Ukraine. In all cases, the area under the curve greater than zero is greater than 0.50, indicating a bias in favor of Ukraine.	120
4.E.1	Predicted bias for all ethnicity, city, and recruitment categories from main empirical models.	125

4.F.1 Predicted preference from models interacting city and ethnicity, weighted using survey data on frequency of recruitment groups in the population of each city. 127

4.F.2 Explicit and Implicit preferences for Ukraine (greater than zero) or Russia (less than zero). 128

List of Tables

3.1 Results of models testing H1 71

3.A.1 Results of models testing H1 (OLS) 78

3.A.2 Rare Events Model Summary 79

4.5.1 Summary statistics of explicit and implicit preferences by ethnicity, with values greater than zero indicating pro-Ukraine bias. Means of the explicit measure are of the full range of possible values (i.e., integers in [-6, 6]), while the means of the implicit measure are calculated from the d-score units (i.e., real values in [-1.15, 1.64]). * indicates mean is distinct from 0 with 95% confidence. 105

4.B.1 Majorities of each of the four ethnic identity categories expressed an implicit and explicit preference for Ukraine (“Pro Ukraine”) or an implicit preference for Ukraine but an explicit preference for Russia (“Pro Russia disassociation”). Respondents who expressed a neutral explicit attitude (they evaluated Russia and Ukraine the same) are coded as dissociating. 118

4.B.2 Majorities in each of the four cities expressed an implicit and explicit preference for Ukraine (“Pro Ukraine”) or an implicit preference for Ukraine but an explicit preference for Russia (“Pro Russia disassociation”). Respondents who expressed a neutral explicit attitude (they evaluated Russia and Ukraine the same) are coded as dissociating. 118

4.B.3	Summary statistics of explicit and implicit preferences by city, with values greater than zero indicating pro-Ukraine bias.	119
4.B.4	Home language use by city.	119
4.B.5	Home language use by reported ethnic identity.	119
4.D.1	Sets of words that IAT participants were required to associated with good/bad and the respective categories. P are political words, and C are cultural words. We found no substantive differences between analyses that used all words, just political words, or just cultural words.	121
4.E.1	124
4.F.1	126

Acknowledgements

This dissertation benefited from many people's perspective, advice, guidance, encouragement and support. First among them is James Long, who chaired my dissertation committee. James saw potential in my ideas, pushed me to refine my thinking about this project, and held high standards. Additionally, James helped me navigate the PhD program, chaired my MA committee, and helped me deal with the world of academic publishing (particularly dealing with reviewers). My work at the University of Washington and beyond has been made far better for his involvement in my academic training.

Victor Menaldo and Scott Radnitz rounded out my committee, and both have made critical contributions to my work. Victor's passion and ambition as a scholar inspired me to think bigger than I otherwise would have and helped me develop the theoretical core of my ideas. Scott was generous with his time on numerous projects I worked on, offering invaluable feedback as a political scientist and area expert with deep knowledge in the post-Soviet region. His work running the Ellison Center for Russian, East European and Central Asian Studies provided a second intellectual home for me outside of the Political Science Department.

I was lucky to be at the Department of Political Science at UW because of the people that were there. Across the board the faculty were engaged, interested, interesting, and supportive, and Geoff Wallace, Chris Adolph, and Caitlin Ainsley merit special mention. They each had an outsize impact on my time at UW. The excellent departmental staff, including Ann Buscherfeld, Susanne Recordon, Meera Roy, Ellen Gregoire, and Steve Dunne

were knowledgeable, supportive, and friendly.

I also benefited from the support of the English Department, where I worked for two years as a writing instructor. Mike Gabbay's research team at the Applied Physics Lab was also an irreplaceable part of my education. Mike, a physicist by training, brought together an amazing group of interdisciplinary researchers from Political Science, Mathematics, and the Information School to study important questions about insurgent violence. Working alongside smart people who approach problems in a fundamentally different way was invaluable career training.

As with faculty, there are too many graduate students who shaped my experience, stretched my thinking, and shared the challenges for me to mention them all, so I'll just mention three. Aaron Erlich was a TA who turned into a collaborator and mentor, giving me the opportunity to co-author with him on several occasions. Will Gochberg was a fantastic office mate who was a valuable sounding board and fellow coffee-enthusiast. We went through many of the departmental rites of passage together. Nora Webb Williams connected me to the opportunity to work at the Applied Physics Laboratory and was a wonderful colleague throughout my time in graduate school.

Finally, during a long graduate school career that has had many ups and downs, my friends and family have been a consistent source of love, comfort, and support.

Dedication

To my parents, for their encouragement; to my wife, for her support; and to my kids, for helping me keep it all in perspective.

Chapter 1

Introduction

This dissertation engages with two questions from the study of authoritarianism and international conflict. First, it identifies an unanswered question in the formal models and political economy research into authoritarian regimes. Chapter 2 starts with an observation: most of the authoritarian regimes in the world today are multi-party authoritarian regimes, and their numbers have been growing relative to one-party states, not shrinking. These multiparty authoritarian regimes allow opposition parties to exist, compete in elections, and enter the legislature and other political offices, so long as they do not pose a meaningful threat to regime survival. These parties are referred to as “systemic” opposition parties because they are often sustained by and work within the confines of the existing political system. In many such regimes, there is a dominant party through which the bulk of rents flow, and which has the greatest capacity to influence policy. This leads to a set of related puzzles: how does the presence of multiple parties benefit the authoritarian leader? Why would someone in an authoritarian regime join a systemic opposition party instead of the dominant party? Why vote for parties that are unlikely to take power?

Formal models of electoral institutions in authoritarian regimes have a lot to say about how parties, elections, and legislatures benefit dictators. Legislatures predict longer leader tenure (Geddes, [1999](#)), the threat of elite defection by an organized opposition can prevent

predation by the autocrat (Magaloni, 2008), and even rigged elections can signal regime strength and deter challenges (Simpser, 2013). These accounts offer little guidance here, however, because they assume a single elite structure, often one that is defined solely by its desire for rents. They don't specify why multiple parties would be better or worse than a single party. Gandhi (2008) argues that institutional choice—such as allowing for multiple parties—may be a function of the strength and organization of the elite, and Lust-Okar (2005) argues that the elite selects the arenas in which it will have to confront the opposition, thereby limiting the threat it faces. While these claims are somewhat relevant to the question at hand, they don't clearly specify the reasons why one dictator would allow multiple parties while another will not, except to refer in a general way to the strength of the regime.¹ There is even less theoretical or empirical focus on why some elites join or remain in systemic opposition parties in authoritarian regimes. The formal models referenced above tend to distill the choices faced by the elite as supporting the regime or opposing it, with opposition defined narrowly as open revolt. They do not consider the middle option of joining the systemic opposition, despite the prevalence of this behavior in reality.

I innovate on the existing theoretical literature in two ways. First, I present a new analogy, likening an authoritarian regime to a discriminating monopsonist. A monopsonist is the same as a monopolist, except a monopolist is the sole seller of a good or service, while a monopsonist is the sole buyer of a good or service.² The dictator needs to buy the labor services of some portion of the population, services that I refer to as political engagement. These labor services cover everything from full time employment (in the case of political elites serving in legislatures and other government roles), to the infrequent act of going to the polls and voting. Because the dictator does not allow any rivals to emerge as competitors, the dictator is a monopsonist with respect to political engagement. Dictators will set the “wage”

¹In fairness, we shouldn't expect these theories to answer the question I pose because it is not their central goal.

²This analogy may seem odd to those who are accustomed to thinking of the state as a supplier of services, such as law and order, roads and bridges, etc. I justify this choice in Chapter 2, but for those who remain unconvinced, the implications of the model hold if we think of the regime as a monopolist.

for political engagement at the point where the marginal value of political engagement to the dictator is equal to the cost of purchase.

My second innovation is that a dictator not just a monopsonist, but a discriminating monopsonist. I model political actors as being interested in maximizing over both material and expressive benefits. Material benefits are the normal stuff of political economy models, namely rents and the expectation of rents in the future. These are the wages that political elites are compensated with for organizing and working on behalf of the regime. Expressive benefits are also a well-developed theoretical concept (Hamlin & Jennings, 2011), albeit one that is newer to the models of authoritarian politics. Expressive benefits generate utility by allowing individuals to take actions that are consistent with their beliefs, attitudes, and values. By allowing political actors to express their views, associate with parties that reflect their values, and vote for and speak on behalf of issues they care about, the dictator gets them to reveal their ideological preferences. Multi-party authoritarianism helps the dictator discriminate between political actors and, ultimately, pay out less rents overall than if all political actors were required to support the same party, working on behalf of the same values and ideas. This model predicts the conditions under which political elites would associate with the systemic opposition instead of the dominant regime party, as well which members of the systemic opposition are most likely to defect to dominant, pro-regime parties.

The third chapter examines my model's predictions about elite behavior in a specific case study: authoritarian consolidation in Russia under Vladimir Putin. I look at how the creation of a pro-regime political party, A Just Russia (AJR), ahead of the 2007 Duma elections created differing incentives for members of the systemic opposition. The new party represented an opportunity to get in on the ground floor of a party that was designed to be a second party of power. As a result, members of the systemic opposition could reasonably conclude that their expected rents would increase by joining this new pro-regime party. However, my model predicts that Duma deputies would be more likely to join AJR if they also stand to gain in terms of the expressive benefit of joining AJR. It expects that those who

are closest to the ideological position of AJR, and farthest from the ideological position of their own parties, are most likely to join AJR. To test this prediction, I collected deputy-level vote data from the 4th Session of the Russian State Duma (2003-2007) and I find preliminary evidence in support of the model's predictions, albeit with small effect sizes. The interaction of ideological proximity to AJR and distance from one's own party translates into greater probability of joining the new party for deputies in the systemic opposition.

Chapter 4, which was co-authored with Aaron Erlich and published in *International Studies Quarterly*, shifts the focus from elite political actors and party politics to how individuals think about themselves and the state when war breaks out. We explore the attitudes of Ukrainian citizens towards Ukraine and Russia after the 2014 Russian invasion of Ukraine and the start of the war between Ukraine, Russia, and Russian-directed proxies. This chapter engages with literature on the “rally ’round the flag” effect (Davies, 2002; Mueller, 1973; Russett, 1990), which expects citizens to have a surge of support for their home state when war breaks out. An open question about this theory, and the evidence that supports it, is the extent to which all citizens rally to their state's side. Indeed, there are numerous examples throughout history of states actively persecuting their own populations when at war due to suspicions that parts of the population—almost always ethnic or racial minorities—will support the adversary. One example is the internment of American citizens with Japanese ancestry during World War II. Scholars have noted that concerns over so-called Fifth Columnists have included ethno-linguistic Russian minority populations that were “beached” outside of Russia following the breakup of the Soviet Union, including in Ukraine (Radnitz & Mylonas, 2022). We show that both ethnic Ukrainians and ethnic Russians favor Ukraine over Russia using implicit association tests. Although respondents could give misleading or untruthful answer on traditional survey questions about these issues, implicit association tests deliver a more accurate view. Additionally, in contrast to expectations that ethnic minorities would feel pressure to mask their loyalties to an ethnic homeland in times of war, we find that Ukrainian citizens that are ethno-linguistically Russian may be over-stating

their positive sentiments towards Russia, based on a comparison of survey responses and IAT results. This chapter presents a counterpoint to a commonly held belief that minority populations can't be trusted during a time of war, a belief that has caused and continues to cause serious harm around the world.

In Chapter 5, I conclude by outlining the findings and significant contributions of each chapter. These include a new theoretical framework to understand authoritarian politics, new data and analysis of the systemic opposition in Russia under Vladimir Putin, and a methodological innovation in the study of public opinion and how the public rallies around the state in wartime. It also details shortcomings and unanswered questions in each of these chapters, along with potential paths for extension and future work.

Chapter 2

Chapter 2

2.1 Introduction

The third wave of democratization saw a drop in the number of single-party authoritarian regimes, as democratic forms of government gained ground around the world. Looking back, the data also show that this period saw the rise of democratic authoritarianism, the incorporation of democratic institutions by authoritarian regimes. Indeed, for the last thirty years the most common form of authoritarianism, by far, has not been a single-party state, or a military dictatorship, but multi-party authoritarian regime with elections and legislatures. This pattern is shown in Figure 2.1, which analyzes data provided by Hadenius and Teorell (2007) and Wahman et al. (2013). Additionally, there has been a steady decline in the share of legislative seats held by the largest party in both military and multi-party authoritarian regimes. The fact that multi-party authoritarianism has been the dominant mode of authoritarian rule for so long suggests that this specific institutional arrangement is not an aberration, but an equilibrium outcome that is, as far as governments go, stable.

I argue that, despite their ubiquity, the presence of multiple political parties in authoritarian regimes presents something of a puzzle. It is the presence of systemic opposition parties in regimes with a dominant party that is most puzzling. Systemic opposition par-

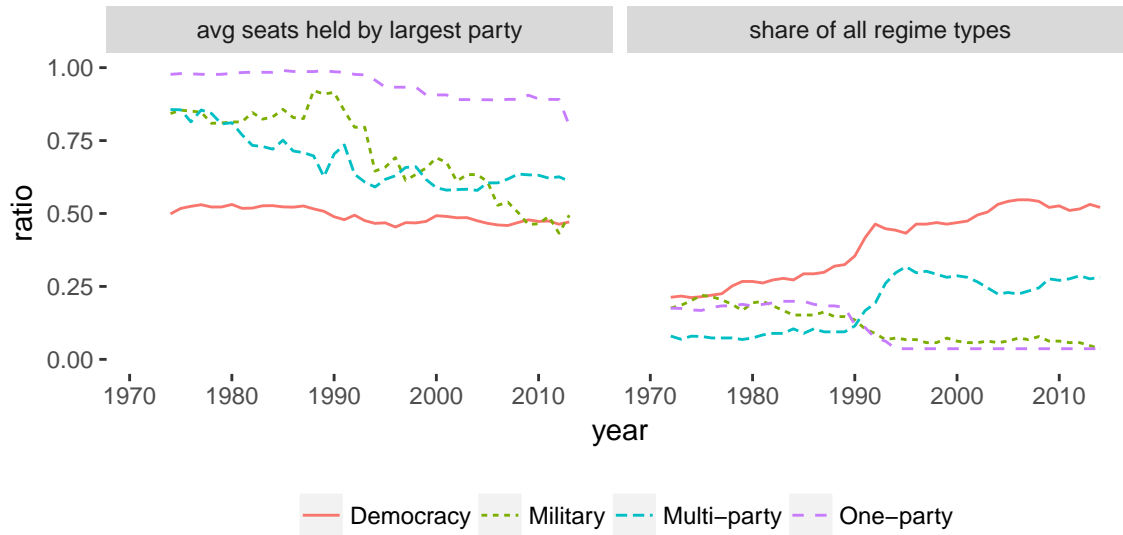


Figure 2.1: Changes in the ratio of democratic, single-party, and multi-party authoritarian regimes

ties are permitted to exist, compete, and win seats, but they do not use the full extent of their formal institutional power to either criticize the authoritarian leader, challenge the authoritarian nature of the system, or maximize their own power. These are parties that are variously characterized as a “loyal” opposition, a “toothless” opposition, or members of “rubber stamp” legislatures.

The puzzle becomes clear when we focus on the separate motivations and considerations of the authoritarian leader, the political elite, and voters. Why does the authoritarian leader in these regimes permit such parties to exist? Organized opposition parties that campaign in the open, cultivate bases of public support, and win seats seem in legislatures must dramatically increase the threats to the regime. Multiple parties increase monitoring costs for the authoritarian leader and reduce collective action problems for members of the elite who might want to challenge the regime compared to a situation in which only one party (or no parties) are permitted. From the point of view of individual members of the political elite, why do some join or support pro-regime parties while others opt for systemic opposition parties? Joining systemic opposition parties foregoes the higher rents and influence associated with joining the dominant political party. Similarly, for voters,

why vote for candidates in systemic opposition parties that are not able to influence policy in a meaningful way? Association with opposition parties is especially puzzling in regimes that have a dominant party, one that is clearly the most efficient avenue through which individuals—both mass and elite—could potentially extract rents from the state. These questions, which address both the demand and supply side for systemic opposition parties, represent an important gap in our understanding of authoritarian politics.

This chapter aims to provide a clear theoretical explanation for the persistence of multiple parties in authoritarian regimes. I argue that formally incorporating ideological preferences into models of political behavior in authoritarian regimes can explain this behavior. Ideology is difficult to pin down, but here I mean the set of views and beliefs that describe the proper way to live in and structure society. I draw on Denzau and North (2000)¹ in particular, who define ideologies as: “the shared framework of mental models that groups of individuals possess that provide both an interpretation of the environment and a prescription as to how that environment should be structured.” By acting in accordance with their ideological preferences, individuals exhibit expressive behavior, in which benefit comes from the “meaning or symbolic significance of actions or choices themselves, rather than their indirect consequences or consumption benefits” (Hamlin & Jennings, 2011, p. 655). My model includes both material returns and an individual’s “expressive benefit” for taking actions and expressing views that are consistent with their ideological preferences. This model shows the trade-off that some individuals will face between taking actions that increase their expected rents from the state and actions that comport with their ideological preferences.

The theory shows that multiple political parties help leaders of authoritarian regimes identify the reservation values and relative preferences for participation among the elite and the population. In so doing, these institutions ensure that authoritarian leaders can maximize their own rents while minimizing the potential for organized opposition to their regimes. Ultimately, the theory I outline can show the conditions under which systemic

¹As referenced in Jost et al. (2009, p. 309).

opposition parties can be an equilibrium outcome. I focus on the expressive benefit that some members of the elite will derive from joining parties in the systemic opposition, but the basic insight is extensible to leaders and voters, as well as those who choose to join true opposition organizations. The model provides a way of reasoning about the types of individuals who will take different types of actions in authoritarian societies. Ultimately, it predicts that leaders will take ideology into consideration when designing institutions and mobilizing popular support; that elites will consider ideology when making their decisions of whether to join the regime or the opposition, and that voters will include an expressive benefit based on ideology in their vote choice.

The rest of this chapter outlines the existing formal theories of democratic institutions in authoritarian regimes. It does this to acknowledge the substantial work that has already been done, as well as to clarify how my contribution fills in gaps in that existing work. I refer to the case of Russia under Vladimir Putin to illustrate with a concrete example how existing theories fall short. I then outline the theory and model and show how the model can be used to generate testable hypotheses. Finally, I end with how the theory contributes to broader questions in political science.

2.2 Literature Review

There is no shortage of theorizing about how democratic institutions can serve anti-democratic ends. An impressive literature has developed over the past 25 years considering why authoritarian regimes adopt democratic institutions such as elections, political parties, and legislatures.² This prior research has identified a number of channels by which adopting democratic institutions can benefit an authoritarian leader. These explanations are, in some cases, quite persuasive, but they do not engage with why *multiple* parties would be more in the regime's interest than a single party. There are also competing accounts about how and whether

²See Brancati (2014) and Gandhi and Lust-Okar (2009) for an extensive review of the institutionalist literature on electoral authoritarianism.

members of the political elite should be expected to join the opposition. Additionally, to the extent that it is addressed, the concept of what “the opposition” means in these theories does not clearly match the empirical reality of opposition parties in authoritarian regimes. This section outlines prior explanations for the institutional equilibria that characterize electoral authoritarianism. After doing so, it takes each explanation in turn and highlights the theoretical weaknesses that persist, adding empirical examples from the Russian case.

2.2.1 Gaps in existing theory

There are many accounts of why democratic institutions would benefit authoritarian leaders. Magaloni (2008) and Magaloni (2006) argue that a political party can help authoritarian leaders resolve commitment problems, manage intra-elite competition, and distribute spoils. Gandhi (2008) claims that the leader’s choice of electoral and legislative institutions will be a function of the leader’s strength vis-à-vis the opposition, and that those institutions will help the leader manage threats posed by the elite. Simpser (2013) argues that elections, even (or especially) those that are rigged, can signal regime strength and deter challengers. Additionally, a recent review by Waldner and Lust (2018) catalogs theoretical frameworks that explain democratic backsliding through institutions, culture, social cleavages, and elite agency.³

Magaloni (2008) acknowledges the potential for a pro-regime elite to defect to the opposition as a threat that helps to prevent predation by the autocrat, but there is no formal model of this concept or empirical test. It presumes that the a member of the pro-regime elite can threaten to defect unilaterally without punishment, but in regimes with weak property rights and selective prosecution of political opponents, such as Russia, such a defection threat may not be credible. Additionally, this account of the threat of defection is in conflict with other accounts of authoritarian politics. The selectorate model, for example, argues that the goal of political actors is to enter the winning coalition, so the threat of an individual defection

³That said, the democratic backsliding literature focuses more on the causes of these changes specifically within democracies, and as a result is less relevant here.

is not terribly threatening to the dictator (Bueno de Mesquita et al., 2003). The defector can be easily replaced, especially when the payoffs for being in the winning coalition are relatively high, as they are in authoritarian regimes. Others argue that it is informal, not formal power that is key to challenging incumbents in authoritarian regimes. Indeed, elites know better than to invest in building opposition parties, Radnitz (2010, p. 6) writes in his analysis of elite politics in authoritarian Central Asia, because authoritarian leaders, “are prone to subvert or ignore formal state institutions.”

On the other hand, if a member of the pro-regime elite can overcome a collective action problem and convince sizable members to join in the defection, then why not seize power? Taken in this sense, providing an institutional vehicle for collective action and defection from the regime—a standing, legal opposition party—seems to be more trouble than it is worth. If, however, these authors are right, and these institutions enhance the security and longevity of the regime, then the puzzle becomes all the more puzzling; why would anyone join the opposition if it is so patently obvious that the regime is strong? The prior work in this area most often focuses on the advantages that an incumbent regime enjoys when it creates a legislature or a political party; it does not address why additional (opposition) parties enhance the regime’s position, nor does it incorporate the latent threat that these opposition organizations pose to the regime.

The existing theoretical and empirical work on authoritarian institutions is vast and impressive. As such, it makes sense to state early on why these problems might need to be re-theorized. First, we should not assume that models that were built to explain why authoritarian regimes incorporate or adopt electoral institutions will necessarily explain why autocratizing regimes will keep those institutions present. Second, I make different underlying assumptions about the motivations of the leader, elites, and citizens than do many of the previous theories of authoritarian politics, shifting the emphasis away from categories of classification (personalist, single party-based, military junta, etc.) and opting instead to focus on the micro-foundations that guide individual action. These different assumptions

lead to a distinct model, one that argues people are motivated by more than just rents.

Democratic institutions, rather than leading to democratization, can increase the durability of authoritarian regimes in several ways: large electoral victory margins can signal regime strength and deter opposition; legislatures can distribute patronage and resolve commitment problems between leaders and regime-loyal elites; local and regional elections can monitor performance of lower-level elites; and acquiring information about societal grievances and demands so as to avoid direct challenges Brancati (2014).

Legislatures and political parties are two democratic institutions that seem to play a very important role in stabilizing authoritarian regimes and ensuring their longevity relative to authoritarian regimes without these institutions (Geddes, 1999), and they tend to predict better outcomes for autocratic leaders once they leave power (Geddes et al., 2014). But we know less from both a theoretical and empirical standpoint about how multiparty authoritarianism differs from single-party authoritarianism.

Lust-Okar (2005) has provided a detailed consideration of how opposition elite strategies in confronting the regime are conditioned by the “structures of contestation” that the regime designs. This work illuminates the empirical reality that not all oppositions are equal. This fact reinforces the notion that incumbents in authoritarian regimes have options as to how to structure the ways in which societal forces outside of the state can push for access to the state’s resources. That said, the argument takes as given that there is an opposition that can be divided, included, or excluded at the regime’s discretion. This account focuses on what happens once the composition of the opposition is relatively static, and it does not interrogate what it means to be opposition, or why anyone would choose to join the opposition in the first place. It does not allow for the identity of the regime or the opposition to shift.

In terms of authoritarian equilibria, Svoblik (2009) presents a model that explains how two different outcomes—established dictatorship or contested dictatorship—emerge, but again, this model does not speak to how those in the opposition get there, or what strategic calculus makes someone at the margin break for the regime instead of the opposition. Further, as the

model tends towards established dictatorship, it seems to predict that leaders of authoritarian regimes will be inclined to, eventually, marginalize opposition parties and then shutter them altogether.

The selectorate model (Bueno de Mesquita et al., 2003) argues that the concentration of power and “affinity” of political elites will help predict coalition formation, but it explicitly avoids a discussion of ideology, suggesting that all individuals would prefer to be in the winning coalition if possible. Indeed, the threat of exclusion from a future winning coalition is a key component of loyalty to the current regime (Bueno de Mesquita et al., 2003, p. 66). The model predicts that repression is more likely in autocratic regimes, but it is silent on the conditions that prompt these regimes to use repression. Conditional on not being repressed, when are those excluded from the winning coalition allowed to form political parties (that are not allowed to win control of government)? Additionally, from the point of view of an excluded member of the selectorate, signing up with an opposition party is a puzzling action to take, as it signals some absence of affinity between that individual and the leader.

My puzzle—why some political elites affiliate with the systemic opposition—and theory are in some ways similar to those addressed by Gandhi (2008) and Magaloni (2006). Magaloni incorporates ideology into her explanation for why regime-loyal elites decide to stay with the regime or split. As for a theory of why multiple parties would be in the regime’s interest, Gandhi (2008) argues that dictators allow opposition parties to emerge because those parties facilitate orderly and non-threatening deal-making with the political opposition. The decision model of political action that I present in section 2.3 is quite similar to arguments from both authors, but it also differs in important ways. First, I argue that the conceptualization of ideology, I , in Magaloni (2006, Ch. 1) as incorporating both ideological preference and policy outcome is overly broad. Second, the focus on actors who care about rents or policy concessions in Gandhi (2008, Ch. 3) is overly narrow and, I would suggest, makes a distinction without a difference. What is a tax break, or a tariff, or access to subsidized loans? I argue that what we call policy concessions are often just rents with some assembly

required. When policies are designed to direct revenue streams to certain groups, then they are for the purposes of my model rents.

Also, institutions, such as elections, legislatures, and parties, are fora in which concessions are granted, but might they not also be concessions in their own right? If individuals care about process, then the means by which ends are arrived at may be important. However, Gandhi's model assumes that opposition groups will not demand these structures in and of themselves, but rather they will make demands in those institutions; dictators will create institutions that deal efficiently with those demands if a dictator is sufficiently weak.

Additionally, Gandhi's model explicitly assumes that the dictator does not share the preferences of the median voter on a given policy space. While this may be the case in some contexts, the fact that many dictators enjoy high approval ratings suggests that the model is unnecessarily constrained. Indeed, in the Russian case, most would concede that Putin has enjoyed the popular support of more than the majority of the population for most of his time as the country's leader; it is likely that he could win freer and fairer elections⁴, but he has chosen not to leave it to chance on this point.

Moreover, Gandhi's formal model presents a dictator and a unified opposition. This model seems to imply that the important question in authoritarian politics is that of the dictator against all, when often it is a question of the dictator and a large coalition against the rest. Again, the recent history of Russia is illustrative in this regard. As Putin was investing in United Russia as a national 'party of power', political elites signed up for the party in droves (see Reuter (2017)). This shifting in allegiances suggests that membership in the opposition is not a permanent condition. Political elites strategically readjust in real time, but most prior work black-boxes the process by which a member of the elite becomes either pro-regime or a member of the 'opposition'.

Finally, and perhaps most problematically, preceding models have not adequately addressed the fact that allowing groups to exist in an organized manner—even in a “sham”

⁴See Frye et al. (2017), H. E. Hale (2017, p. 275), and Treisman (2011, 2014).

legislature or a “loyal opposition” political party—gives that group additional resources, reduces its coordination costs, increases its societal legitimacy, and facilitates collective action to a far greater extent than when the group is forbidden from existing openly. The effect of the ability to openly organize and meet on the distribution of power between leader and opposition must be incorporated into any model that seeks to explain why a leader would allow such organization to occur or persist. This would seem to be doubly important for a theory that argues that weak leaders are those that are most likely to empower the opposition by granting institutions in which they can conspire against the leader.

2.2.2 Empirical problems: systemic opposition in Russia

Previous work on authoritarian political institutions has focused on the adoption of nominally democratic institutions by authoritarian leaders. They have shown the various ways in which these institutions, once in place, can be manipulated by the leader to shore up the regime. However, there are several ways in which this prior work does a poor job explaining some aspects of Russian politics in the Putin era (1999-present), particularly the persistence of systemic opposition parties. First, the opposition does not appear to have a meaningful role to play in the policy making process in either government or the legislature. Second, it appears that affiliation with United Russia, the party of power, is strongly associated with financial success out of politics. Given these two facts, why would anyone remain in the systemic opposition? If opposition politicians are motivated by rents and concessions, and if these flow through affiliation with United Russia, then the rational move would be to join United Russia, as many Russian politicians did in droves; but many others did not.

Some explanations for the persistence of democratic institutions within Putin’s Russia highlight international norms that require the forms of democratic governance, the threat of pressure of international donor groups and agencies if Russia were to eliminate all opposition, or Putin’s own preferences for playing by the rules. But these explanations are unconvincing. Part of Putin’s political power has come from a deliberate flouting of western political norms.

Further, Russia has paid no meaningful penalty for curtailing freedoms of political opposition or the media, so it is hard to understand why fear of such a penalty would now limit its actions. Indeed, quite the opposite; the Putin regime has been historically very willing to throw out of the country the very international NGOs who might presumably prevent a full reversion to authoritarianism by withhold aid dollars. Finally, attributing actions to Putin's preferences for playing by the rules doesn't really explain much; it just shifts the phenomenon under study into Putin's head and limits our capacity to compare across other cases.

Russia under Putin is not a case of an authoritarian regime choosing to introduce democratic institutions to deal with emerging challenges or shore up the regime in some other way. Rather, it is a case of democratic backsliding or authoritarian re-consolidation. Under Putin, Russia saw a series of democratic institutions weakened or curtailed, including (but not limited to) the elimination of direct election for governors and some mayors; manipulation of electoral laws and registration procedures to keep some political parties off the ballot; and selective and targeted prosecution to make candidates ineligible for running for office⁵. Thus, Russia is not a case of adopting these institutions for regime survival; it is one in which some democratic institutions are left standing. Existing theory does not help us understand which such institutions we should expect to be left in place and which should be eliminated. An examination of the logic of these theories illustrates their shortcomings in explaining the logic of the rollback of democratic institutions in Russia.

The puzzle of the leader

The theory laid out above is motivated by a set of puzzles that I argue are unavoidable in the existing literature on electoral authoritarianism. Russian politics under Putin lays each of these puzzles bare. At the level of the leader, why would Putin allow an organized opposition to persist? There is little in his track record that suggests he has a particular interest in generating the kind of vigorous policy debate that an organized political opposition

⁵This quick summary says nothing of the pressure put on the judiciary, the media, and civic organizations that could potentially pose a challenge to the stability of Putin's regime.

can supply. Furthermore, there are historical and contemporary examples of how organized opposition politicians can become unexpectedly powerful, rapidly gathering enough strength to challenge not just incumbent leaders but the very foundation of the regime. In the waning days of the Soviet Union, Boris Yeltsin was able to use his position as an opposition leader and the president of the Russian Republic to tear down the USSR from within. Just years later, the Communist Party of the Russian Federation emerged resurgent and posed a serious challenge to Yeltsin's grip on power. In the early years of Putin's time in power, so-called Color Revolutions toppled incumbents in Georgia (2003), Ukraine (2004), and Kyrgyzstan (2005). In each case, disputed elections were the flash point for unrest and citizen protest that led to the incumbents losing power. While in the case of Kyrgyzstan informal networks, rather than political party structures, proved decisive in toppling the incumbent (Radnitz, 2010), a disputed multi-party parliamentary election was the catalyst for the unrest. Clearly, there was reason for Putin to believe that an organized political opposition posed risks to his long-term hold on power.

If there was a motivation to remove organized opposition, perhaps Putin did not have the means to do so? Again, the historical record suggests that this is not necessarily the case. Putin was able to alter electoral laws and regulations multiple times, in each instance reducing the authority of groups and individuals who could challenge his rule. Three examples illustrate this power, and all come from early in his time in office, before he amassed the authority that he has today. First, there was the elimination of direct election of governors in 2004, which made the regional elites more dependent on him. Second, there was the shift from a mixed single-member district and proportional representation electoral system to a national-list proportional representation system, further hobbling the regional elites that had been so troubling to Yeltsin. Third, he pushed through new policies that raised the threshold to register as a party in Duma elections, effectively preventing Yabloko and other opposition parties from challenging his hold on power. Putin's time in power has been marked by changes in formal institutions that increase his authority; there was clearly the

means to consolidate. Further complicating things, Putin actually introduced new systemic opposition parties as he was consolidating his power.

The puzzle of elites

If Putin's behavior is puzzling because he is allowing potential sources of opposition to remain unified, then elite behavior is puzzling because some elites continue to participate in or join what appears to be a toothless opposition. Reuter (2017, p. 28) writes that, "United Russia membership provides elites with a number of significant benefits, including access to rents, policy influence, and career advancement opportunities. From 2003 onward, almost all legislative bargaining has run through United Russia legislative factions.... In elections, United Russia membership provides access to a valuable party brand, ballot access, and state administrative support." He reports that by 2008 94% of regional governors were UR members, and by 2011 the party had majorities in 82 out of 83 regional legislatures and 92% of city councils in large cities (above 75,000 inhabitants). Moreover, Krol (2017) shows that Duma deputies aligned with United Russia exact far more in policy concessions than members of the systemic opposition.⁶ Frye et al. (2014) have shown how economic elites mobilize voters for United Russia (24% of firms report doing so, and 25% of voters claim to feel such pressure), but that means that a substantial proportion of the elite do not; what keeps them on the sidelines or, alternatively, mobilizing for the systemic opposition?

These benefits are so attractive that it prompts a question: why would anyone run for office as a member of an opposition party? To date, however, there is relatively little empirical work that explores these various potential explanations for why potential candidates would elect to join any part but the dominant one. Yet, we see large, stable 'systemic opposition' parties that have proliferated in Russia and in other authoritarian or hybrid regimes.

⁶That said, Szakonyi (2017) has shown that businesspeople who win elected office are able to deliver greater profits for their firms, an effect that is present for members of opposition parties as well as those aligned with United Russia.

The puzzle of the voter

The final and interrelated piece of the puzzle is found when one looks at voter behavior. Cox (1997) argues that voters understand whether their preferred candidates are likely to win, and engage in strategic voting when their candidates are likely to lose. In this light, sustained voter support for systemic opposition parties is a puzzle because the nature of the system implies that these parties are extremely unlikely to seize power through electoral means. Admittedly, there are cases where voters seem to overestimate the probability of their preferred candidates winning. Drawing on political psychology, Horowitz and Long (2016) show how voters in democracies where ethnicity is highly salient in politics can discount readily available information showing that their preferred candidates will lose. This explanation, however, does not explain the Russian case because Russian politics, and in particular, Russian systemic opposition parties, are not organized around ethnic lines, nor do they exhibit the type of geographic concentration that Horowitz and Long (2016) identify as important to sustaining misperceptions about candidate viability.

Another possible explanation is that opposition elites are simply buying the loyalty of some percentage of the voting public. However, this, too, is not a sufficient explanation. The option of “against all” was on the Russian ballot for president and Duma (both in single-member districts and proportional representation list voting) from 1993 until 2006, when it was abolished by Putin. The share of the vote going to “against all” climbed in each election from 1995 to 2004 in both parliamentary and presidential contests. In single-member districts in 2003, 12.9% of voters cast their ballots “against all” (McAllister & White, 2008). Even if it won the most votes, “against all” could not take office, so why vote for it? If votes are just bought and sold, then who would buy an “against all” vote?⁷ The willingness of so many voters to forgo potential pay-offs suggests that voters value not just the material returns to their voting. These votes are difficult to reconcile with analysis of Russian politics

⁷It is worth mentioning that even under the single-party Soviet Regime, political scientists were attempting to discern traces of dissent in voter behavior or non-behavior (Gilison, 1968; Karklins, 1986).

that turns on either a purely instrumentally rational voter who pursues policy objectives or one that assumes voters are essentially paid off or coerced to vote a certain way (See, for example, Frye et al. (2018)). This pattern of voting, and the fact that it was accelerating under Putin, and that Putin abolished this option presumably to prevent it from growing to be an embarrassingly large share of the vote, demands a model that incorporates voters' expressive benefit over voting.

There is evidence of expressive voting even in the early days of the construction of United Russia as a party, as it was able to articulate an ideology that voters found attractive. Indeed, as early as the legislative and presidential elections of 2007-08, there was emerging a loyal United Russia voter who identified with the technocratic, pro-market, and relatively pro-Western positions that the party espoused (Colton & Hale, 2009). Further, a "defection cascade" from United Russia in 2011-2012 seemed to be led by those who were less socially vulnerable segments of the population, those for whom the potential clientelistic payoff of voting for the dominant party was relatively meager (H. E. Hale & Colton, 2017). This suggests that something other than the rents offered by the regime predicted voter behavior. What is less clear is the degree to which systemic opposition parties were able to identify and mobilize core voters based on ideological appeals.

2.3 A theory of rents and ideas

This section presents formal models that form the backbone of my theoretical argument; these models show that by incorporating both ideology and material interests we can derive hypotheses that explain the anomalies that persist in the existing explanations. To address the theoretical shortcomings and empirical puzzles identified above, and in an attempt to build a model that better describes the empirical reality in Russia under Putin, my model diverges in the following ways from the framework laid out by Gandhi. First, I propose a model in which the leader offers a combination of material and ideational compensation to

all members of society, as opposed to the rents-concessions offer in Gandhi's model. I also isolate the strategic interaction surrounding institutional equilibrium from the haggling that may or may not take place inside that institutional setting. Second, I allow for the dictator to share, broadly speaking, the ideological position of the majority. Third, I examine the strategic decision of political elites to join the dictator's coalition, thereby getting a share of the rents that Gandhi assumes accrue to the dictator, or remain opposed. To clarify, these relaxations of the assumptions in Gandhi's model do not fundamentally challenge its conclusions. Rather, I hope to provide a more encompassing framework of which Gandhi's model is a special case that places constraints on leader's position relative to the majority.

My theory rests on a micro-foundation of individual preferences over two goods. All actors in my model—leaders, elites, and citizens in authoritarian settings—seek to maximize utility across their preferences for material and ideational goods. Material goods are best thought of as private goods in the classic economic definition in that they can be exhausted and others can be excluded from using them. I call material goods, when they are distributed by the state, rents. Ideational goods are best thought of as a collection of attitudes about the right and proper way to govern and organize society. I will use the terms rents and ideas to distinguish between the material and ideational goods that citizens, elites, and the leader will seek to maximize from their respective relationships to the state. Not all individuals care equally about ideational goods relative to rents, and I represent this individual-level difference with a weighting term ($\lambda \in [0, \infty]$) that alters the value of ideational goods to material goods.

Individuals will seek to maximize across rents and ideas such that they will take any increase in one that does not cost them any of the other. Additionally, they will trade a reduction in rents for a sufficiently large increase in ideas, and vice versa. Figure 2.2 shows the ideas-rents frontier. Individuals want to be on the dashed line, and they are indifferent in trading off between rents and ideas at some rate. The shape of this frontier depends on each individual's combination of preferences for more rents or ideas that more closely match their

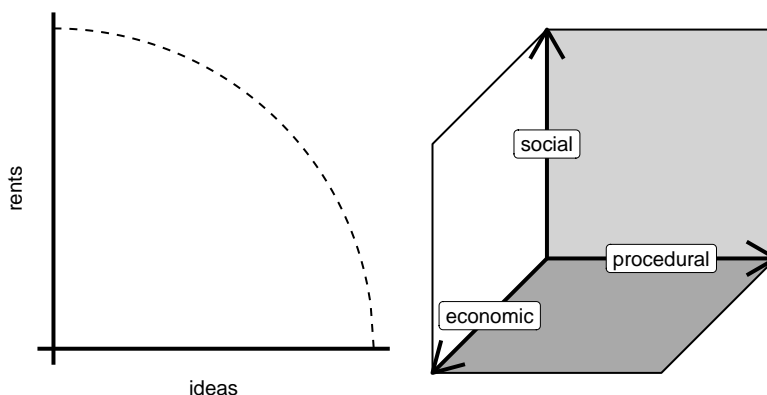


Figure 2.2: The ideas-rents preference frontier (left) and a three-dimensional ideas space (right)

own, which I model as λ , the preference for rents over ideational goods. An individual's ideal point is determined by their preference over economic issues, social issues, and procedural issues.

The concept of ideational goods is a new one to many of the formal models of authoritarian politics, but it is in many ways quite intuitive to grasp. Ideational goods are government policies and positions that do not directly affect the material well-being of citizens, elites, or the leader, or only do so in a manner that is exceedingly difficult to quantify. Rather, they are government actions, policies, and statements that are judged against individual standards of the right and proper way to live life and order society. In the context of the trade policy, for example, these are sociotropic policy preferences, i.e., those that are not perceived to directly affect the individual⁸. Individuals value ideational goods that are consistent with their own ideological preferences because of the expressive benefit they receive; this benefit should be thought of as one that can be added or subtracted from overall utility, just like the returns from material gains (Hamlin & Jennings, 2011, p. 655). Some examples include policies that

⁸See Mansfield and Mutz (2009) for an example showing how individuals appear to vote more based on sociotropic perceptions, i.e., how they perceive policies will affect the country, than how they perceive those policies will affect themselves individually.

are designed to: support religious liberty or require adherence to religious norms; ensure justice or economic equality; promote national honor and prestige; protect the environment; or advocate for human rights internationally. Many factors can affect the rents (r) associated with a given political action that has some ideational component ($idea$) to it, but generally speaking I argue that an individual (i) will choose the action ($a \in \{a, \dots, A\}$) that maximizes the utility function:

$$u_i(a) = r_a - \lambda_i(idea_a - idea_i)^2$$

As an example, I map preferences for ideational goods onto a three-dimensional idea-space, in which there are preferences for economic, social, and procedural values. Economic values may run from individualist to collectivist in orientation. Social values may run from conservative to liberal. Procedural values may run from indifferent to absolutist, and they refer to the degree to which a person feels the rule of law should govern society. As one becomes increasingly absolutist about procedural values, insistence on the rule of law grows accordingly. Thus, ideational goods can be effective courts and the rule of law, economic equality, or military spending (treated as a concessions to the military in Gandhi), provided such spending is deemed an important part of building national prestige or defending the nation from attack. Thus, rents for one group (the military) can be ideational goods for others (nationalists, militarists, revanchists).

2.3.1 Strategic objectives

I assume that the authoritarian leader wants to retain this position as it allows the greatest possible ability to maximize across ideational and material goods. To retain office, the leader needs people to perform tasks that ensure regime survival. The tasks that individuals perform, from working in the security services, to running for office, to attending rallies and voting, can be grouped under the notion of supporting the regime. Support takes many

shapes, all of which convey **active acquiescence** to the fact that the political regime as established is authoritarian and that the leader is preeminent in that system. Elites and citizens engage in active acquiescence by engaging in economic activity and engaging in political activity in approved ways. The leader provides incentives to elites to support the regime, and elites in turn provide incentives to citizens to act in ways that support the regime. Ideally, the leader wants to leverage the masses against the elites in order to ensure behavior in line with the leader's preferences. If possible, the leader will structure the institutions in such a way that it forces the elites to reveal their preferences and their reservation values for complying with the regime and the institutional configuration established by the leader.

I assume that elites are elite by virtue of their ability to mobilize individual members of society. This conceptualization is especially important in the Russian case, as networks of personal connection are critically important to all facets of life⁹. I allow for members of a given network to be heterogeneous in their preferences for rents over ideas, as well as for their location in the ideas space, but I assume that on average there is also heterogeneity between networks, so that these different networks can be characterized by different preferences for rents and different average values in the ideas space. Finally, I assume that the elite is better informed about these preferences than the dictator. While elites have preferences over ideas, these are in practice swamped by the fact that their position at the head of a network enables them the ability to reap very large rents in proportion to the size of the network and their ability to keep that network together. As such, they tend to adopt the point in ideas space that makes them most likely to keep that network together and grow it.

Individuals look to maximize their utility from rents and from the political actions they take, with the utility from political actions being a function of their position in the ideas space. Individuals maximize rents by staying in their network. However, there is an inherent tension between these rents and ideas for the individual; if the network moves to an ideas space that is too far from the individual's preferred point, then the rents associated with

⁹For the ways in which networks pervade modern Russian politics, see H. Hale (2003), H. E. Hale and Colton (2017), and Petrov (2011).

network membership may not be worth the expressive cost of engaging in political (or other) activity that is contrary to their own ideational preferences. Critically, individuals are not always able to discern whether policies implemented by the government directly help or hurt them; what is the effect of interest rates on employment? Not even all economists agree. However, individuals are able to tell, with an approximation, whether the ideas advocated for by a government are or are not in line with their own position in the ideas space.

2.3.2 A choice theoretic model of party affiliation

When choosing how to engage politically, most people operate as if they were selling their labor services in the open market. With respect to whether the regime is authoritarian or not, or what parties exist, an individual's choice is not influential enough to change the choice set for others. They do not have the resources to fundamentally alter the parties that exist. This makes the individual decision a choice-theoretic problem. I will illustrate this concept, and how it can apply to electoral politics, by providing a model of individuals considering whether to align with a political party and run for office.

When deciding which political party to associate with, individuals are best conceptualized as wage-takers because they must, in most instances, take potential party offerings as given. In reality, very few people have the capacity to start a new party from scratch. Out of n potential candidates, each one ($i \in \{1, \dots, n\}$) cares about the material and ideological benefits associated with running for elected office and maximizes utility over the choice of political organizations ($j \in \{1, \dots, J\}$). Specifically, they make a party choice that will maximize their utility over rents and ideology based on the various rents offered and ideological positions of the parties ($rents_j$ and $ideology_j$, respectively). Candidates enjoy diminishing marginal returns to rents and want to minimize the distance between the ideology of the party they choose and their own ideological preferences ($ideology_i$). The relative value of material and ideological benefits can vary from person to person according to a weighting parameter ($\lambda_i \in [0, \dots, \infty)$). Candidates assess the probability of winning elected office across

each party, and this term is a function of both the candidate and the party ($\pi_{i,j}$). Finally, regardless of winning or losing, each candidate bears a cost of running with a given party that varies by candidate (κ_i) and is related to the distance between the ideological preferences of the candidate and the party. Thus, each *candidate*_{*i*} $\in \{1, \dots, C\}$ chooses *party*_{*j*} $\in \{1, \dots, J\}$ where $C > 1$ and $J > 1$ such that the choice maximizes:

$$u_{\text{candidate}_i} = \pi_{i,j}[\lambda_i * \text{rents}_j - (\text{ideology}_j - \text{ideology}_i)^2] - \kappa_i(\text{ideology}_j - \text{ideology}_i)^2$$

The preceding analysis implies that the appropriate framework in which to analyze much of authoritarian politics is that of the market. The authoritarian leader plays the role of the monopsony buyer of political participation, i.e., active acquiescence of the population, and rents are the price paid to individuals for that acquiescence (Figure 2.3). Taken together, the willingness of individuals in society to acquiesce to the regime form an upward sloping supply curve. As a monopsony buyer chooses the quantity and price at which the market clears, there is no meaningful interpretation of a demand curve. Rather, it is instructive to think about a level of active acquiescence that a leader will seek to buy, q^* , which maximizes the simply utility function that a leader faces with respect to total participation of society in the market: $U_L = PV - C$, where PV is the protective value of buying off a given proportion of the population and C is the cost of buying off that proportion. The level of participation q^* is the level at which the marginal protective value (MPV) is equal to the marginal cost (MC). Conditional on the regime staying in power, I assume that there is a net positive value to the regime of people participating in and legitimizing the political process through voting *even if they vote for the opposition*¹⁰.

The monopsonist analogy may be surprising to many in light of the long tradition¹¹ of conceptualizing the state as a supplier of goods and services, such as roads, schools, the military, law and order, etc. I don't dispute that the state provides services to its popula-

¹⁰This argument is based on the argument by Simpser (2013) that rigged elections can still send beneficial messages.

¹¹see, for example, Hobbes (1651), Weber (1946), and Olson (1993), to name just a few.

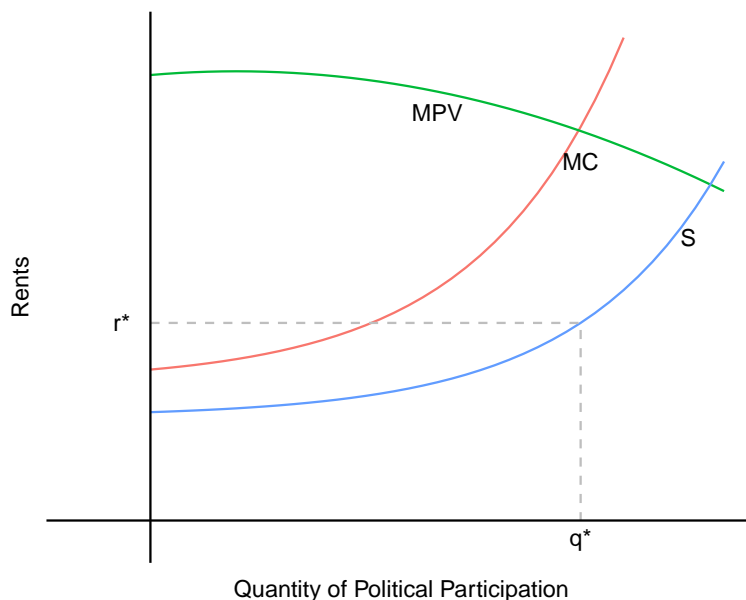


Figure 2.3: Equilibrium level of political participation and rents. the points q' and r' .

tion, and I don't see this role as inconsistent with the conceptualization of an authoritarian political regime as a consumer of services. Consider, for example, a factory in a closed town. The factory must secure a workforce to produce widgets. If it is the only employer, it is a monopsonist with respect to labor. It then must sell the widgets, and if there are no other widget manufacturers, it is a monopolist. The questions that I ask in Chapter 2 have to do with the choices that a dictator makes in securing the labor force (politicians) necessary to prop up the regime.¹²

2.3.3 A game theoretic model of party affiliation

Ultimately, the choice of many individual political actors, aggregated up, determines the equilibrium outcome of interest. I have presented a market in which the leader wants to purchase political engagement, but citizens and elites can always sell their labor services elsewhere. They can stay in the private sector. They may exit the market, either quietly,

¹²For those who are not convinced by the concept an authoritarian regime as purchasing labor, the math works the same if you frame the regime as a monopoly provider of a service that the elite are buying, although this framing is strange because the elite are not, in any meaningful sense buying the service of being a politician.

by staying home, or loudly by revolting and attempting to redefine the terms of the market. To better understand this relationship, I model the strategic interaction between a dictator (D) and a member of the political elite (E), first with and then without the presence of an opposition party. The leader wants to maximize rents retained but avoid being overthrown. In the absence of an opposition party, the elite utility function comprises both rents and ideology, as described above, but the choice set that they consider ($\{Accept, Reject\}$) does not offer them different expressive benefits; they are simply choosing between given level of rents as payment for associating with the single pro-regime party, on the one hand, and the expected utility of revolution, on the other. The game is a modified divide-the-dollar format, in which the leader makes an offer to the elite, the elite accepts or rejects the offer. If it rejects the offer, the elite revolts, the leader pays a cost κ of putting down the revolt, and nature (N) determines whether the revolt is successful with probability ϕ . Whoever is in power after the revolt gets $1 - \kappa$. Figure 2.4 illustrates the sequence of play.

In this setup, the leader will make an offer of r that is sufficient to keep the Elite from rejecting, an amount that is defined by the elite's expected utility from revolt conditional on the probability of success. In other words, the leader will offer $r \geq \phi(1 - \kappa)$ and the elite will accept. I define this level of r_p as the maintenance cost for the leader's dominant party. I assume that the only alternative to accepting the offer is rejecting and rebelling, although this is obviously a simplification. In reality, the alternative is usually to exit the political process. However, the portion of society that rejects the offer is clearly worrisome from the point of view of the dictator, because they are willing to express their discontent.¹³ The dictator must pay high costs for surveillance and to repress these individuals.

Now let the dictator propose an institutional arrangement in which elites can choose to affiliate with either a party of power or an opposition party. Both parties will receive some rents, but those that go to the dominant party will exceed those that go to the opposition ($r_p > r_o$). The party of power supports the leader in every instance, but the opposition is

¹³The Putin regime has been specific about these numbers in some elections, claiming at one point that their aim was to win 75% of the vote on 70% voter turnout.

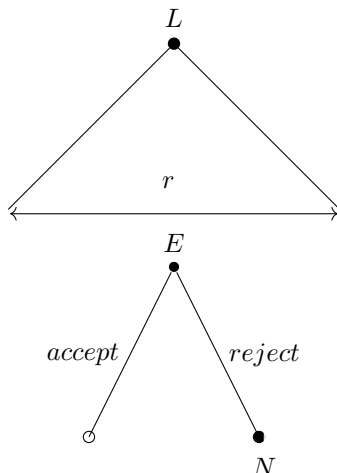


Figure 2.4: The classical divide-the-dollar game with the possibility of successful revolt.

different in that it can express ideas that are distinct from those of the party of power even if this does not meaningfully change the outcome of policy. This configuration amounts to an expressive benefit (ϵ) for those who join the opposition.

In order to isolate the change from the first setup of the game, I assume that the dictator and the dominant party have reached an equilibrium that requires the dictator to share some portion of total rents ($r_p = \rho$) with members of the dominant party. The dictator's choice in this context is the level of rents to offer candidates joining opposition parties ($r_o = x$).¹⁴

After observing the offers, the elite can choose to join the dominant party (d) or the opposition party (o), taking the offers of ρ or x , respectively, or reject both and revolt with probability of success ϕ . Similar to the previous setup, if rebellion occurs, the dictator pays costs of repressing the rebellion κ , and the winner of the rebellion receives the remainder. Thus, the utility function of the dictator is:

$$u_D = [1 - \rho]d + [1 - x]o + [(1 - \kappa)(1 - \phi)]r$$

¹⁴While it may seem odd to have a dictator financing an opposition party, this conceptualization of the relationship between dictator and opposition reflects the fact that the dictator allows the opposition to exist rather than expropriating it. Thus, any rents they receive, and the level of those rents, can be thought of as only existing due to the forbearance of the dictator. For evidence of this in the Russian context, see Reuter and Robertson (2015).

And the utility function for the candidate is:

$$u_C = [\rho]d + [x + \epsilon]o + [(1 - k)\phi]r$$

2.3.4 Results

The game will never result in rebellion provided one of the following conditions is met:

$$\frac{x + \epsilon}{1 - k} \geq \phi$$

or

$$\frac{\rho}{1 - k} \geq \phi$$

The above conditions show that joining either the dominant or opposition party is preferable to risking rebellion, and provided that these hold, the candidate faces a choice between either d or o .

The candidate will join the dominant party whenever the rents associated with dominant party membership are equal to or greater than the rents associated with the opposition party plus the expressive benefit that candidate derives from associating with the opposition party ($\rho \geq x + \epsilon$). In response, the dictator will offer the opposition party members rents that are equal to the dominant party rents, less the expressive benefit ($x = \rho - \epsilon$). In practical terms, the results show that the cost to the dictator of an opposition party is strictly decreasing in the value of the expressive benefit ϵ ; as the value of speaking one's mind increases, candidates joining the opposition need less compensation in rents (x) to keep them from rebelling. This result shows the value of having ideologically coherent opposition parties, as more of total resources are retained by the dictator. If we assume that the counterfactual is a candidate who may only choose to join the dominant party or reject the offer, then it becomes clear that allowing (and even financing) opposition parties is a welfare increasing prospect for the dictator. Indeed, an opposition party can be a way of limiting the total payout that a dictator

must make to mobilize a preferred proportion of the population. When an opposition party is present, members of the political elite can derive an expressive benefit by joining it, which illustrates how the presence of an organized political opposition will force members of the elite to reveal their type. The leader can then offer less rent because members of the elite receive the expressive benefit, enhancing welfare for the leader.

2.3.5 Ideological heterogeneity and sustainable systemic opposition parties

The logical next step for this game is to allow types of candidates to vary, such that some derive relatively high expressive benefits, while others derive low or no expressive benefits from ϵ . This situation is the result of allowing ideological distance between the individual and the dominant party to vary. To the extent that the dictator knows this information, the dictator can act as a **discriminating monopsonist**.¹⁵ By allowing for the existence of an opposition party that is ideologically attractive to some (but, critically, not most) of the population, the dictator can economize on staying in power.

It is clear that the dictator can save money by offering members of the systemic opposition rents as well as the expressive benefit from choosing a party that is closer to their ideal point. Thus, creating an opposition party allows the dictator to claim the surplus that would otherwise be enjoyed by the member of the opposition in a one-party regime. Moreover, the dictator also secures the surplus that otherwise would have gone to members of the dominant party. Put another way, the presence of an opposition party forces elites to reveal their type and show their preferences over ideology. The dictator can then offer differing levels of rents for the dominant and opposition parties that make the elites indifferent between the two, and the dictator pays all elites, including those in the dominant party, less than would otherwise be the case.

¹⁵One interesting mechanism for the game would then become whether the candidate had any ability to bluff and elicit a higher level of ϵ .

Further, as the ideological point of the opposition party (*ideology_o*) approaches that of the individual (*ideology_i*), the rents that individual will need to participate in the political process approach zero (see 2.5). This figure shows the rent ratios under which an opposition party is sustainable as a function of distance between the dominant party and the individual, holding the expressive benefit of opposition (χ) constant. The point r'' indicates that maximum level that a leader will be willing to pay to sustain such an opposition party, indicating that at the extremes of ideological distance multipartyism may not be sustainable. There is also a maximum value that the leader will place on keeping an opposition party, which is the ratio of the premium that the dominant party receives over the opposition party to the marginal cost of sustaining an opposition (r''). The implication of this threshold value is that at relatively low and relatively high levels of ideological distance, systemic opposition parties may prove unsustainable over the long term, as their members would demand rents that the regime would not be willing to pay considering the risks and costs associated with allowing these parties to exist.

The question remains, however, as to what sorts of inducements can keep at least some potential candidates signing up for the opposition. In other words, what determines the size of ϵ ? How can the dictator keep it sufficiently large to increase rents but not destabilize the *status quo* by granting meaningful freedoms and power to the opposition? The dictator can control these aspects of the game by actively managing the position of the party of power and the opposition party within the ideas space. Thus, in contrast to Gandhi's argument, there may be times when the best course of action for the authoritarian leader is to dominate the center of the idea-space, essentially delivering ideational goods that are broadly in line with what the majority of the elite want. However, the leader will also allow the formation of smaller parties that occupy the margins and convince elites who might not be willing to affiliate with the party of power a way to still reap the benefits of association with the regime.¹⁶

¹⁶Gandhi (2008, 79) considers the possibility that broad 'fronts' of political parties could give elites the benefits of association with the regime but reduce the costs of association with the ruling party in particular,

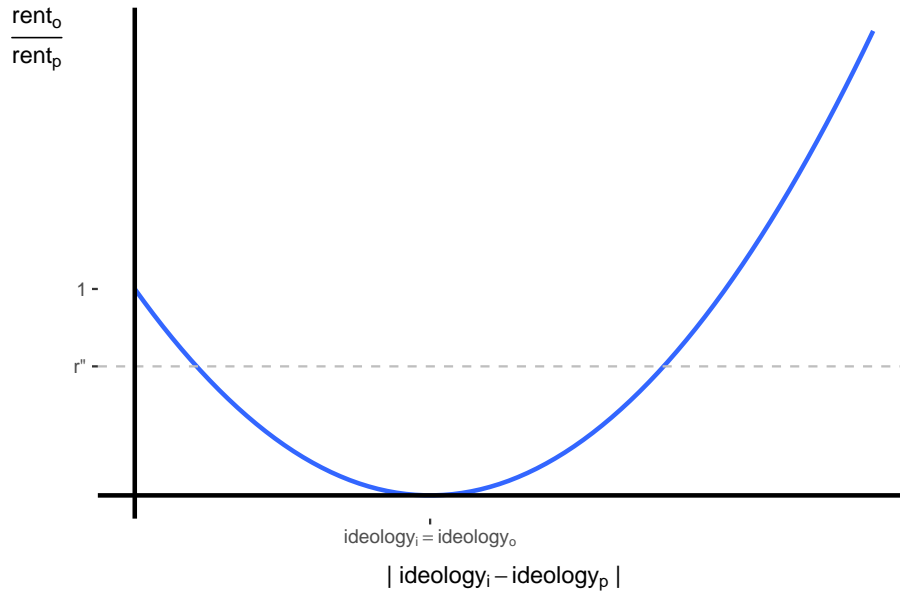


Figure 2.5: The rent-ratio necessary to sustain an opposition party as a function of ideological distance

2.4 Testable implications

The model above implies that certain things should be true in the empirical record. I focus here on ideological distance and ideological importance with respect to rents as they apply to three broad areas of behavior: defecting from the systemic opposition; new entrants deciding to ally with the systemic opposition; and the willingness of the systemic opposition to criticize the dominant party. With regards to ideological distance, the model predicts that as the ideological distance between a ruling party and a party in the systemic opposition grows (shrinks):

- 1.a. members of the systemic opposition are less (more) likely to defect and join a ruling party.
- 1.b. the systemic opposition is more likely to attract new candidates and supporters.
- 1.c. the systemic opposition is more (less) likely to spend time criticizing the regime.

We can also think about how shifts in dominant party rents or systemic opposition party

a situation in which multiple parties create a “menu” of options among which elites can choose given different degrees of “opportunism”. However, this possibility does not appear explicitly in her model.

rents will affect behavior. All else equal:

- 2.a. when dominant party rents increase, members of the systemic opposition who are ideologically closer to the dominant party will be more likely to join the dominant party.
- 2.b. when systemic opposition party rents decrease, members of the systemic opposition who are ideologically farther from the dominant party will be more likely to exit politics.

With respect to the importance of ideology vis-à-vis rents, the model predicts that as the weight on ideology (λ) grows (shrinks):

- 3.a. members of the systemic opposition are less (more) likely to defect and join a ruling party.
- 3.b. the systemic opposition is more (less) likely to attract quality candidates and electoral support.
- 3.c. the systemic opposition is more (less) likely to publicly criticize the ruling party.

The parameters in the model allow for other tests, as well; the list above is not comprehensive. In the following chapter I focus on hypothesis 2.a., examining how the creation of a second pro-regime party impacts members of systemic opposition parties. I conceptualize such an event as an expansion of the rents that are on offer via pro-regime activity, and hypothesize that those members of the political elite with the least ideological distance to the new party are most likely to join it.

2.5 Conclusion

The model predicts that at the extremes of ideological distance, multipartyism is unsustainable, while in the middle of the range, systemic opposition parties can exist indefinitely. In the latter scenario, introducing or allowing the continued existence of opposition parties represent Pareto improvements for the regime, as they maximize regime durability while simultaneously reducing the costs associated with mobilizing electoral support. Leaders and elites, who could potentially seek to eliminate multipartyism altogether, have come to understand that pushing for the total elimination of democratic institutions would reduce the

efficiency of their regimes, decrease individual welfare, and potentially even reduce their probability of staying in power. Additionally, as the Marginal Protective Value of political engagement falls, there is less incentive for the regime to purchase high levels of engagement. When this happens, multipartyism is less attractive because the monitoring costs for multiple parties are high and fixed, but having multiple parties delivers less efficiency in terms of economizing on the rents needed to sustain opposition parties. This type of situation is most likely to arise when the state has large repressive capacity. Thus the model provides a theoretical framework to explain the empirical finding that large repressive capacity is associated with greater levels of autocracy, even within autocratic regimes (Albertus & Menaldo, 2012).

This chapter built the theory with a formal model that specifies the ways in which leaders balance risks of revolt, ideological preferences, and maximizing rents, to select from an array of institutional designs. This initial decision of institutional design creates incentives for some elites and voters to join the political opposition *even though they do not expect it to take power*. The theory engages with and builds on recent work on the institutions of authoritarian regimes and how they enable regime survival. Specifically, it builds on prior formal work that has modeled both the outcomes and choice of institutions. On a conceptual point, the project seeks to bring ideology back into the study of institutions and institutional design and selection in authoritarian politics. In integrating ideas and ideology, this chapter argued that there is an intrinsic, expressive benefit to voting and political participation in authoritarian regimes, as has been claimed about voting in democratic systems (Brennan & Lomasky, 1997). In providing a model that incorporates both the ideological benefits and material concerns that link elites to voters, the theory also speaks to studies of clientelism and vote-buying in unconsolidated democracies (Stokes, 2005; Stokes et al., 2013). Acknowledging that voting rules and institutional design are often endogenous to the authoritarian leader's preference, this chapter seeks to think critically about the ways in which voters behave given an absence of a reasonable chance of engaging in strategic voting. As such, it seeks to

expand our formal models of strategic voter behavior from those in which rules are considered exogenous and party entry considered to be relatively free¹⁷. Finally, the model presented in this chapter can be applied to questions of political behavior in authoritarian regimes, including why candidates run for office and why voters support parties and candidates that they cannot meaningfully expect to alter the political process.

More broadly, integrating ideology into our formal theories of authoritarian politics helps us understand the resurgent authoritarian regimes that we see around the world today, especially in Eastern Europe, including the role that populist politics plays in the consolidation of political power.

¹⁷See, for example, Cox (1997, pp. 17–19), for a study of the ways in which voters engage in strategic voting in environments in which institutional rules and presumed to be exogenous in most cases most of the time and in which party entry is not a political decision of the incumbent.

Chapter 3

Chapter 3

3.1 Introduction

This chapter takes the model outlined in section 2.3 and evaluates how its predictions map onto decisions of legislators in the Russian Duma under President Vladimir Putin. Russia has seen dramatic change in its political institutions since Putin came to power over two decades ago. There has been a considerable marginalization of any meaningful political opposition or critical media, as well as growing restrictions on the ability of civic organizations challenge the power of the state. Most important for testing the theory presented here, however, is the fact that while Putin has overseen the construction of the dominant political party United Russia¹; he has also overseen the consolidation of ‘systemic opposition parties’ that compete regularly and win seats in federal, regional, and local legislatures. These same three or four parties do not meaningfully challenge the president, unlike during the 1990s, when a constantly changing slate of political parties formed and merged, some posing serious threats to then-President Boris Yeltsin’s grip on power.² Thus, Russia will serve as a case of a state being ‘treated’ with a new institutional design. My theory expects there to be heterogeneous

¹For a full account on the creation of United Russia as a dominant party, see Reuter (2017).

²For an account of the inability of parties in Russia to consolidate in the 1990s and early 2000s, see H. E. Hale (2007).

treatment effects based on ideology. I test these expectations by reviewing changes in party affiliation of Duma deputies, the legislators in the lower (main) law-making body at the federal level.

3.2 Case and Setting

By 2006, the Russian political system was increasingly authoritarian in nature. Additionally, Putin's popularity was routinely recorded at over 70% by reputable pollsters, and the dominance of United Russia was increasingly apparent, though the party lagged Putin in its popularity. Against this backdrop, Putin's deputy presidential chief of staff, Vladislav Surkov, orchestrated the merger of the relatively insignificant Party of Life with a number of smaller and even less significant parties to create A Just Russia³. In a speech to a Party of Life congress in early 2006, Surkov is quoted as explaining that the regime needs a "second leg" to walk on when the first "goes numb," betraying concern about the ability of United Russia to provide consistent support to Putin if its popularity sags⁴. A Just Russia, the Communist Party of the Russian Federation (KPRF), and the Liberal Democratic Party of Russia (LDPR)⁵ would be the three systemic opposition parties to win seats alongside the dominant United Russia party in the 2007 Duma elections.

The goal of A Just Russia was to create a second party of power, one that could occupy the center-left ideological space, ensuring no threat from the communists, and provide support for Putin. The party formed in October 2006 to compete in the 2007 Duma elections. Rather than happening behind the scenes, Kremlin involvement was on display, so it was readily apparent to observers and political elites alike that this party had Putin's blessing.

³The party's Russian name, *Spravedlivaia Rossiia*, is sometimes translated as simply "Just Russia" instead of "A Just Russia." The Russian language does not have articles, with the definite or indefinite nature of nouns often conveyed by context, and I am not a sufficiently skilled linguist to know which is more appropriate in this case. I have opted for "A Just Russia" because the alternative, "Just Russia" could be confused by an unfamiliar reader for "Only Russia," which is definitely wrong.

⁴As quoted in March (2009), which contains a detailed account of the party's creation.

⁵LDPR is widely understood to be neither liberal, nor particularly democratic, with its politics usually described as ultra-nationalist, xenophobic, and populist.

In case there was any doubt, “...Putin admitted he helped create Just Russia... [AJR] leader, Shestakov, was Putin’s judo trainer and his cousin Igor Putin (director of Samara Reserve factory) reportedly switched to [AJR] from United Russia”(March, 2009, p. 516). After a strong showing in regional elections in early 2007, and before the Duma elections in December of that year, AJR was gaining momentum and was performing as planned: “bringing different elite groups into the spoils system and allowing disaffected elites to engage in opposition without being seen as outright ‘traitors’” (March, 2009, p. 521). As such, its creation, and the reaction of the Russian political elite, offer an occasion to test the theory.

3.3 Inference strategy

The theory outlined in chapter 2.3 makes causal claims about how a period of authoritarian consolidation and the creation of a dominant party affects the choices of political elites. In the language of experimentation, consolidation and dominant party creation are the treatments whose impact I wish to assess. Because it would be impossible and unethical to randomly assign this treatment to political elites, I need to specify a causal inference strategy that will work with observational data. I propose using the authoritarian consolidation and creation of a dominant party under Putin was an unanticipated shock that, as a treatment, is applied to all of society. The variation comes not from having a control group against which to test the behavior of political elites, but from examining how the treatment will have differential impacts on members of the elite depending on how that shock affects the parameters of the model differently. I am testing the theory by looking for evidence of heterogeneous treatment effects consistent with what the theory predicts. In this sense, my research design benefits from my ability to exploit uniform treatments across dissimilar groups (e.g., members of the political elite with different ideological preferences), then test for heterogeneous treatment effects in the way that my theory would predict.

Authoritarian consolidation and dominant party creation under Putin is a single case,

albeit one that offers many potential ways to test the theory. The advantage of using a single case as an initial test of this theory is that virtually all important background variables are held constant. Admittedly, any findings are more limited in scope and the support for the theory that they offer is lessened by virtue of coming from a single case, but I am making a conscious decision to pursue stronger internal validity rather than external validity at this point. I discuss additional tests of the theory in the conclusion, tests that are better suited to make a case for the broad applicability of the theory across many different contexts.

3.4 Russia under Putin: a relevant case?

This chapter uses Russia under Vladimir Putin as a test case for the theory that understanding the intersection of material incentives and ideology can explain elite behavior in the face of authoritarian resilience. For Russia during this period to be a good test case, at least three things need to be true. First, it needs to be the case that Russia became more authoritarian during this period. Second, it needs to be the case that a dominant political party came to power during this period. Third, it needs to be true that ideology is, in fact, a meaningful way to understand Russian politics and the motivations of Russian political elite. This section reviews the evidence on all three counts.

3.4.1 Before Putin: central authority weakens under Yeltsin

To make the case that authoritarian consolidation occurred under Putin, we need first an understanding of the Russian political regime under the prior president, Boris Yeltsin. Yeltsin rose in popularity and political power throughout the 1980s by criticizing the slow paces of reforms under Gorbachev, and was elected president of the Russian Republic of the Soviet Union in May 1991. With the dissolution of the Soviet Union in December 1991, Yeltsin was elevated to head of the newly independent and constitutionally democratic Russian Federation.

One of the initial characteristics of Russian federal system was the relative strength of the regions⁶ with respect to the central government. The devolution of power to the regions stemmed from Yeltsin's 1990 struggle for power with Gorbachev, in which Yeltsin famously urged to the leaders of ethnic republics, "take as much sovereignty as [they could] swallow." This policy largely continued throughout the early 1990s, as Yeltsin traded central government authority to the regions in exchange for support for his liberalization policies. While Yelstin made an attempt to regain control after regional leaders attempted to dissolve the federal government in 1993, Yeltsin once again traded autonomy to the regions for political support during the 1996 presidential election. In the final years of his presidency, he was too politically weak to wrest meaningful control back from the regions (see Tolz and Busygina (1997), Zhuravskaya et al. (2010), and Kahn (2002) for details of the back-and-forth power struggles with the regions).

The state's monopoly on the use of force was also substantially weakened under the Yeltsin regime to a degree not seen since the Russian Civil War (1918-1921). Perhaps most visible to outside observers was the First Chechen War (1994-1997) in which the larger, better equipped, and better trained Russian military was unable to subdue Russia's breakaway Chechen Republic. Despite brutal attempts to regain control over the region, the Russian government ended its military campaign, and Yeltsin signed a peace treaty with Chechen leader Maskhadov in 1997 that all but acknowledged Chechen independence. At the same time, the Russian state proved unable to stop growing lawlessness, as criminal networks that had long existed under the Soviet regime flourished in the freer post-Soviet environment. According to World Bank data, homicides in Russia doubled from 1991 to 1994, and public opinion data showed that large majorities of Russians believed, year after year, that law and order was deteriorating⁷.

⁶Russia's largest sub-national units have many designations: region, republic, subject, krai, okrug; there was not uniform autonomy across all of these types of sub-national units, but the important point is that many enjoyed substantial autonomy in economic and political matters.

⁷See, for example, Reisinger et al. (2007, p. 33) for the results of surveys conducted in 1990, 1991, 1992, and 1995 asking about Russians' beliefs as to whether law and order was grown strong, stayed the same, or grown worse.

Russian politics in the 1990s was marked by pluralism, meaningful electoral competition, and, critically, genuine uncertainty about the outcome of elections. Parties opposed to the president captured seats in the Duma and regional legislatures, as well as governorships. Even Yeltsin's re-election as president in 1996 was by no means a guaranteed outcome. The communists (KPRF) had scored a substantial victory in the 1995 federal legislative elections, winning more votes than any other party and controlling nearly a third of the federal legislative chamber. Yeltsin's eventual 1996 victory was only achieved by striking a deal with wealthy private interests, who agreed to fund the Yeltsin campaign in exchange for controlling interests in many state enterprises. Even then, Yeltsin just barely won a plurality of votes in the first round, and bested his communist opponent Genadii Zyuganov by 54% to 40% in the run-off. Moreover, contemporary observers saw the potential for continued democratic reforms and consolidation, while acknowledging substantial challenges. Writing after parliamentary, presidential, and gubernatorial elections that took place in 1995 and 1996, McFaul (1997) noted high levels of participation, the fact that most elections were deemed free and fair, and that voters in the critical presidential race "rejected the Soviet past and a fascist future."

At the federal level, Yeltsin enjoyed substantial authority and was afforded more power than the legislature. Admittedly, the president had used military force in 1993 to resolve a dispute with communist hardliners in the parliament, after which a new constitution gave him even greater powers with respect to the parliament than he had enjoyed previously. Still, parliament was able to act as more than just a rubber stamp during the rest of his presidency. Reflecting on the evolution of parliament throughout the 1990s, Sakwa (2002) wrote "Large areas of policy-making have become the direct prerogative of the president, above all in foreign policy and national security policy. The president, moreover, has extensive decree powers. Yet all of this does not make the president all-powerful. The need to achieve parliamentary approval for legislation places it at the shared centre of Russian politics. The constitution, moreover, requires parliamentary confidence in the prime minister, and thus

by implication in the government.... The Duma [has been] able to carve out an important role for itself despite the formal provisions of the constitution but within the constitution's framework." Critically, Yeltsin was unable to control the legislature or build a party that could pose a serious challenge to the communists. While Yeltsin made an attempt to construct a dominant "party of power" in 1995, the Our Home is Russia party ended up neither dominant, nor powerful. Yeltsin and regional power brokers were reluctant to support the nascent party, it won only 10.1% of the vote in the 1995 legislative elections, and fewer than one third of regional governors were ever members of the party (Remington, 2008, p. 965).

Furthermore, there were centers of power independent of Yeltsin and outside of politics. Privatization in the early 1990s had created a class of wealthy, independent businessmen who controlled important economic resources and could translate them into political power. By the late 1990s, powerful independent voices existed in the media, too, and these voices had the capacity to change presidential policy. Media coverage of Russia's disastrous campaign to restore control over the breakaway republic of Chechnya contributed to Yeltsin's decision to withdraw government forces, accept an embarrassing military defeat, and hand the republic *de facto* independence. Furthermore, there were signs that the potential for independent, well-funded, critical voices in the media were likely to grow. Oligarchs that got rich in the 1996 loans-for-shares deals already owned or acquired media interests. Indeed, Vladimir Gusinsky, an oligarch who felt cheated by these deals, waged sustained campaign through his national television station, NTV, that resulted in the removal of Yeltsin's Deputy Prime Minister, Anatoly Chubais, in 1997 (Lipman & McFaul, 2001).

3.4.2 Authoritarian consolidation under Putin

When Putin first became prime minister in 1999, Russia could not be considered an authoritarian regime, and was more often described as a weak, developing, or hybrid democracy. The country suffered from high levels of corruption, organized crime, and weak political institutions, but the country was still rated as "partly free" by Freedom House. Shortly after

winning the presidency in 2000, however, Putin began to implement policies that were the first steps in his campaign to eliminate what democracy existed in Russia. His stated desire to build a “vertical of power” created alarm in some observers (Tyler, 2000), while his attacks on large independent media outlets rattled others (Lipman & McFaul, 2001). That said, at least some observers felt that Putin’s efforts, while worrying, were not inconsistent with continued democratic development in Russia (Nichols, 2002). Within a few years, however, it could not be denied that there had been a qualitative shift in Russian politics, with Putin amassing substantially more formal and informal power than Yeltsin had enjoyed. By 2004, Putin curtailed regional autonomy by pushing through constitutional reforms that made governors appointed by the president rather than elected, changed that caused Freedom House in its *Freedom in the World 2005* (2005) report to downgrade the country from “partly free” to “not free.” A critical component to this rating change was the degree to which electoral politics in Russia came to be dominated by United Russia, a new, permanent political party that supported Putin and acted as a permanent “party of power”.

3.4.3 United Russia as a dominant party

Since the early 2000s, the leadership of Vladimir Putin—first as president, then prime minister, then president again—has coincided with the dominance of United Russia and the presence of a small number of additional political parties that are legally allowed to form, nominate candidates, and win seats. United Russia was created in 2001 as the country’s “party of power,” a political organization that was explicitly in favor of the policies pursued by the president. The fact that United Russia became a permanent party of power was a departure from party politics of the 1990s, however, when parties were fluid affairs that would often morph from election to election. The existence of a permanent, national party structure was made possible by the rise in global oil prices in the early 2000s, which brought enormous revenue to the Russian state (see Figure 3.1).

United Russia won 300 out of 450 seats in the 2003 Duma elections, largely on the

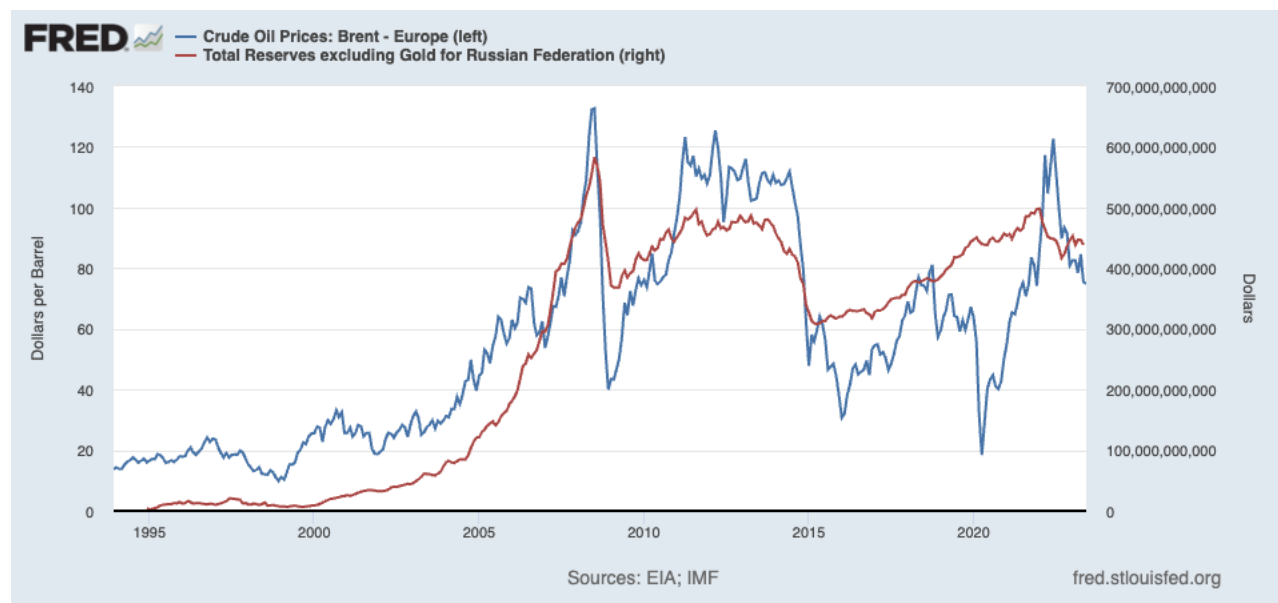


Figure 3.1: The rise in global oil prices in the early 2000s gave the Russian state new resources, enabling it to construct and support a permanent dominant party organization, United Russia.

back of Putin’s popularity and endorsement of the party⁸ While previously the party had focused on delivering Putin’s pro-market economic reform agenda, in the Fourth Duma (2003–2007) it passed legislation to centralize state control, limit political party organization, and marginalize non-governmental organizations, as well as provide substantial new patronage opportunities to state officials and party members (Remington, 2008). The transition to dominant party status was swift, as Russian politics became more authoritarian. As Reuter (2017, p. 28) writes, “United Russia membership provides elites with ... access to rents, policy influence, and career advancement opportunities. From 2003 onward, almost all legislative bargaining has run through United Russia legislative factions.... In elections, United Russia membership provides access to a valuable party brand, ballot access, and state administrative support.” By 2008 94% of regional governors were UR members, and by 2011 UR majorities in 82 out of 83 regional legislatures, majorities in 92% of (large) city councils (Reuter, 2017). This state of affairs was a far cry from the situation in the 1990s and first years for

⁸Putin is not, technically, a member of the party, illustrating the degree to which United Russia is a party that supports the president, not the other way around.

the 2000s, when the presence of viable “party substitutes” inhibited the creation of stable, national political parties (H. E. Hale, 2007).

With the creation of a dominant party, elections had become a far less viable means to capture power in any other party because it was clear that United Russia was the party favored by Putin. As such, United Russia would continue to reap the spoils, and meaningful electoral challenges to its authority would not be allowed. In the words of former Kremlin advisor Gleb Pavlovsky (2016, p. 15), “The electoral system in Russia is well developed and highly sophisticated. It is also completely useless. Elections are separated from the process of endowing the state with power; they amount to nothing more than an expensive ritual.” The surge in oil revenue that severed Putin’s dependence on the oligarch class and facilitated the creation of a dominant party are a clear example of how otherwise beneficial economic developments can have baleful political consequences when state institutions are weak (Menaldo, 2016).

3.4.4 The relevance of ideology in Russian politics

To what extent does political ideology matter in Russian politics? The fact that Russian politics, especially party politics, was turbulent in the 1990s⁹ could lead some to believe the political ideology is not a useful way to analyze or explain politics or political behavior. Additionally, Russian voters’ lack of experience with electoral competition between parties organized on clear ideological lines could mean that they did not understand political ideology as it is understood in other countries. However the evidence suggests this is not the case. There are important instances of political ideology explaining specific outcomes, such as the inability of the Russian parliament¹⁰ to pass constitutional reform in 1990-1993 (Andrews, 2009). Broadening the temporal scope, there are also strong arguments to be made that

⁹See H. E. Hale (2007) for a discussion of the inability of Russian political parties to consolidate during this period.

¹⁰Specifically, the data analyzed are the votes of the Supreme Soviet, a law-making body elected by the 1000 members of the Congress of People’s Deputies, which had been elected in the Russian Republic’s March 1990 elections.

the degree of ideological polarization explains patterns of economic policy reform in the first fifteen years after the fall of the Soviet Union (Frye, 2010).

3.5 Empirical test: who joins “A Just Russia”?

3.5.1 Theoretical expectations

The creation of AJR as the presidential administration’s “second leg”, a party that would support Putin but occupy different ideological territory than United Russia, is an occasion to test the observable implications of the theory by looking at the differential impact the event had on members of the systemic opposition. Because AJR was explicitly pro-administration and because its creation was supported by Putin, members of the political elite should have expected that it would provide as good or higher rents than the systemic opposition parties KPRF and LDPR were able to provide.

Additionally, in light of the growing authoritarian nature of the Russian political regime, the rents to KPRF and LFPR membership were declining, but the rents associated with United Russia membership were on the rise. Party membership, in this case, gives us our best proxy for the return to rents associated with switching parties: for members of the systemic opposition, switching represented the chance for greater rents in the future than they would otherwise expect; for United Russia deputies, there was little reason to expect this to be the case. This situation most closely reflect hypothesis 2.a., which predicts that when the rents on offer for the dominant party increase, members of the systemic opposition that are ideologically closest to the dominant party should switch to that party.

The logic behind hypothesis 2.a. reflects a single systemic opposition party. However, there are multiply parties that qualify as systemic opposition at this point, specifically KPRF and LDPR. Because the model predicts that legislators will consider both the material returns of political party membership and the expressive benefits that they derive from that membership, those members of the systemic opposition who are farthest from the ideological

center of their own party and closest to the ideological center of AJR to be most likely to switch. As a result, we need to control for each member's distance from their own party when establishing the relationship between their proximity to AJR and their probability of joining it. The creation of AJR represents the chance for all deputies who are not in United Russia to increase the rents they get from political participation; this will prove most attractive to deputies who would derive a relatively greater expressive benefit from AJR. Given that we can't measure the difference in rents directly, there are two main concepts that can be measured in order to determine whether, in fact, those who would gain a relatively greater expressive benefit from AJR are more likely to join the party: **ideological proximity to AJR** and **ideological distance from their own parties**.

3.5.2 Data

My outcome of interest is joining A Just Russia, which is a straightforward binary variable (join=1). I measure the two ideological concepts by generating two metrics for each deputy: the Hamming proximity to AJR and Hamming distance from that deputy's party. The Hamming distance is a distance metric that comes from information theory and quantifies the distance between two vectors. In this case, the vectors represent sequences of votes taken by Duma deputies. The Hamming distance is equal to the sum of the elements in a vector that have to change for the two vectors to be identical. For example, if deputy one voted $\{Y, Y, N, Y\}$ on four votes, and deputy two voted $\{N, Y, N, N\}$, Hamming distance would be 2 because to make the two vectors identical two of the elements need to be switched. This metric is well-suited to measuring vote data because the outcome—whether deputies voted the same way—is binary. Using roll call votes to measure ideological distance follows work by Andrews (2009), who used voting data to measure ideological polarization among Russian lawmakers and explain the failure to pass constitutional reform in 1990-1993.

The ideological proximity to AJR metric is defined as the distance each deputy is from the central point of AJR. To calculate this distance metric, I first compute the modal vote

among all deputies of the Third Duma (2003-2007) who eventually joined AJR in the Fourth Duma. Then, for all deputies, I count the number of votes that they took where they did not agree with the AJR modal vote. Because this is a distance metric, the proximity is one minus the Hamming distance. Similarly, for each deputy, I calculated their Hamming distance to their own party’s central point, similarly defined as the modal vote vector. This count, divided by the total number of votes taken, generated a normalized score, where the metric is the count of elements in a vector that have to change proximity to AJR and distance from the deputies’ own party by analyzing vote data for every member of the Third Duma (2003-2007) and measuring the Hamming distance between.

I collected the data for this analysis from two sources. The Duma Deputy party membership data came from the Russian Federal Elections Commission’s elections page.¹¹ Voting in Russia’s Duma is recorded electronically, and I collected the electronic vote records by scraping the data from the Duma’s legislative record page.¹² Votes were recorded as “for”, “against”, “abstained”, or “no vote”.¹³ Votes “for” were coded as 1, “against” as 0, and both “abstained” and “no vote” coded as 0.5.¹⁴ I also collected the names of all legislators in every Duma session since the first Duma (1993). This data allowed me to determine how many sessions each deputy had previously served in, as well as to create a variable for whether the candidate was male or female.¹⁵

3.5.3 Statistical model

In this section I outline the statistical model I use to test the theoretical expectations I lay out in Section 3.5.1. To recap, I have argued that the increasingly authoritarian nature of

¹¹<http://www.vybory.izbirkom.ru>, accessed in April 2018.

¹²<http://vote.duma.gov.ru>, accessed in November 2018.

¹³Za, Protiv, Vozderzhals’ia, Ne golosoval, respectively.

¹⁴The Hamming distance and proximity metrics generate a binary indicator of whether a deputy agreed with, for example, the modal vote of AJR on a given electronic vote. As a result, the numerical distance implied by this coding is immaterial; a deputy who voted “against” or “abstained” when AJR’s modal vote was “for” would get a distance of 1 for that vote.

¹⁵I did this by analyzing the patronymic (middle name) of each Duma deputy, which according to Russian tradition varies based on sex.

politics in Russia represents a negative shock to the rents for all members of the systemic opposition. At the same time, the creation of AJR and poses the opportunity to join a political party that is supported by the regime, implying greater potential rents over time, albeit not as great as those that could be captured through membership in United Russia ($\rho_{syst.oppos} < \rho_{AJR} < \rho_{UnitedRussia}$). The model predicts that the greater stream of rents will prove enticing enough for some deputies to switch. Specifically, for members of the systemic opposition that are close enough to AJR in ideological terms ($|I_{candidate} - I_{AJR}| \leq \tau$), switching makes sense, and we should observe deputies leaving their systemic opposition parties to join AJR. I have no theoretical or empirical basis upon which to define the threshold of ideological proximity at which switching should occur. As a result, I operationalize the model's expectations as saying that the closer a deputy is to AJR in ideological terms, the more likely they are to switch.

Identification strategy

I cannot randomly assign authoritarian consolidation to study whether its effect on elite behavior is consistent with my model. In order to test the model and identify the impact of authoritarian consolidation on party switching, the case presented here follows the logic of a natural experiment. Specifically, the rise of Putin, his consolidation of power, and the creation of a dominant party (United Russia) and a second pro-regime party (AJR) were unexpected. These events ran completely contrary to the experience of the 1990s, when decentralization of power and weakening of the state seemed to be the Russia's post-Soviet fate. The consolidation under Putin was only possible due to the steep and unanticipated surge in global oil prices that resulted in an increase in financial resources for the Putin government in the early 2000s. This unanticipated shock to members of the Russian political elite predicts different responses depending on whether members of the elite are ideologically closer to or farther from AJR and their own parties.

In any observational study, endogeneity and omitted variable bias are serious concerns.

When claiming that an observational study approximates a natural experiment, the greatest threats to inference come from the possibility that the treatment was anticipated, and that a simple pre- post- comparison will be contaminated because the values of the outcome variable observed prior to the treatment could reflect anticipation of the treatment. If this were the case, and political elites anticipated consolidation and acted accordingly beforehand, then any estimates my model returns would represent an underestimate in the effect. In the particular case I study—Russia under Putin—some could argue that the treatment was actually not much of a treatment at all. According to this thinking, critics could argue that because Russia did not sufficiently democratize in the 1990s, the “shock” of authoritarian consolidation under Putin was not much of a shock at all. Because political elites’ expectations were not that they would continue to live in a stable, functioning democracy, authoritarian consolidation under Putin does not pose a substantial enough shock or treatment to generate any change in behavior. If this is correct, any detected effects would be noise. I deal with each of these considerations in turn.

Not unanticipated: One threat to my inference strategy is the potential that either Putin’s authoritarian consolidation or the creation of a dominant party was foreseen by political elites in the late 1990s and early 2000s. If this is the case, then my assertion that these events were an unanticipated shock falls apart because it would mean that any effects that my model would measure in the mid-2000s would be substantially biased downward. Duma Deputies, anticipating rising authoritarianism would have hedged their bets early and ensured that they had other ways to secure their rents that did not rely on formal institutional politics. I would argue, however, that Russian political elites in the late 1990s could not have known that they were so swiftly headed towards a period of re-autocratization. Indeed, the degree to which the central state had been weakened, combined with Yeltsin’s erratic behavior, a military defeat in the breakaway republic of Chechnya, and a ruinous debt default, would have left observers in the 1990s deeply skeptical that Russia was on the verge of re-autocratization.

Not a shock: A potential threat to my inference strategy is that I have incorrectly

characterized the nature of the Russian political regime under Yeltsin. If Russia never democratized, then perhaps I am seriously overstating the degree to which Duma deputies would have experienced a negative shock to rents with the rise of Putin and United Russia. Admittedly, it would be a mistake to overstate the degree to which Russia democratized during the 1990s. Some go so far as to characterize Russia from the 1980s to the early 2000s as undergoing a long transition from one authoritarian regime (communist, single-party) to another (personalist, multi-party). According to Snegovaya (2023, p. 105), for example, the failure to replace communist era elites meant that Russian democratization never happened in any meaningful sense, and “the eventual re-autocratization of Russia was just a matter of time.” This analysis may be correct, but it was not the consensus view in the late 1990s or even early 2000s, and the phrase “a matter of time” is doing a lot of work.

Moreover, while there are clear examples of democratic transitions undermined by authoritarian elites who managed to protect their wealth and power via constitutional design, (Albertus & Menaldo, 2017), Russia is not one of them. The oligarchs that came to power in the 1990s did not come from the top echelons of Soviet political life. And while Putin eventually rose to the height of Russian politics, he had never been more than a low-level KGB functionary in the waning days of the Soviet Union, a far cry from the authoritarian “elite”. In the 1990s, he was a marginal player in the St. Petersburg municipal administration until he was plucked from obscurity by the Yeltsin administration. Regardless, it can be true that both 1) Russia failed to fully democratize and 2) central authority weakened dramatically in the 1990s. It is exactly this weakening of central authority that set the stage for authoritarian resurgence under Putin, and is what is required for a test of the theory.

Even if I am correct in describing this episode as a natural experiment, there are other theoretically plausible explanations for why some Deputies would join AJR, while others would not. For one, assuming that there are costs to switching parties, candidates with a longer time horizon are more likely to be willing to pay those costs. Moreover, there are reasons to assume that the costs of switching are not fixed, to the extent that switching to a

new party means sacrificing the existing relationships and networks that a deputy has built up over time. In Russia in particular, there alternative explanations merit consideration. Informal and reciprocal exchanges of favors (*blat*) and the importance of securing protection from powerful patrons (your *krysha*, literally “roof”) are important features of Russian society, politics included. I see time spent in politics as a reasonable proxy for both the length of time horizon and the relatively greater costs in the form of lost connections that some would have to pay to switch parties. I measure the number of sessions in which a Duma members has served to capture and control for these alternative explanations.

Model specification

The outcome variable, whether a candidate joins AJR, is binary, and I model the relationship between the deputy’s ideology and this outcome with a logistic regression model. Specifically, I model each deputy joining AJR ($join_i=1$) as:

$$join_i \sim \text{Bernoulli}(\text{logit}^{-1}(\mathbf{x}_i\boldsymbol{\beta})) \quad (3.1)$$

where \mathbf{x}_i is a covariate vector for each deputy that includes the deputy’s ideological distance to his or her own party in the Fourth Duma (2003-2007) (**party ideol. distance**); that deputy’s ideological distance to what will be the center of AJR (**AJR ideol. proximity**); the interaction of these two variables; the count of sessions that the deputy has been elected to (**Duma experience**); candidate **sex** (female=1); and a fixed effect for party membership.

I consider four separate model specifications. First, I consider a base model that includes the two ideology variables and all covariates. Second, I consider a model that just includes the interaction of the ideology covariates. Third, I add all covariates to the interaction model. Fourth, I use the same interaction model with additional covariates, but drop observations from deputies who had been members of the two main parties that merged to create AJR. The

reason to do this is that including these deputies and the fixed effect for party membership would likely introduce very strong colinearity between the outcome and the party fixed effects.

The model results for each of these specifications are shown in Table 3.1. As expected, models (1) and (3), the fully specified models with all Deputies, suffer from extremely high colinearity, which can be seen in standard errors that are unreasonably large given the parameter estimates. Between models (2) and (4) there are theoretical reasons to prefer model (4). In addition to the explanatory ideological variables, there is reason to think that factors like whether the deputy is male or female, or the amount of experience, or indeed, even the party of which the deputy is a member could have some explanatory power. Including these variables, as specification (4) does, is desirable for this reason. Looking at the model results, the Akaike Information Criterion (AIC) is substantially lower for model (4) than (2), meaning that the additional variables do lead to better model fit.

3.5.4 Model Interpretation

While the parameter estimates for the key variables are signed as expected and significant at the 95% level or greater in most or all models, logistic regression coefficients are difficult to interpret directly because their linear combination with observed covariates is transformed to the logistic scale. To address this challenge in interpretation, I follow King et al. (2000) to generate quantities of interest and estimate uncertainty in those estimates, reporting what the model expects based on cases that we can theorize about.

Using the `clarify` package (Greifer et al., 2023) in R, I estimate the change in model-predicted probability that a Duma deputy will join AJR if we change the values on the proximity to AJR and distance from own party variables from low (10th percentile of observed values) to high (90th percentile of observed values). I do this once for each party in model 4 (KPRF, LDPR, and UR), in each case using quantiles generated within those parties. Generating the quantiles for counterfactual simulation within each party is important because

Table 3.1: Results of models testing H1

	<i>Dependent variable:</i>			
	(1)	(2)	(3)	(4)
party ideol. distance	-0.001 (0.030)	-0.801*** (0.297)	-0.795* (0.461)	-1.374** (0.679)
AJR ideol. proximity	0.509*** (0.154)	0.144 (0.093)	0.274 (0.205)	0.383 (0.242)
party ideol. distance * AJR ideol. proximity		0.014*** (0.005)	0.014* (0.008)	0.023** (0.011)
Duma experience	0.025 (0.388)		0.028 (0.439)	0.410 (0.450)
sex: female	0.966 (1.168)		0.596 (1.215)	0.694 (1.230)
party: KPRF	-2.991** (1.448)		-2.629 (1.630)	-4.636** (2.220)
party: LDPR	-3.498* (1.903)		-2.055 (2.323)	-3.878 (3.034)
party: AJRM	18.305 (1,630.920)		18.835 (1,624.730)	
party: NPUM	2.274** (1.088)		2.914** (1.284)	
Constant	-31.618*** (8.203)	-12.230** (5.534)	-18.534* (10.721)	-23.863* (12.445)
Observations	468	476	468	419
Log Likelihood	-36.584	-54.787	-34.772	-28.731
Akaike Inf. Crit.	91.169	117.575	89.544	73.462

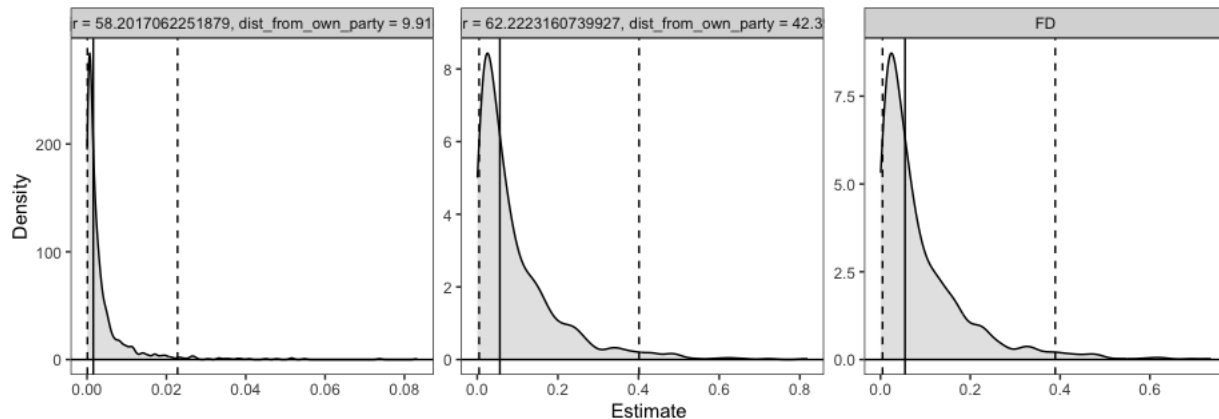
Note:

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

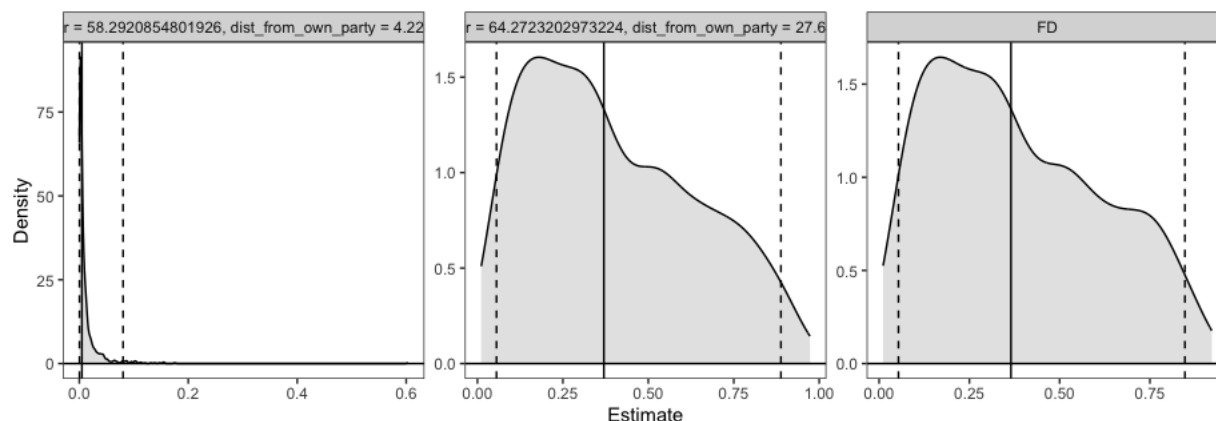
it ensures that I only consider values for these variables that actually exist in the data.

As expected, moving from low to high values on proximity to AJR and distance to own party increases the probability that KPRF or LDPR deputies will join AJR. For KPRF, the probability increases by 5.4 percentage points (ppt) (95%CI: [0.5ppt, 38.9ppt]) from 0.2% to 5.6%. For LDPR, the probability of joining AJR increases by 36.5ppt (95% CI: [5.3ppt, 88.7ppt]), from 0.5% to 40.0%. See Figures 3.2 and 3.3, respectively. There is no change in the probability that a UR deputy will join AJR (first difference=-0.2ppt, 95% CI: [-2.2ppt,3.8ppt]); regardless of the value of the proximity to AJR or distance to own party, the probability of UR deputies joining AJR remains less than 0.3%. The results for deputies from the two systemic opposition parties (KPRF, LDPR) and the results from the United Russia deputies are aligned with what the model would expect.

Figure 3.2: First Difference Estimate: KPRF



There are far more deputies who do not join AJR than who do, leading to a class imbalance problem that can result in statistical noise or cause computational problems for logistic regression. Even more concerning, class imbalance problems in logistic regression can lead to bias in coefficient estimates that could understate the probabilities estimated from the statistical model (King & Zeng, 2001). As it applies here, this means that the model could understate the probability associated with joining AJR. To deal with this possibility, I used the same model specifications and fit a variation of the Firth's rare events logistic regression

Figure 3.3: First Difference Estimate: LDPR

following Pühr et al. (2017), and the results were only slightly different (see Appendix Table 3.A.2). For the deputies at the 10th and 90th percentiles of the explanatory variables in KPRF, the first difference estimate was an increase of 6.0 percentage points, compared to 5.4 percentage points in the regular logistic regression model. For LDPR deputies, the first difference estimate actually decreased; the rare events model resulted in a first difference of 29.3 percentage points, compared to 36.5 percentage points in the regular logistic regression model. There was no meaningful change in the first difference for the United Russia deputies (first difference -0.1ppt, 95% CI: [-2.5ppt, 3.6ppt]).

I interpret the simulation results from the statistical model as preliminary empirical support for my formal model and theory. As expected, there is no strong incentive for UR deputies to leave their party because it would mean reduced access to rents and no meaningful difference in ideology. AJR was expressly founded to support Putin’s agenda, which is what UR already was doing. There was no benefit to be had in expressive benefit, so no incentive to lose out on rents. For KPRF and LDPR deputies, on the other hand, the story was quite different. There was reason to believe that the rents associated with this party, which was explicitly pro-regime, would be at least as good as those that they received in their current position. For those who were ideologically distant to their own parties *and* close to AJR, the incentive was greater to switch.

3.6 Potential for additional empirical tests

The creation of AJR and subsequent decisions by LDPR and KPRF deputies to join it or stay in their own parties offered a good case in which to test the predictions of hypothesis 2.a. While the data gathered for that test are not ideal for testing the other possible hypotheses outlined in Chapter 2.4, there are different cases in Russia under Putin that would be amenable to the expectations in Hypotheses 1 and 3.

Hypotheses 1.a, 1.b, and 1.c, all predict political elite behavior under conditions when the ideological distance between the dominant and systemic opposition parties is changing. A good case to test these hypotheses would be the increasingly bellicose and nationalistic actions and rhetoric of the Putin administration and United Russia in the late 2010s through today. Starting with the 2008 war with Georgia, Putin employed increasingly anti-western rhetoric and pursued aggressive, expansionist policies. This represents a marked shift from the relatively pro-market and pro-western policies and statements emanating from the Kremlin in the mid-2000s. This change in the policies and statements of the regime represents an ideological shift that narrowed the distance between it and LDPR, traditionally a vocal and organized home for nationalist political views. The model would predict that over this period LDPR members would be more likely to defect to United Russia and that LDPR would have a harder time attracting new candidates. Those who were interested in and attracted to nationalist views were increasingly likely to find a home in United Russia, particularly after Russia's 2014 invasion of Ukraine.

Hypotheses 3.a., 3.b., and 3.c. examine the impacts of increasing or decreasing weight on ideology. While hard to measure directly, background economic conditions could be a good proxy. In periods of a strong and growing economy, the rents offered by the dominant party are worth relatively less, which would appear in the model through a growing weight on λ , which determines the individual-level relative value of rents or expressive benefit and ideology. A booming economy means that individuals would see themselves as less dependent on rents directly, as the opportunity for other income streams are relatively greater.

Conversely, as economic conditions deteriorate, the weight on ideology would shrink relative to the importance of rents. To test this we could examine periods of negative economic shock and examine patterns of defection, candidate recruitment, public criticism of the regime to test the predictions of H3a-c.

3.7 Conclusion

This chapter has presented a first test of predictions that flow from my theory of authoritarian regimes as monopsony buyers of political activity that I presented in Chapter 2. The model I presented in that chapter argues that legislators consider both the material returns to elected office as well as the expressive benefit they derive from publicly associating with and advocating on behalf of a political party that advocates for policy positions they care about. I tested this theory by focusing on the creation of a regime-endorsed political party in Russia in 2006. This chapter uses the creation of A Just Russia and that party's ability to recruit candidates from existing political parties, testing whether ideological proximity to the new party predicted that legislators would switch party. It found that moving from low to high values on the explanatory variables among members of the systemic opposition were associated with small but statistically significant increases in the probability of switching to AJR. Similar analysis showed no statistically significant change in probability for members of United Russia, consistent with the models predictions about the heterogeneous effect of the creation of a dominant party and rising authoritarianism.

There are also important limits to the findings presented here. First, they are from a single study of the creation of a single political party, so they are necessarily limited in their implications for the theory more broadly. A natural extension of this finding would be to replicate the study at the regional level, seeing whether the patterns identified here are reinforced by regional legislators' decisions about whether to switch parties. Additionally, aside from controlling for number of sessions in which a deputy served, the model does

not yet account for the possible confounding role of network ties on the decision to switch parties. But we know that these informal networks matter a great deal in Russian politics, and finding a better way to measure these relationships would increase confidence in the statistical model's ability to capture reality.

Additionally, further research could explore additional and more sophisticated ways to model individual-level desire for rents and ideological preferences. While using party as a proxy for each deputy's material returns to switching is a reasonable first step, more could be done to determine whether some individuals care more about rents relative to others. Additionally, vote history is a logical way to measure ideological similarity and dissimilarity between legislators, but additional research into preferences over particular issues could enrich the modelling approach taken here. One possible path forward would be using topic modeling or large language models to encode legislative text and floor speeches associated with each deputy, adding richness to the measurement of the ideology variable. Preliminary analyses showed no strong relationships between topics of speech, as determined through topic modeling, and party membership, but recent advances in language modeling could potentially offer an interesting path forward.

Finally, the statistical analysis conducted here treated deputies as independent of the actions of other deputies, but this is likely too strong an assumption. A model that incorporated the probability of joining AJR conditional on other similar deputies doing so could offer greater insights and more explanatory power. Indeed, given the important role of personal and familial relationships among Russian political actors in predicting outcomes (H. Hale, 2014), a model that took into account network structures would be a logical extension of this work. The data collected for the analysis in this chapter could form some of the node level data for a network based model, but it would require edge data, too. This could be an encoding of qualitative data, such as an analysis of known patterns of relationships, or quantitative, such as creating an edge-level metric for Hamming Distance between each pair of deputies.

Appendix

3.A Additional regression analyses

Table 3.A.1: Results of models testing H1 (OLS)

	<i>Dependent variable:</i>			
	(1)	(2)	(3)	(4)
party ideol. distance	0.001** (0.001)	-0.085*** (0.007)	-0.039*** (0.007)	-0.042*** (0.007)
AJR ideol. proximity	0.014*** (0.002)	0.004 (0.002)	0.0004 (0.003)	0.002 (0.004)
party dist. * AJR prox.		0.002*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)
Duma experience	-0.006 (0.009)		-0.002 (0.008)	0.002 (0.009)
sex: female	0.021 (0.025)		0.014 (0.024)	0.012 (0.024)
party: KPRF	-0.131*** (0.030)		-0.075** (0.031)	-0.092*** (0.035)
party: LDPR	-0.132*** (0.035)		-0.024 (0.038)	-0.040 (0.045)
party: AJRM	0.604*** (0.060)		0.647*** (0.058)	
party: NPUM	0.616*** (0.048)		0.591*** (0.046)	
Constant	-0.729*** (0.105)	-0.153 (0.131)	-0.002 (0.156)	-0.067 (0.197)
Observations	468	476	468	419
R ²	0.786	0.632	0.802	0.219
Adjusted R ²	0.782	0.630	0.798	0.205

Note:

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

Table 3.A.2: Rare Events Model Summary

	Coefficient	Standard.Error	Lower.95..CI	Upper.95..CI
party ideol. distance	-0.95	0.54	-2.34	-0.03
AJR ideol. proximity	0.30	0.22	-0.10	0.77
party dist. * AJR prox.	0.02	0.01	0.00	0.04
Duma experience	0.45	0.42	-0.43	1.23
sex: female	0.90	1.13	-1.55	2.73
party: KPRF	-3.42	1.87	-7.99	-0.02
party: LDPR	-2.45	2.58	-8.65	2.04
Intercept	-19.73	11.28	-41.84	2.38

Model fitted by Penalized ML Method: 1-Wald, 2-Profile penalized log-likelihood, 3-None

Likelihood ratio test=41.30064 on 7 df, p=7.088e-07, n=419 Wald test = 78.03769 on 7 df, p = 3.463896e-14

3.B 2003 Candidates Who Repeat

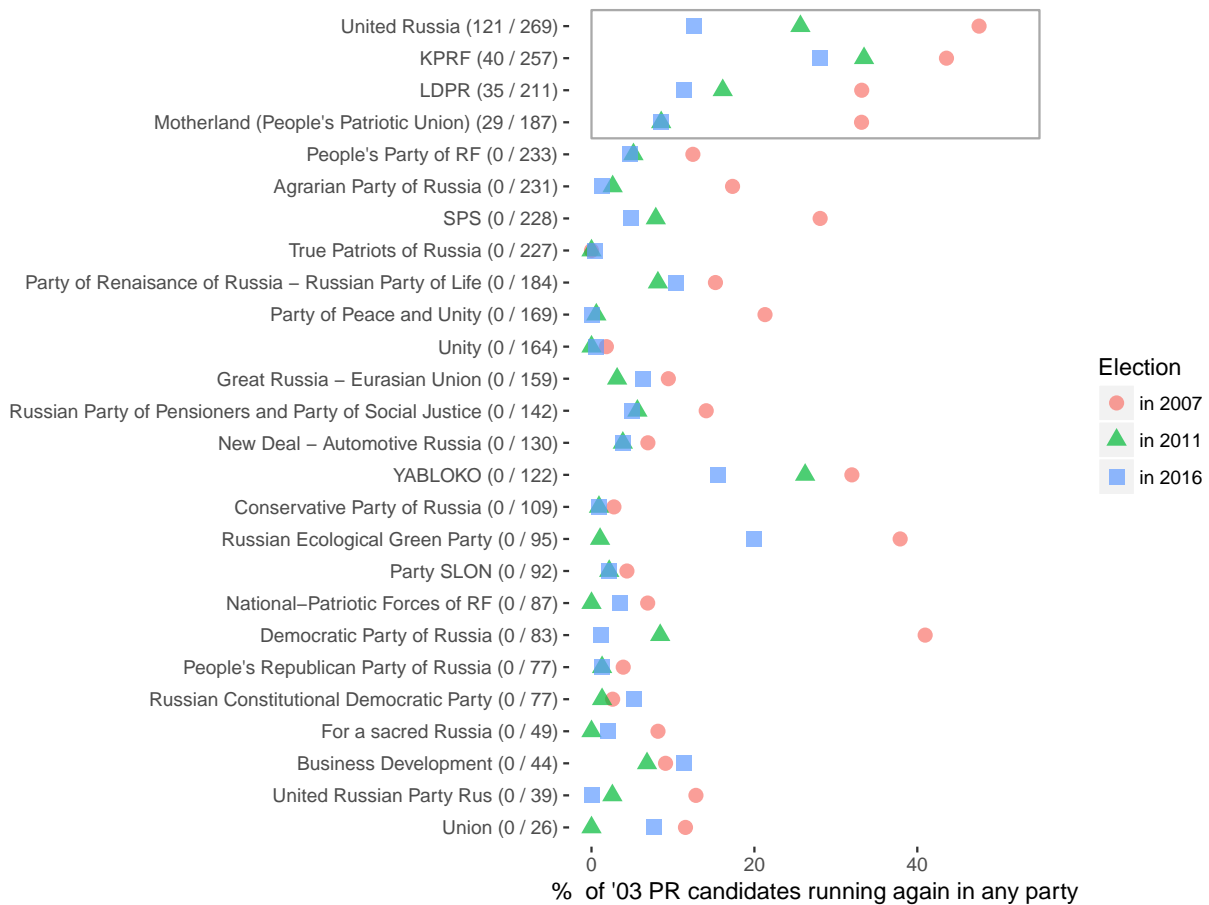


Figure 3.B.1: The figure above shows what percentage of candidates who ran for the Duma in 2003 on a party list for a given party and then ran again in 2007, 2011, and 2016 for any party. The party labels include each party's total seats won out of total seats nominated in parentheses. The parties that won any seats in the 2003 election from the party list are boxed.

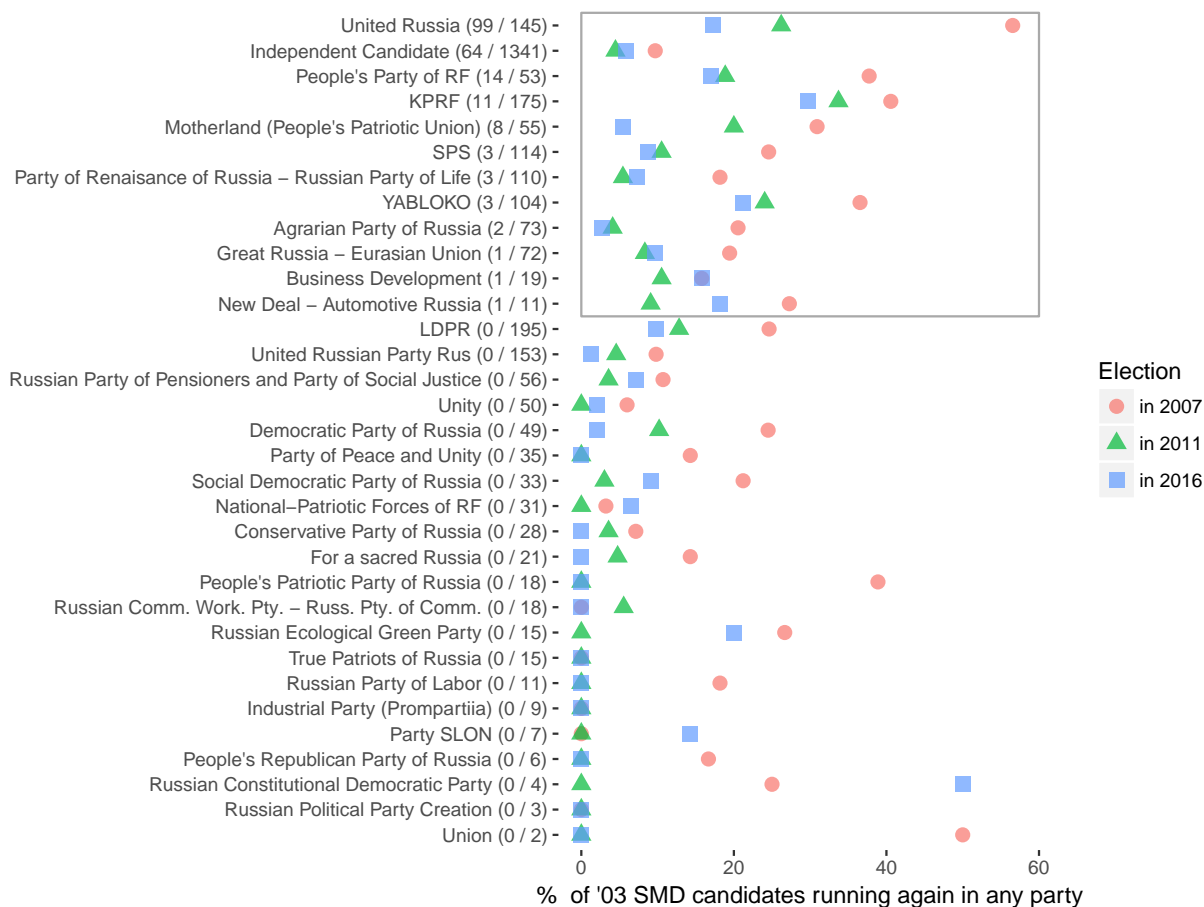


Figure 3.B.2: The figure above shows what percentage of candidates who ran for the Duma in 2003 in a single member district for a given party and then ran again in 2007, 2011, and 2016 for any party. The party labels include each party’s total seats won out of total seats nominated in parentheses. The parties that won any seats in the 2003 election from the party list are boxed.

Chapter 4

Chapter 4

Co-authored with Aaron Erlich

4.1 Introduction

A well-developed international relations literature demonstrates that when a country is at war, its population will “rally ’round the flag.”¹ Prior work on diversionary war and the “rally effect” argues that the new, highly salient threat posed by a wartime adversary causes a surge in individuals’ support for their home state and its leaders (Davies, 2002; Lai & Reiter, 2005; Mueller, 1973; Russett, 1990), anger, and support for military response to the threat (Lambert et al., 2010). In this article, we seek to improve our understanding of the

¹This article was co-authored with Aaron Erlich and published in *International Studies Quarterly* (Erlich & Garner, 2021). This text deviates only slightly from that publication, mainly with figure captions being moved to the main body of the text in several places. In that publication, we thanked the editors at *International Studies Quarterly*, two anonymous reviewers, Tymofii Brik, Megan Bradley, Michael Burch, Sarah Bush, Henry Hale, Egor Lazarev, Volodomyr Kulyk, Matthew Light, James Long, Krzysztof Pelc, Lee Seymour, Gavin Slade, Geoff Wallace, Nora Webb Williams, and seminar participants at Centre for International Peace and Security Studies (CIPPS) at McGill University; the Severyns Ravenholt Seminar in Comparative Politics at the University of Washington; Centre for European, Russian, and Eurasian Studies (CERES) at the University of Toronto; Midwest Political Science Association (MPSA) 2018 Annual Meetings; Urban Space (Kyiv); and Temple University. All errors remain ours. The authors collected these data in collaboration with the National Democratic Institute, which funded the research. The analysis and conclusions contained here solely represent those of the authors. The research comes under either University of Washington IRB protocol #00007684 or McGill University REB protocol #390-0217. The data underlying this article are available on the ISQ Dataverse, at <https://dataverse.harvard.edu/dataverse/isq>.

rally effect by engaging with three questions about attitudes in the context of a wartime rally.

First, how do ethnic ties to a wartime adversary state relate to attitudes towards the home state and the aggressor? On the one hand, a comparative politics literature on ethnicity and conflict suggests that the outbreak of a conflict can create or exacerbate ethnic divisions within a given state and prolong fighting (Buhaug et al., 2014; Fearon, 2004; Kaufmann, 1996). If ethnic identity becomes a salient division in a society at war, members of an ethnic minority group may express attitudes about that war that differ from those of the majority ethnic group, particularly when the minority group has ties to the wartime adversary (Mylonas, 2012). Suspicions that minority populations are disloyal can exacerbate tensions between these populations and their governments and, in some cases, lead governments to commit civil and human rights abuses (Wallace, 2015). On the other hand, there is evidence that the threat of external conflict strengthens ties to the home state across ethnic groups and forges new, pro-state identity. Increased pro-state identity may occur because the state becomes more intimately involved in the lives of its citizens to meet the exigencies of war (Herbst, 1990; Tilly, 1990). A newly salient, hostile outgroup raising the importance of civic identity, regardless of ethnicity, may also contribute to higher levels of pro-state identity (Saideman & Ayres, 2008).

Second, does proximity to conflict predict different levels of support for the home state and antipathy towards the aggressor within the context of a rally? There is considerable variation in how individuals experience war on the local level, with these experiences often diverging from higher-level narratives about the conflict (Balcells, 2012; Kalyvas, 2006; Wood, 2008). Spatial variation in the risks associated with the conflict could lead to differences in the salience of the conflict, impacting attitudes towards the home state and the aggressor (Bakke et al., 2009; Tellez, 2019).

Third, can we trust the data? Examining attitudes about the home state and the aggressor during wartime presents a major methodological difficulty because these are the opinions

that respondents would most likely feel pressure to falsify. Survey respondents may feel considerable pressure to voice support for their home country and hostility towards the adversary in times of conflict, especially when asked the question by a survey enumerator whom they have just met. This problem is acute when trying to elicit the views of ethnic minority populations who have ties to the wartime adversary. History shows that such sub-groups are vulnerable to persecution based on their identity, so they may be ready to overstate their loyalty to the state and their hostility towards the aggressor.²

To answer these questions, we examine Ukrainian citizens' attitudes towards Ukraine and Russia in the context of the ongoing conflict between the two countries. The case is an ideal one to study our questions because the conflict between Ukraine and Russia led to a documented "rally effect" in Ukrainian public opinion that saw greater support for Ukraine and hostility towards Russia (Korovkin & Makarin, 2019; Kulyk, 2016). Ukraine's substantial self-identified ethnic Russian minority could plausibly feel allegiance to Russia and may not feel the same pro-Ukraine sentiment as other respondents. Moreover, the risks associated with the conflict have always been greatest in the east of Ukraine, meaning that spatial variation in the salience of the conflict could lead to regional variation in attitudes towards the home state and the aggressor. Finally, the survey work that documents the rally effect in Ukraine does not explicitly engage with the fact that respondents may feel intrinsic or extrinsic pressure to state pro-Ukraine and anti-Russia views, potentially biasing researchers' estimates of public opinion.

We measured 600 respondents' attitudes towards Ukraine and Russia in four Ukrainian cities in April 2015, roughly a year after the start of hostilities. We used a traditional questionnaire to evaluate explicitly held attitudes towards Ukraine and Russia, then took the difference of these attitudes to calculate each respondent's *explicit bias* in favor of one country or the other. Respondents also took a computer-based *implicit association* test

²Examples include persecutions of "fifth column" populations, such as the internment of U.S. citizens of Japanese descent during World War II and the ethnic cleansing in the Soviet Union before, during, and after World War II (see Martin, 1998).

(IAT) that revealed their automatic, reflexive bias in favor of one of the two countries. Implicit associations are valuable to researchers because they are often strong predictors of behavior, sometimes stronger than explicitly held views (Roccatto & Zogmaister, 2010). Implicit associations do not always correlate with explicitly stated attitudes, a phenomenon called *dissociation*. Dissociation is particularly likely in settings in which respondents may feel reluctant to state their opinion frankly or when they may have multiple, cross-cutting loyalties to consider in forming an explicit opinion. If individuals are dissociating, then traditional survey research, which records only an explicit attitude, does not tell the whole story. Our research design allows us to test for congruence of implicit and explicit bias, as well as both *pro-Ukraine dissociation*, i.e., holding implicit preferences for Russia but voicing a pro-Ukraine explicit bias, and *pro-Russia dissociation*, in which respondents state preference for Russia when they actually have an implicit bias for Ukraine.

To preview our results, we find that, on average, members of all ethnic identity groups, including ethnic Russians, in all four cities express implicit and explicit preference for Ukraine over Russia, although in some cases not at levels of statistical significance. That said, self-reported ethnicity does serve as a meaningful predictor of relative levels of pro-Ukraine sentiment: respondents who identify as ethnic Russians express a weaker preference for Ukraine than do those who identify as ethnic Ukrainians. We find no consistent support for conflict proximity as a driver of bias. We use nationally representative survey data and weights based on city-level election results to check for robustness.

Comparing the results of the IAT and the explicit questionnaire, we find moderate evidence of attitude dissociation, with just under two in five respondents reporting explicit biases that are at odds with their implicit biases. Ethnicity is less important than the nature of respondents' implicit biases in predicting dissociation; dissociation is only somewhat more common among ethnic Russian respondents than among ethnic Ukrainians, but those who are implicitly pro-Russia are more than twice as likely to dissociate than those who are implicitly pro-Ukraine. Still, to our surprise, we found a large number of respondents disso-

ciating in favor of Russia—that is, having a pro-Ukraine implicit bias but stating a neutral or pro-Russia explicit bias. This pro-Russia dissociation runs counter to what one would expect from a classical view of attitude dissociation, i.e., that the war would induce those who are implicitly pro-Russia to conceal this implicit association by stating pro-Ukraine explicit attitudes, while those who were implicitly pro-Ukraine would be free to express this opinion. This unexpected finding merits further study in its own right and in other contexts.

This study makes important advances in our knowledge about conflict, identity, and research methods. First, it dispels the notion that ethnic minorities will necessarily feel allegiance to their purported external ethnic homeland in times of conflict. Second, it shows that ethnic identity is an important predictor of the degree of home-state bias in a society where a rally effect has occurred. Third, it documents how those who implicitly favor the wartime adversary dissociate at a higher rate than those who are implicitly biased in favor of the home state. This last finding underscores the value in studying implicit associations as well as explicit attitudes, building on work that has used the IAT to measure potentially sensitive opinions in similarly challenging settings, such as attitudes towards the leader of an authoritarian regime (Truex & Tavana, 2019). While others have deployed tools to measure social desirability bias or other forms of deliberate attitude falsification in war-zones (see Blair et al., 2014; Lyall et al., 2013), we believe we are among the first to study the congruence or dissociation of implicit and explicit biases towards the belligerents in a military conflict. Had we only looked at explicit survey responses we might have misstated the role of ethnicity in predicting attitudes about the belligerents and we would have completely missed the sizable proportion of our sample that dissociated *in favor* of the adversary.

4.2 Case selection

After several months of political turmoil in Ukraine, Russian forces invaded and seized the southeastern Ukrainian region of Crimea in February and March 2014. Shortly thereafter,

the Russian military began to provide support to an insurgent movement fighting government forces in eastern Ukraine. In the summer of 2014, Russian troops invaded eastern Ukraine to support these insurgents in their fight against Ukrainian military. A ceasefire in September 2014 was abandoned amid heavy fighting, but was followed by a second ceasefire signed in February 2015. This second ceasefire has held and led to the establishment of a “line of contact” monitored by the Organization for Security and Cooperation in Europe (OSCE), although violations have occurred frequently, and both civilian deaths and military casualties continued.

There are at least three reasons why this conflict provides a good case for study. First, the conflict led to a pro-Ukraine, anti-Russia rally in public opinion (see Figure 4.2.1). In September 2014, Ukrainian survey respondents reported that their attitudes towards the Ukrainian national anthem, flag, language, and Ukraine’s independence, had improved by at least 25 percentage points from before the conflict onset (Kulyk, 2016, p. 599). Attitudes towards the Russian state dropped by nearly 50 percentage points. Nationally representative public opinion data from before and after the conflict confirm respondents’ recollections by showing a sharp deterioration in public attitudes towards Russia after the start of the war.³

Second, Ukraine’s large ethnic Russian minority is a compelling group to study within the context of a wartime rally. On the one hand, relations between ethnic Ukrainians and ethnic Russians were peaceful (Posen, 1993), identity was not politicized to the degree it was in other post-Soviet states, and Russians did not suffer civic discrimination after independence (see Fournier, 2002). Further, in the first decade of independence, politics hinged less on ethnic identity cleavages than it did on left-right ideological divides, arguments over economic integration with Europe versus the Russia-led Commonwealth of Independent States (CIS), attitudes towards the Ukrainian state, and elite politics (Abdelal, 2005; D’Anieri, 2007; H. E. Hale, 2011; Shevel, 2002). On the other hand, Russian government messaging and media framing at the start of the war, much of which broke through to Ukrainian audiences,

³Korovkin and Makarin (2019) further discuss the way that Ukrainians’ attitudes towards Russia deteriorated after the conflict’s onset.

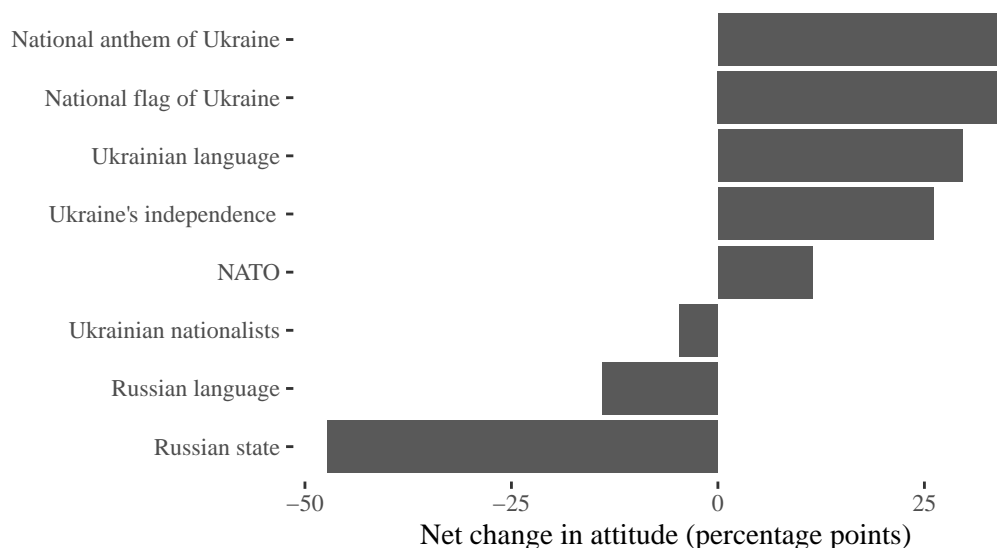


Figure 4.2.1: Data from September 2014 reported in Kulyk (2016) show a substantial rally effect as measured by respondents' reported changes in attitudes towards subjects associated with Ukraine and Russia. The values in the figure are the difference between the percentage of respondents who said that their attitude towards a subject improved a lot or a little and the percentage who said their attitude worsened a lot or a little.

framed the conflict as one where radical, neo-Nazi elements in the Ukrainian state and society committed atrocities against ethnic Russians (Lankina & Watanabe, 2017). This appeal could have fragmented society along ethnic lines, particularly in light of the fact that Ukrainian national identity has been contested and weakly-defined in the post-independence period (Korostelina, 2013).

Research conducted since the conflict's onset appears to confirm the low political salience of ethnic identity in Ukraine. Frye (2015) shows that whether a candidate for office was ethnically Russian or spoke Russian was less important to voters in Ukraine than policy positions. There is also evidence of renewed civic identity in Ukraine, at least relative to civic identity in Russia (Goble, 2016). While Sasse and Lackner (2018) show that the increased feeling of Ukrainian identity is stronger among the Ukrainian-speaking and bilingual population as opposed to monolingual Russian-speaking Ukrainians, Giuliano (2018, p. 170) documents a "trend of non-mobilization" along ethnic lines persisting after the outbreak of the war, arguing that material interests and responses to national-level political developments were

more important than ethnic identity (Giuliano, 2015). All of these studies examined the explicitly held views, meaning that we still know very little about the implicit associations of these respondents. Precisely because ethnic identity appears to be low-salience in Ukraine, application of the IAT will help us better understand the degree to which explicitly stated attitudes match implicit associations.



Figure 4.2.2: The four cities in which IATs were conducted

Third, all the cities in which we ran the lab have both self-identified ethnic Russian and Ukrainian populations, but their different locations mean they have had different experiences with the conflict. The map in Figure 4.2.2 shows the location of these cities, with shaded regions indicating Crimea, which was occupied and then annexed by Russia in March 2014, and regions partially controlled by separatist forces at the time of survey enumeration. Some cities were more affected and more at risk of being affected than others, with Kharkiv at the greatest risk during the study period. Of all the cities in which we collected data, Kharkiv was the closest to the heavy fighting around Debaltseve in February 2015. It was also the scene of intense riots and unsuccessful attempts by anti-Ukraine insurgents to seize government buildings and territory earlier in the conflict. Moreover, the city is closest to

Ukraine’s long border with Russia, making it most at risk from the buildup of Russian military equipment and personnel along the border that began early in the conflict. Data from a nationally representative survey conducted in May 2015, one month after our data collection, showed that respondents in Kharkiv had the highest expectation of conflict-related violence compared to the other three cities.⁴ In short, at the time of data collection, Kharkiv was the city out of our four locations most likely to come under threat from Russia-backed insurgents or Russian military action.⁵

4.3 Theoretical framework

The “rally around the flag” effect is a surge in public support for a political leader or government in the wake of a major international crisis or war. These rallies have occurred in countries and conflicts as diverse as the United Kingdom during the Falklands War and Persian Gulf War (Lai & Reiter, 2005), the United States following the terrorist attacks of September 11, 2001 (Kam & Ramos, 2008), and Russia after it invaded the Ukrainian region of Crimea in 2014 (H. E. Hale, 2018; Theiler, 2017). Cross-national quantitative analysis finds that leaders may initiate international conflict in order to increase their support at home by exploiting this rally effect (Davies, 2002).

There are multiple proposed mechanisms for the rally effect, but one of the most studied is the “patriotism” mechanism Mueller (1973). Patriotism links an individual’s reaction to international conflict and their change in opinion about their government, drawing on the logic of social identity theory (Tajfel & Turner, 1979). According to social identity theory, individuals derive personal value from group identity, want to have positive feelings toward their group, and develop an in-group bias relative to out-groups. Proponents of the patriotism explanation for the rally effect argue that military conflict creates a newly salient, foreign enemy, a new out-group, casting the entire nation as the relevant in-group.

⁴For more information on the survey, see the Results section, below.

⁵This grouping is subjective, and a case can be made for considering all of the cities except Kyiv as proximate to the conflict and vulnerable to spillovers in violence, a possibility we consider in Appendix 4.A.

This stronger sense of identification with the entire nation during times of crisis and war leads to a more positive appraisal of the nation and, by extension, the political figures and institutions that represent it. Individuals feel a strong bias in favor of their home country relative to the adversary, and political leaders benefit by way of a surge in popular support.

While generally supporting the rally effect and the patriotism mechanism, prior work has also noted that the effect of the rally is mediated in important ways by political ideology (Kobayashi & Katagiri, 2018; Merolla & Zechmeister, 2013) and prior support for the political leader and government policy (Sigelman & Conover, 1981), as well as identity factors like race and gender (Perrin & Smolek, 2009). Still others have noted that sub-populations experience very different emotional reactions to events that produced a rally within the population at large (Ojeda, 2016). We extend this research by examining how ethnic identity and geography may predict different attitudes towards the home state and the aggressor against the backdrop of a wartime rally caused by a conflict that has potentially important ethnic and spatial dimensions.

Researchers of conflict, civil war, and the comparative study of ethnicity have shown that ethnic divisions can become highly salient in conflict settings, marking the boundaries of relevant in- and out-groups in a way that could cause different ethnic groups within a given state to develop different views on the same conflict. Ethnic divisions and ethnic exclusion can contribute to explaining both the duration (Fearon, 2004) and onset of civil wars (Buhaug et al., 2014; Wimmer, 2013), particularly over territory seen as a group's homeland (Toft, 2010). Acts of violence, manipulation of information, or perceived injustices by the ethnic other can activate ethnic animosity, even where it did not exist previously (Petersen, 2002). At the most extreme, conflict that is fought along ethnic lines "destroys the possibilities for ethnic cooperation" (Kaufmann, 1996, p. 137), making multiethnic society impossible.⁶ In these instances, conflict is fundamentally about questions of whether a group is ruled over by its own members, defined in ethnic terms.

⁶Admittedly, many downplay the role ethnic diversity plays as an *independent* cause of civil war (Collier & Hoeffler, 2004).

Ethnicity can become salient or emerge as an imagined category of social distinction Anderson, 1991, even in societies where individuals historically attached relatively low salience to ethnic identity or often did not think of themselves in such terms, such as many in Ukraine before the start of the conflict in 2014 (Shevel, 2002). Kuran (1998) argues that an exogenous shock can cause *ethnification*, whereby individuals increasingly coordinate behavior along ethnic lines in expectation that others will also do so. This reputational cascade can lead some to express greater ethnic attachment than they feel, a type of ethnic preference falsification. This process is similar to the process of “ethnic polarization” identified by Somer (2001, p. 128), in which an exogenous shock pushes individuals to embrace a “divisive image” of ethnicity, one that “implies a definition of ethnic identities as mutually exclusive and incompatible with belonging to the same nation.” In this account, individuals begin to behave in ways consistent with their new conceptualization of polarized and mutually exclusive ethnic categories.

We argue that war between states with different majority ethnic groups could provide an exogenous ethnification shock to those groups. Indeed, the rally effect, which demonizes the adversary state, could create or increase feelings of persecution, insecurity, and fear among the ethnic minority with ties to the aggressor. Under such conditions, ethnic identity would become more salient for ethnic minorities who are co-ethnics with the majority group in the adversary state.⁷ Even if they had no feelings of antipathy toward the host state previously, ethnic minorities could come to fear that state during a conflict with their external sponsor because states in low-information environments use ethnicity as a marker for loyalty, repressing those groups they perceive as disloyal (Blaydes, 2018; McNamee & Zhang, 2019; Wallace, 2015). Such a scenario could cause ethnic identity to become the salient group-level identity *within each belligerent state*. Under such circumstances, ethnic identity would predict attitudes towards the belligerent states, with the non-core ethnic group in each country sympathizing with its external supporter, i.e., the adversary of the home state, as

⁷This situation would be similar to the one described by Mylonas (2012, pp. 26–32) of a “non-core group” in a state that is in conflict with the “external power” that supports it.

stated in **H1**:

- **H1.** Ethnic identity will predict bias, with the dominant ethnic group expressing bias for the home state and the minority ethnic group expressing bias in favor of the external adversary with which it has ties.

If **H1** is correct, then we would expect citizens of Ukraine who identify as ethnic Russians to express pro-Russia attitudes, while those who identify as ethnic Ukrainians would be pro-Ukraine.

The logic that leads to **H1** above assumes that the conflict will fracture society along ethnic lines, but this need not be the case. War with a foreign enemy can minimize the role of ethnic identity relative to state identity, producing a nationalism that is based on the state and is necessarily more civic in nature than ethnic.⁸ Indeed, drawing on the European political development literature Herbst (1990, p. 122) argues, “the threat of a palpable external threat may be the strongest way to generate a common association between the state and the population.”

In such a situation, all ethnic groups express a bias in favor of the home state, but ethnic minorities who are co-ethnics with the majority group in the adversary state may feel less antipathy towards the adversary than would individuals of other ethnic groups. The relatively less pronounced feelings of antipathy towards the aggressor would result in a less pronounced pro-home state bias. Thus,

- **H2.** Ethnic identity will predict variation in the degree of bias in favor of the home state.

If **H2** is correct, then we would expect both ethnic Russian and ethnic Ukrainian residents of Ukraine to express pro-Ukraine attitudes, although at different levels; those who identify as ethnic Russians would have a weaker bias in favor of Ukraine over Russia.

⁸For instance, Saideman and Ayres (2008) demonstrate how the external threat of Soviet invasion in 1956 resulted in increased pro-Hungary feelings, which helped unify the multi-ethnic state and inculcate a strong new Hungarian identity.

Contrary to group-level theories of identity, others have focused on physical proximity to and experience with wartime violence as the key factor in explaining a range of attitudes and behaviors during and following conflict. Gibler et al. (2012) suggest that individuals in the state that is attacked will identify more with that state as they need its protection, but that this effect is stronger closer to the conflict zone. Similarly, Tellez (2019) shows how those living closer to a conflict zone express different attitudes towards the conflict than those who are farther away. While a different kind of international security situation, Cortina (2020) shows that attitudes towards a highly politicized issue—the proposed border wall between the U.S. and Mexico—are a function of distance from the border.

If proximity to conflict can affect identification with the state, attitudes about the conflict, and policy preferences, then we would expect individuals in areas closer to fighting to express attitudes about the belligerents that are different from those in areas farther away. There are many ways that this could play out. Individuals who are closer to the conflict could overstate their loyalty to the home state in an attempt to demonstrate their allegiance and win protection from possible attack. On the other hand, if they foresee being caught up in the conflict, individuals may hedge their responses by downplaying their preference for one state over the other, trying to stake out a neutral position. Alternatively, if they view a takeover by the adversary state as likely they may express a preference for the adversary over the home state. While the exact nature of the difference is difficult to predict, what is consistent across all these scenarios is that those who are closer to the fighting are likely to express views towards the home state and the belligerent that are at odds with respondents who are more removed from the conflict:

- **H3** Greater proximity to the conflict will predict different levels of bias towards the home state and the aggressor.

Coming up with an *a priori* distance within which attitudes will be swayed by risk of exposure to conflict is difficult. Bakke et al. (2009), for instance, use a 50km threshold in their operationalization of proximity to conflict in the North Caucasus, but that was due to

an analysis of the geography of violence in that case. As we have only four cities in which we ran the IAT, taking a similar approach and finding the appropriate cutoff on a continuous distance measure is not feasible. Instead, looking at the four cities in which we ran the IAT, our **H3** expects that respondents in Kharkiv would hold the most conflict-affected attitudes in their views of Ukraine and Russia than respondents of the three other cities (Kyiv, Kherson, and Odesa). We believe that, at the time that the study was in the field, Kharkiv was at greater risk than the other cities of becoming embroiled in the conflict due to its geographic proximity to active or potential conflict and the violence that had already occurred within the city. We acknowledge that proximity to the conflict is not the only dimension along which Kharkiv is different from the other cities, making it an imperfect measurement. To address this problem, as well as the subjective nature of these types of groupings, we also consider an alternative operationalization of **H3**, in which Kharkiv, Kherson, and Odesa are considered more proximate to the conflict, and attitudes in those cities are contrasted to those in Kyiv (see Appendix 4.A).

4.3.1 Implicit association and explicit attitudes

Traditional surveys evaluate respondents' explicit attitudes, which are the product of a psychological process that includes the interaction of both affect and cognition. As a result, it can be difficult to tell whether a respondent's answer to a question is the result of a process of reasoning and deliberation by which they arrive at an answer (cognition) or a strong underlying feeling (affect). These underlying feelings—implicit associations—are worth studying and understanding because they have been shown to predict important political behavior and attitudes, including party choice and turnout (Arcuri et al., 2008; Ryan, 2017), attitudes towards female politicians (Beaman et al., 2009; Mo, 2015), and immigration policy preferences (Pérez, 2010), to name just a few. In some cases, implicit attitudes are more accurate predictors of behavior than explicit attitudes (Roccatto & Zogmaister, 2010). Studying implicit attitudes is particularly important in sensitive situations or when seeking to under-

stand how individuals relate to taboo topics because it is difficult for respondents to fake their answers—social desirability bias and other forms of preference falsification do not affect implicit associations.

We follow Truex and Tavana (2019), who studied attitudes towards political leaders in an authoritarian setting, by comparing implicit and explicit biases to understand public opinion about an ongoing military conflict. We use the associative-propositional evaluation (APE) model (Gawronski & Bodenhausen, 2006) of attitude formation to understand the progression from implicit associations to explicit attitudes. According to this model, implicit associations are automatic responses to a stimulus. Whether this response is negative or positive is a function of the associations that the stimulus generates within a respondent. In forming explicit attitudes, respondents evaluate their implicit associations in light of propositions, i.e., statements that have truth-value for each respondent. If the propositions lead to an explicit attitude that is consistent with the association, e.g., both are positive, then there is *congruence* between the implicit and the explicit. Previous work has demonstrated that we should generally expect to find congruence (Nosek et al., 2002), but the process of evaluating the association with propositional reasoning may lead some respondents to state an explicit attitude that is at odds with their underlying, implicit association, a phenomenon referred to as *dissociation*.⁹

The examination of respondents' implicit bias toward Ukraine or Russia will reflect their automatic, underlying feelings, while comparing their explicit attitudes towards the two countries will filter those implicit association through a process of cognitive reasoning. As there are two possible outcomes each for implicit associations and explicit attitudes, there are a total of four potential outcomes:

- **Pro-Ukraine congruence:** the respondent is implicitly and explicitly pro-Ukraine.
- **Pro-Russia congruence:** the respondent is implicitly and explicitly pro-Russia.

⁹Neither implicit associations nor explicit attitudes should be thought of as necessarily more fixed or as true beliefs, and changes in the relevance of associations or propositions can cause shifts in both implicit and explicit biases (Gawronski & Bodenhausen, 2006).

- **Pro-Ukraine dissociation:** the respondent is implicitly pro-Russia but is explicitly pro-Ukraine or neutral. This type of dissociation is what one would most expect to find if the conflict setting is leading people to engage in a cognitive process that reverses their implicitly pro-Russia orientation.
- **Pro-Russia dissociation:** the respondent is implicitly pro-Ukraine but is explicitly pro-Russia or neutral. We expect that this will be quite rare in light of the fact that the ongoing conflict should encourage and support favorable assessment of Ukraine, but not of Russia.

It is important to keep in mind that dissociation is not necessarily the same thing as social desirability bias, in which respondents misrepresent their *explicit* attitudes or behaviors to avoid reprisal or some social sanction. When preference falsification may be a problem,¹⁰ scholars have employed sensitive survey methods such as list experiments or randomized response to remove the effect of social desirability bias on understanding corrupt practices (Brierley, 2020; Malesky et al., 2015), support for the leader in authoritarian regimes (Frye, 2015), or attitudes towards militant groups in conflict settings (Blair et al., 2014; Lyall et al., 2013). These techniques allow respondents to state an *explicit* view that they would otherwise not feel comfortable sharing. While these methods allow researchers greater confidence that they have elicited a truthful expression of an explicit view, they provide no information about how respondents may feel about a given subject on an automatic or reflexive level, i.e., their implicit association towards a given subject. These methods cannot detect dissociation—the divergence of implicit association and explicit bias—which may still occur even if respondents are truthfully stating their explicit attitudes.

4.4 Research Design

After conducting a pilot in January of 2015 in Kharkiv, we ran the IATs in April 2015 in four cities: Kyiv, Kharkiv, Kherson, and Odesa (See Figure 4.2.2). These IATs, as well as nationally representative survey data used in robustness checks, were part of a research

¹⁰Kuran (1991) provides a theoretical account of this process.

effort that was commissioned and funded by the National Democratic Institute. The Kyiv International Institute of Sociology (KIIS) carried out the data collection.

4.4.1 Subject Recruitment

In each of the four cities, we recruited 150 subjects to take part in our lab-in-the-field from pre-specified sampling locations throughout each city. All subjects were told that they would be given a test about their attitudes and we reported the implicit score to them at the end of the test. A pre-test had shown that we were able to recruit a sufficiently wide range of individuals who self-identified with Russian or Ukrainian ethnicity, but we were concerned about our ability to recruit a sufficient number of individuals from across the political spectrum (for descriptive statistics on our respondents by city, see Appendix 4.B). To address this potential problem, we used quotas to allocate the sample within each city, recruiting on reported voting behavior from the 2014 parliamentary election (abstainers, pro-Russia voters,¹¹ and anti-Russia voters), age, and gender. Allocating quotas in this way only omitted voters who supported far-right parties in the 2014 elections, a small fraction of the adult population in our cities of interest.

Within each city, we recruited approximately 50 subjects who had voted for pro-Russia parties, 50 who had voted for anti-Russia parties, and 50 who had abstained. Our rationale for this recruitment strategy was that pro-Russia voters might be more likely to explicitly or implicitly favor Russia and, in the case of having pro-Russia implicit attitudes, would be most likely to dissociate. Including abstainers was particularly important because those who did not vote could be systematically different from those who did. While quota sampling can introduce bias, no laboratory sample is representative, and we wanted the ability to analyze

¹¹See Appendix 4.C for party classification. Abdelal (2005, p. 116) describes the parties these voters support as “not anti-Russian.” We use the term “pro-Russia” because these parties are more in favor than other parties of working to normalize relations and deepen ties with Russia. Voters may support these parties for many other reasons besides their stance towards Russia, such as Soviet nostalgia or support of individual local politicians. Giuliano (2018) and Toal (2017) further discuss the motivations of “pro-Russia” party voters in the context of the ongoing conflict in eastern Ukraine.

different patterns from different sample populations.¹²

4.4.2 Measuring implicit and explicit attitudes

The IAT has a long track record of uncovering respondents' implicit biases, and it has several qualities that aid researchers in studying these biases. The IAT has strong test-retest reliability (Lane et al., 2007), it is hard for respondents to manipulate (Greenwald et al., 2003, 2009), and its measures of implicit attitudes have been shown in multiple contexts to be strong predictors of political behavior (Arcuri et al., 2008; Pérez, 2010; Roccoato & Zogmaister, 2010). Following Greenwald et al. (2003), we measure IAT respondents' explicit bias as the difference ($Ukraine_i - Russia_i$) between their stated responses to two 7-point thermometer questions, which asked about their attitudes toward Ukraine and Russia.¹³ The assumption in measuring differences is that the difference in preference is what matters, not the overall level of the preference. Therefore, a respondent who rated Russia 1 and Ukraine 2 would have a pro-Ukraine explicit preference score of 1, as would a respondent who ranked Russia 5 and Ukraine 6.

IATs were used to measure respondents' implicit bias towards either Ukraine or Russia.¹⁴ These computer-based tests prompt respondents to associate words with a given *category* (Ukraine or Russia in our case) or a given *attribute* (, Positive or Negative in our case). In the test, a word that is associated either with a category or an attribute is displayed in the middle of the screen, while the corresponding categories or attributes are displayed in the upper left and upper right corners. Using the 'E' and 'I' keys on the keyboard, the respondent is asked to associate the word in the middle of the screen with the relevant category or attribute as quickly as possible.¹⁵ The computer tracks the time that it takes a respondent to perform

¹²We opted for a lab environment instead of online samples because the lab setting allowed us to explain and debrief to subjects the study in more detail, which we deemed important. The lab environment also delivered near perfect compliance with the IAT protocol.

¹³The questions read: "How do you feel about the following countries?" Respondents were then asked about several countries including Ukraine and Russia. The scale ranged from "Strongly negatively" (-3) to "Strongly positively" (3).

¹⁴The IAT software *Implicit Millisecond* was used to carry out the tests.

¹⁵The computer keyboard used in the lab had both English and Cyrillic letters pasted on the keys.

each such association task, generating a metric referred to as the *response latency*.

The validity of the IAT comes from the fact that if a respondent does not associate a category (e.g., Russia) with the attribute (e.g., Negative) listed in the same corner of the screen, then the respondent will be much slower in choosing the side of the screen to which the word in the middle belongs. For example, respondents will be slower to associate positive words (such as “love”) from the middle of the screen when a country they dislike is paired with the Positive category. Each respondent’s implicit bias towards either Ukraine or Russia is the standardized difference (IAT *d*-score) between that respondent’s response latencies on blocks where the negative attribute is paired with Ukraine and positive attribute with Russia, and blocks where the negative attribute is paired with Russia and positive attribute with Ukraine.¹⁶

Our IAT proceeded as follows. First, respondents completed training rounds on words that are associated with our two categories: Ukraine and Russia. Next, they completed training rounds on words associated with our two attributes: positive and negative. After the training rounds, respondents proceeded to the main blocks of the test. In a first set of blocks, each of the two categories (Ukraine and Russia) were shown to respondents on either on the upper left or right hand corners of the screen, next to one of the two attributes (Negative or Positive). In a second set of blocks, the category associated with the attribute is reversed: if Negative was shown next to Russia and Positive next to Ukraine in the first set of blocks, then Negative was shown next to Ukraine and Positive next to Russia in the next set of blocks.

A sample of the IAT screen is shown in Figure 4.4.1. In all blocks with both categories and attributes, categories are displayed in white at the very top of the screen, with attributes immediately beneath them (the test computer displayed attribute text in green). In the example shown, the Ukraine category is paired with the “Positive” attribute on the left, and the “Russia” category is paired with “Negative” on the right. The word that the respondent must

¹⁶Appendix 4.D explains how the *d*-score is calculated and lists the words used in the IAT.

match with the corresponding category or attribute appears in the middle of the screen. In this example, it is a negative attribute (“Crime”), so it appears in green. To answer correctly, respondents need to associate “Crime” with “Ukraine or Negative.” English translations are shown in Figure 4.4.1 in white boxes, but they were not present in the actual test.

Respondents took the IAT in Russian because we wanted to standardize the language in which subjects took the test. However, we acknowledge that this is a potential source of bias because it could exclude some people from sampling if they do not read Russian. We discuss this possible source of bias in our discussion of robustness in the results section.

TEST TEST TEST

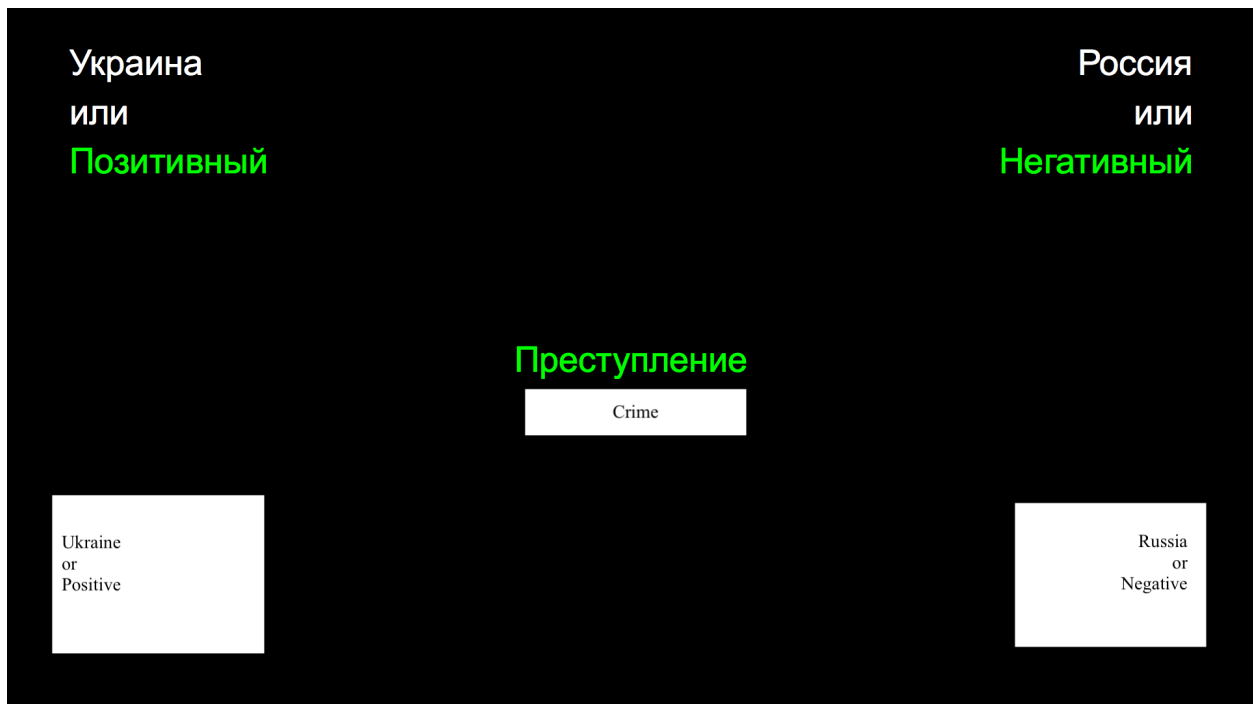


Figure 4.4.1: IAT screen example

We construct both explicit and implicit scales such that values greater than zero imply preference for Ukraine, while values less than zero imply preference for Russia. We report summary statistics (see Table 4.5.1) for the explicit scores in the units of the difference in Likert scales (integers in [-6, 6]) and for the implicit bias in *d*-scores, a standardized measurement of implicit bias. In both cases, greater than zero implies a pro-Ukraine bias, but the scales are not directly comparable to each other.

4.4.3 Operationalizing predictors and other control variables

To test the role of **ethnicity** (H1 and H2), we ask respondents to list the ethnic group (*etnicheskaia grupp*a) or groups to which they feel they belong, allowing respondents to list multiple ethnic groups.¹⁷ We use two measures of ethnicity in our analysis. Our first approach groups anyone who lists Russian as a response in the Russian category and places all others in the non-Russian category. Our second approach, and the model specification results on which we focus, separates out four groups: those who identified as only Ukrainian, those who identified as only Russian; those who identified as both Russian and Ukrainian; and all other respondents.¹⁸ By holding the lab in four different cities, our research design allows for a test of **proximity to conflict** (H3) by including controls for city in the statistical model. Due to findings that ethnicity can have differential effects in different regions of Ukraine (Erllich & Garner, 2020), we interact our city variable with our four-category measure of ethnicity in our main model specification.

Covariates other than ethnic identity and proximity to conflict may also predict implicit or explicit bias. To address this, we control for economic, social, and political variables that may be correlated with our outcomes of interest. We include controls for whether respondents are **unemployed**, because unemployment could drive less connectedness or

¹⁷The exact wording of the open-ended question was: “Many people identify with one or more ethnic groups. Please state the ethnic group or groups that you identify with.”

¹⁸This second measurement strategy emulates the new approach taken by the United States census and allows us to determine whether those who identify with overlapping identities are different from those who do not, mirroring a concern of Onuch and Hale (2018, p. 91) in Ukraine.

greater dissatisfaction with Ukraine; whether a respondent has **family in Russia**. We also include controls for **age**, **gender**, and whether the respondent's **home language** was Russian or Ukrainian, as Onuch and Hale (2018), and Kulyk (2011) show that reported ethnic identity is not the only relevant or salient component of ethnicity in Ukraine, with language playing a meaningful and distinct role. For the model of explicit bias we also control for **voting behavior** in the 2014 parliamentary election (on which we recruited and which is an important residual category of sentiment).

In our statistical models of implicit and explicit bias, we regress these outcome variables on the predictors and control variables using OLS regression with robust standard errors, although there are two important differences between the models. First, our model of explicit bias controls for respondents' implicit association scores, but the opposite is not true because in the APE model the explicit attitude is formed only after the implicit association has been made. Second, and for largely the same reason, we do not include voting behavior in the model that predicts implicit bias because we believe voting behavior is a product of implicit bias, not the other way around.¹⁹

4.5 Results

In our combined sample, the average respondent expressed a pro-Ukraine explicit attitude and a pro-Ukraine implicit bias. Moreover, in each of the four ethnic identity categories, the average respondent was pro-Ukraine on both explicit and implicit measures (see Table 4.5.1). That said, the mean values for explicit and implicit bias among those who identify as ethnic Russian are not greater than zero at standard levels of statistical significance, indicating that ethnic Russian identity may be an important predictor of relative levels of bias. There were differences in the mean levels of bias from city to city, too, although all were on average pro-Ukraine (see Appendix 4.B.1). The explicit bias measure ranged from an average of

¹⁹Results are substantively unchanged in alternate specifications that include voting behavior in the model that predicts implicit bias.

+1.4 in Odesa to +2.6 in Kyiv. This variation was driven mostly by assessment of Russia, as explicit attitudes towards Ukraine ranged from an average of 5.4 in Kharkhiv to 5.7 in Kyiv, whereas differences in average explicit attitudes towards Russia were wider, ranging from 3.2 in Kyiv to 4.1 in Odesa. These initial summary statistics run counter to H1, but they provide suggestive evidence to support H2, that in the context of a wartime rally, ethnic identity predicts relative levels of preference for the home state over the adversary. They also suggest that, as predicted by H3, proximity to conflict can predict bias.

ethnicity	Pro-Ukraine Explicit					Pro-Ukraine Implicit				
	mean	SD	min	max	n	mean	SD	min	max	n
Ukrainian only	2.42*	2.22	-5.00	6.00	386	0.53*	0.39	-0.73	1.64	386
other	1.63*	2.53	-5.00	6.00	54	0.28*	0.45	-0.84	1.05	54
Ukr and Rus	0.91*	2.36	-6.00	6.00	80	0.26*	0.45	-0.87	1.32	80
Russian only	0.12	2.43	-6.00	6.00	80	0.12	0.56	-1.15	1.33	80
All groups	1.84*	2.44	-6.00	6.00	600	0.42*	0.46	-1.15	1.64	600

Table 4.5.1: Summary statistics of explicit and implicit preferences by ethnicity, with values greater than zero indicating pro-Ukraine bias. Means of the explicit measure are of the full range of possible values (i.e., integers in $[-6, 6]$), while the means of the implicit measure are calculated from the d-score units (i.e., real values in $[-1.15, 1.64]$). * indicates mean is distinct from 0 with 95% confidence.

4.5.1 Regression analysis

Our main empirical models provide evidence that is largely consistent with the findings above (see models 5 and 6 in Appendix Table ??). The top-left and top-right panels of Figure 4.5.1 present the model-based predicted values for those respondents who self-identify as ethnic Russians only (see Appendix Figure 4.E.1 for a complete set of predicted probabilities by ethnicity and city). After controlling for relevant covariates, those who identify as ethnic Russians are consistently pro-Ukraine (both explicitly and implicitly), although the predictions are not always statistically distinct from zero at the 95% confidence level. This finding is a clear rejection of H1; on average, ethnic Russians neither explicitly nor implicitly favor Russia over Ukraine. While we are unable to say that pro-Ukraine views were caused by the

conflict in the east, we can say that in the context of a wartime rally, those who identify with the ethnic minority with ties to the aggressor are still, on average, biased in favor of the home state on both an implicit and explicit level.

The bottom two panels of Figure 4.5.1 show the first difference associated with a change in ethnic identity from Ukrainian to Russian to test whether ethnic identity may predict meaningful differences in pro-home state bias, on average, per H2. In all cities, and for both explicit and implicit biases, modeling a change from ethnic Ukrainian to ethnic Russian identity leads to less of a pro-Ukraine bias, although not always at statistically significant levels. Thus, even after controlling for theoretically relevant factors, it appears that ethnic Russian identity is associated with a weaker bias in support of the home state over the adversary.²⁰ Taken together, the model-based predicted values and first differences for ethnic Russian identity provide clear support for **H2**. Ethnic Russian respondents are not pro-Russian, but they are less pro-Ukraine than ethnic Ukrainians.

Finally, Figure 4.5.1 allows us to examine whether Kharkiv, the city most proximate to the conflict, is systematically different from those cities farther away, per H3. While there are differences between the cities, no clear pattern emerges. In Kyiv and Kherson there is a large and statistically significant difference between the explicit biases of ethnic Ukrainians and ethnic Russians, with no statistically significant difference between these groups in Odesa or Kharkiv. However, the pattern does not hold for implicit bias. In Kyiv, there is no statistically significant difference in implicit bias between those who identify as Ukrainian and those who identify as Russian, while in Odesa, Kherson, and Kharkiv there is a larger and statistically significant difference. These findings show no consistent difference between Kharkiv and the other cities, and thus no clear association between proximity to conflict and whether respondents are more or less pro-Ukraine in their explicit or implicit biases.

²⁰As a robustness check, we conduct a coarsened exact matching (CEM) exercise to verify that there is a relationship between Russian ethnicity and lower levels of pro-Ukraine bias.

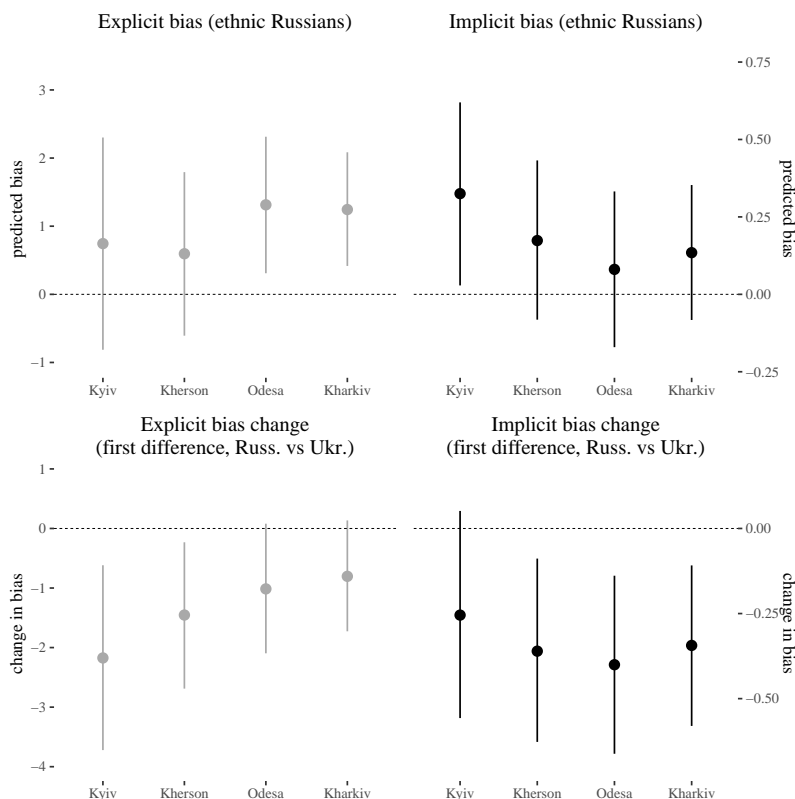


Figure 4.5.1: Predicted levels of bias and first differences from main statistical models (see Appendix Table ??). Implicit bias is measured in d -score units, while the predicted values for explicit bias reflect the range of possible values in $[-6, 6]$.

4.5.2 Dissociation in implicit and explicit attitudes

Comparing individual respondents' implicit and explicit biases reveals that some respondents are stating explicit attitudes that differ from their implicit biases. Pooling all respondents, explicit and implicit attitudes were correlated at $\rho = 0.33$ (see Figure 4.5.2), a level that indicates a reasonably high degree of attitude dissociation. Slightly more than one in three respondents (37%) dissociated. To put these figures in perspective, we compare them to the correlation of explicit and implicit attitudes in a very permissive context and a relatively repressive context. Explicit and implicit attitudes towards United States President George W. Bush during the 2004 presidential race were correlated at $\rho = 0.73$ (Greenwald et al., 2006), while research conducted in 2016 showed Egyptians' attitudes towards President Abdel Fattah El-Sisi were correlated at $\rho = 0.17$ (Truex & Tavana, 2019). The level of

dissociation we find in our study indicates that something in some respondents' process of propositional reasoning is leading to an explicitly stated preference that is at odds with the implicit association. This could be driven by internal processes or by social desirability bias or other forms of preference falsification.

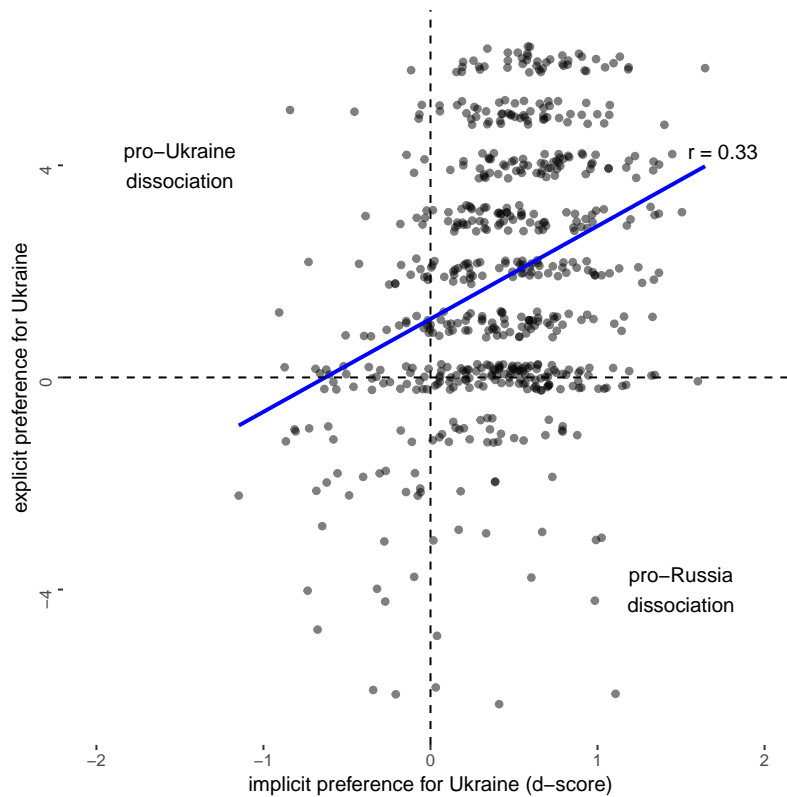


Figure 4.5.2: The majority of IAT participants expressed pro-Ukraine implicit and explicit preferences (upper right).

Pro-Ukraine dissociation was much more common than pro-Russia dissociation, consistent with our expectation that dissociating in favor of the home state would be more common during war. Out of all implicitly pro-Russia respondents ($n = 100$), 70% reported a pro-Ukraine or neutral explicit attitude, while among implicitly pro-Ukraine respondents ($n = 500$) just 30% stated a neutral or pro-Russia explicit view. This finding suggests that a conflict environment can have a differential effect on the rates of dissociation between categories. While pro-aggressor dissociation was much less common, we were surprised that so many respondents dissociated *in favor* of Russia, with 150 individuals stating neutral or

explicitly pro-Russia views despite having an implicit bias for Ukraine. These individuals included respondents from all ethnic groups, including 22% of those who identified as only Ukrainian and 32% of those who identified as only Russian, and in all cities.

Finally, an analysis of those identifying as ethnic Russian belies the notion that ethnic minorities are likely to hide pro-aggressor views. Indeed, more ethnic Russians reported a pro-Russia explicit view despite being implicitly pro-Ukraine than expressed explicit bias for Ukraine when they implicitly favor Russia. Ethnic Russians were least likely, when compared with other ethnic groups, to state a pro-Ukraine explicit view if they were implicitly pro-Russia. This finding contradicts the classical expectations about how dissociation would work in an ethnic minority group with ties to the adversary and challenges the notion that ethnic minorities are secretly part of a so-called fifth column. After controlling for theoretically relevant covariates, the relationship between Russian ethnic identity and dissociation is inconsistent across our different model specifications, with no statistically significant relationship in our fully-specified model. The coefficient associated with being implicitly pro-Russia is always large and significant, regardless of model specification (see Appendix Table 4.E.1). Further, after controlling for theoretically relevant covariates, dissociation appears more likely to occur in Kharkiv, but the relationship is never statistically significant at standard levels across our various model specifications.

4.5.3 Additional findings

In addition to the findings regarding our main hypotheses, our model reveals three additional important findings. First, those who identify as only ethnic Ukrainian express the strongest pro-Ukraine implicit and explicit biases. After controlling for theoretically relevant covariates, predicted values for ethnic Ukrainians are consistently more pro-Ukraine for both explicit and implicit biases in all four cities (see Appendix Figure 4.E.1). Second, those who claim both Russian and Ukrainian ethnic identities are more similar in their explicit and implicit biases to those respondents who claim only Russian ethnicity than to those

who claim only Ukrainian identity, particularly in Kyiv. Third, there was an unexpected finding regarding having voted for a pro-Russia political party. Supporters of pro-Russia parties expressed an explicit bias for Ukraine over Russia, although they were less explicitly pro-Ukraine than those who voted for anti-Russia parties.

4.5.4 Robustness

The quota sampling strategy is one possible vector by which we may inadvertently bias our results. To address this possibility, we conduct two analyses. First, we re-weight the data on our recruitment categories (pro-Russia voting, anti-Russia voting, and abstaining) using city-level data from Ukraine’s Central Election Commission on the outcome of the 2014 parliamentary elections. Specifically, we weight the pro-Russia vote and anti-Russia vote observations in our data to the number of votes cast in each city for pro-Russia and anti-Russia parties on the national party list component of the ballot. We weight the abstainers to the difference between votes cast and eligible voters in each city. After weighting the data, the predicted probabilities of interest are virtually unchanged (see Appendix Figure 4.F.1).²¹ Additionally, the level of dissociation that we see is only slightly different after weighting the data, falling from $\rho = 0.33$ to $\rho = 0.31$ (see Appendix Figure 4.F.2), and we find no meaningful differences in the significance or signs on our coefficients of interest (see Appendix Table 4.F.1).

Second, we compare the explicit evaluation of Russia that we collected from respondents in the lab to data from a nationally representative survey. Data from this survey show that the majority of respondents in each city where we fielded the IAT and in each ethnic group believe Russia has a negative or very negative influence on Ukraine.²² This response

²¹The sign on the predicted bias for ethnic Russians in Odesa switches from positive to negative, but the confidence interval overlaps zero.

²²The question was: “In your opinion is the influence of the following, in determining Ukraine’s future, very negative, negative, neither negative nor positive, positive, very positive?” The face-to-face survey of 5,847 residents of Ukraine carried out by KIIS in May 2015 included oversamples ($n = 600$) in three of the four cities in which we ran our lab: Kharkiv, Kyiv, and Odesa. Kherson had fewer observation ($n = 191$), leading to greater uncertainty, but we believe this sample provides a reasonable comparison.

suggests that the population we recruited, which gave Russia roughly equivalent scores on the explicit questionnaires, is not vastly different on our explicit outcome of interest than the populations of each city in which we fielded the IAT.

The language of the IAT (Russian) is another possible vector by which bias could have been introduced into our results by essentially requiring that potential respondents speak Russian, but even if there was an effect, we believe it would be quite small. In Ukraine, particularly in 2015 in the cities in which we fielded the IAT, we believe very few potential respondents would have been excluded from recruitment by this requirement. Ukrainian society is largely bilingual, and many people prefer to speak Russian in their daily lives. Data from a May 2015 survey, just one month after our data collection, corroborate this, as the vast majority of randomly selected respondents in the cities where we ran our labs answered the survey in Russian rather than Ukrainian (99.3% in Kharkiv, 100% in Kherson, 84.8% in Kyiv, and 96.0% in Odesa).²³ Still, we cannot eliminate the possibility that our results are biased by the exclusion of those who do not speak Russian or choose not to do so. However, if this were the case, we believe it would bias our estimates of attitudes in favor of Russia and against Ukraine, as we believe that not knowing Russian or refusing to speak Russian would be associated with stronger pro-Ukraine feeling. In other words, we think fielding the IAT in Russian would lead us to underestimate the amount of pro-Ukraine sentiment, if anything.

4.6 Conclusions

The wartime “rally ‘round the flag” effect has occurred in many countries around the world and across many types of conflict, from conventional wars to terrorist violence. Our research contributes to understanding public opinion against the backdrop of a rally in two ways. First, we showed that ethnic minority populations are not necessarily favorable to the country

²³Our survey measures actual language usage during the survey, rather than stated preference, which likely explains the high levels of Russian.

with which their ethnicity is associated. Second, we find that ethnicity can be an important predictor of the intensity of bias in favor of the home state for ethnic minority populations with ties to the aggressor state.

Taken together, these findings make an important contribution to the literature on ethnicity, identity, and conflict. In light of prior work showing that conflict involving multi-ethnic societies can exacerbate ethnic divisions and stoke irredentism, the pro-Ukraine implicit and explicit biases of ethnic Russians in Ukraine urge caution towards assumptions that ethnic minority populations will form so-called “fifth columns.” We interpret our findings as being consistent with work that demonstrates how identity can cohere in a country that has been the target of aggression, reinforcing a supra-ethnic national identity rather than fracturing society along ethnic lines. This interpretation is consistent with other findings that have shown renewed strength of civic identity in Ukraine in the wake of the conflict.

Methodologically, we show how an innovative application of the IAT can further both our understanding of attitudes and our confidence in the responses that we obtain through traditional surveys. We find that most respondents have congruent implicit and explicit biases, but dissociation—the divergence of implicit and explicit biases — is not uncommon, with implicit bias for the aggressor being its strongest predictor. Against the backdrop of a conflict with Russia, respondents who are implicitly pro-Russia are much more likely to state explicit biases that are at odds with their implicit biases, a finding that holds after controlling for theoretically relevant covariates. Ethnic identity and proximity have a weaker and less consistent relationship to dissociation. The weak relationship between ethnicity and dissociation further undermines the stereotypical view of ethnic minority groups with ties to the aggressor as dissembling or untrustworthy members of a so-called fifth column.

In addition to adding to what we know, our findings raise important questions that urge further study. First, more study is needed to determine a clearer relationship between proximity to conflict and attitudes about conflict. Second, what accounts for the dissociation that we see in our sample deserves further scrutiny. Follow-up studies using sensitive

survey techniques could determine whether dissociation is the result of social desirability bias or other external factors. Third, exploring whether having an implicit bias for another state over one's home state predicts dissociation across contexts merits further investigation. Replicating our study in a country that is not at war with its neighbor could help determine the degree to which this finding is associated with interstate conflict, in particular. Fourth, the mechanism which determine why some ethnic Russians—and some ethnic Ukrainians—exhibit pro-aggressor bias could also be explored further. Possible explanations could be differences in education, media consumption habits, or whether the respondent migrated from Russia to Ukraine. Fifth, whether implicit biases for the home state persist over time is another valuable question. Data on the rally effect shows that, in most cases, the effect is temporary, but there has been no research that we are aware of on how implicit bias changes over time in a society that has undergone a wartime rally. Pursuing these questions will further deepen our understanding of public opinion during wartime, along with the interplay of identity and attitudes towards conflict.

Appendix

This appendix contains supplemental material referenced in Chapter 4.

4.A Alternative grouping of cities for H3

We examine an alternative operationalization of proximity in **H3** in which Kyiv is considered distant to the conflict relative to the other three cities, which we treat as more proximate to the conflict for a number of reasons. This alternative operationalization of proximity also results in a null-effect for proximity, in line with what we find when considering Kharkiv vs. the others.

4.A.1 Justification for Kyiv vs. Other Cities comparison

At the time that the study was in the field, Kharkiv, Kherson, and Odesa could all be seen to have been at greater risk and experienced greater perceived risk than Kyiv of becoming embroiled in the conflict due to their geographic proximity to active or potential conflict. In the early period of the war, Kherson was considered to be potentially at risk because Russia's strategic goals may have included creating a land bridge to Crimea (Marten, 2015, p. 191). Odesa is close to Russian troops in both the breakaway Moldovan region Transnistria and in Crimea. Moreover, these three cities were all more economically impacted by the disruption in trade with Russia than was Kyiv, and they were also more likely to bear the brunt of the internally displaced person population.

All three cities are in regions that Russian President Vladimir Putin specifically mentioned in April 2014 as being part of the historical Russian empire and not Ukraine (Taylor, 2014), a broad swath of southeastern Ukraine claimed by Russia-backed separatists as Novorossiia.²⁴ These cities were all included in the separatists’ project to create *Novorossiia*, a pro-Russian independent state, from much of southern and eastern Ukraine. Figure 4.A.1 presents a propaganda map circulated by the anti-government insurgent group Donbass People’s Militia. This map, variants of which were widely circulated on social media, shows Kherson, Kharkiv, and Odesa as part of Novorossiia, which we argue could have created a sense of greater risk among the residents of these cities compared to those in Kyiv.



Figure 4.A.1: Russian propaganda map depicting half of Ukraine as part of “Novorossiia.”

The propaganda map above, which claims to be the work of the Donbass People’s Militia, a separatist force fighting in the east, was circulated online in late 2014. It portrays nearly half of Ukraine (the south-eastern pink part of the map), including all of Kharkiv, Odesa, and Kherson oblasts, taken over and turned into Novorossiia. Images like this were common propaganda items on social media and websites that advocated for the Russian-backed fighters in eastern Ukraine.

²⁴For more on the conflict, see (Laruelle, 2016). For a detailed analysis of Novorossiia as a “geopolitical imaginary” in the late-Soviet period, 1990s, and today, see O’Loughlin et al. (2017).

4.A.2 Expectations of conflict: Kyiv vs. Other Cities

Nationally representative survey data from May 2015 indicates that there was indeed greater perceived risk in Odesa, Kharkiv, and Kherson compared to Kyiv. The average survey respondent in these three cities was roughly one point higher than the average respondent in Kyiv on a ten-point scale measuring whether they expected to be exposed to air strikes, ground combat, and shelling over the course of the coming year (see Figure 4.A.2). These differences in means were consistently statistically significant in Kharkiv and Odesa, but not in Kherson, which may be due to the much smaller sample size in Kherson relative to the other cities.

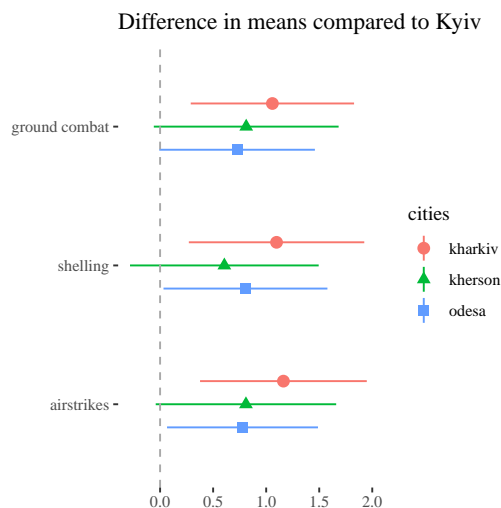


Figure 4.A.2: Differences in average expectation of conflict compared to Kyiv.

Respondents in the three cities that we classify as the East reported greater average expectation of conflict-related violence over the course of the coming year than respondents in Kyiv, with Kharkiv respondents expressing the greatest expectation of conflict. (95% survey design-based confidence intervals shown.)

4.B Summary Statistics

Table 4.B.1: Majorities of each of the four ethnic identity categories expressed an implicit and explicit preference for Ukraine (“Pro Ukraine”) or an implicit preference for Ukraine but an explicit preference for Russia (“Pro Russia disassociation”). Respondents who expressed a neutral explicit attitude (they evaluated Russia and Ukraine the same) are coded as dissociating.

category	Ukrainian		Other		Ukr & Rus		Russian		Row Total
	n	%	n	%	n	%	n	%	
Pro Ukraine	271	70	26	48	31	39	22	28	350
Pro Rus. Dissoc.	84	22	15	28	25	31	26	32	150
Pro Ukr. Dissoc.	27	7	9	17	17	21	17	21	70
Pro Russia	4	1	4	7	7	9	15	19	30
Total	386	100	54	100	80	100	80	100	600

Table 4.B.2: Majorities in each of the four cities expressed an implicit and explicit preference for Ukraine (“Pro Ukraine”) or an implicit preference for Ukraine but an explicit preference for Russia (“Pro Russia disassociation”). Respondents who expressed a neutral explicit attitude (they evaluated Russia and Ukraine the same) are coded as dissociating.

category	Kyiv		Kherson		Odesa		Kharkiv		Row Total
	n	%	n	%	n	%	n	%	
Pro Ukraine	107	71	87	58	76	51	80	53	350
Pro Rus. Dissoc.	27	18	37	25	40	27	46	30	150
Pro Ukr. Dissoc.	13	9	17	11	19	13	21	14	70
Pro Russia	3	2	8	5	15	10	4	3	30
Total	150	100	149	99	150	101	151	100	600

Table 4.B.3: Summary statistics of explicit and implicit preferences by city, with values greater than zero indicating pro-Ukraine bias.

city	Explicit					Implicit				
	mean	SD	min	max	n	mean	SD	min	max	n
Kyiv	2.56	2.56	-6	6	150	0.53	0.41	-0.74	1.40	150
Kherson	1.77	2.37	-6	6	149	0.46	0.49	-0.91	1.64	149
Odesa	1.44	2.49	-5	6	150	0.3	0.46	-1.15	1.60	150
Kharkiv	1.6	2.22	-5	6	151	0.4	0.44	-0.87	1.45	151
All cities	1.84	2.44	-6	6	600	0.42	0.46	-1.15	1.64	600

Table 4.B.4: Home language use by city.

category	Kyiv		Kherson		Odesa		Kharkiv		Row Total
	n	%	n	%	n	%	n	%	
Russian	88	59	90	60	132	88	117	77	427
Both	30	20	41	28	10	7	26	17	107
Ukrainian	32	21	17	11	8	5	7	5	64
Other	0	0	1	1	0	0	1	1	2
Total	150	100	149	100	150	100	151	100	600

Table 4.B.5: Home language use by reported ethnic identity.

language	Ukrainian eth.		Other eth.		Ukr & Rus eth.		Russian eth.		Row Total
	n	%	n	%	n	%	n	%	
Russian	248	64	41	76	61	76	77	96	427
Both	77	20	10	19	17	21	3	4	107
Ukrainian	59	15	3	6	2	2	0	0	64
Other	2	1	0	0	0	0	0	0	2
Total	386	100	54	101	80	99	80	100	600

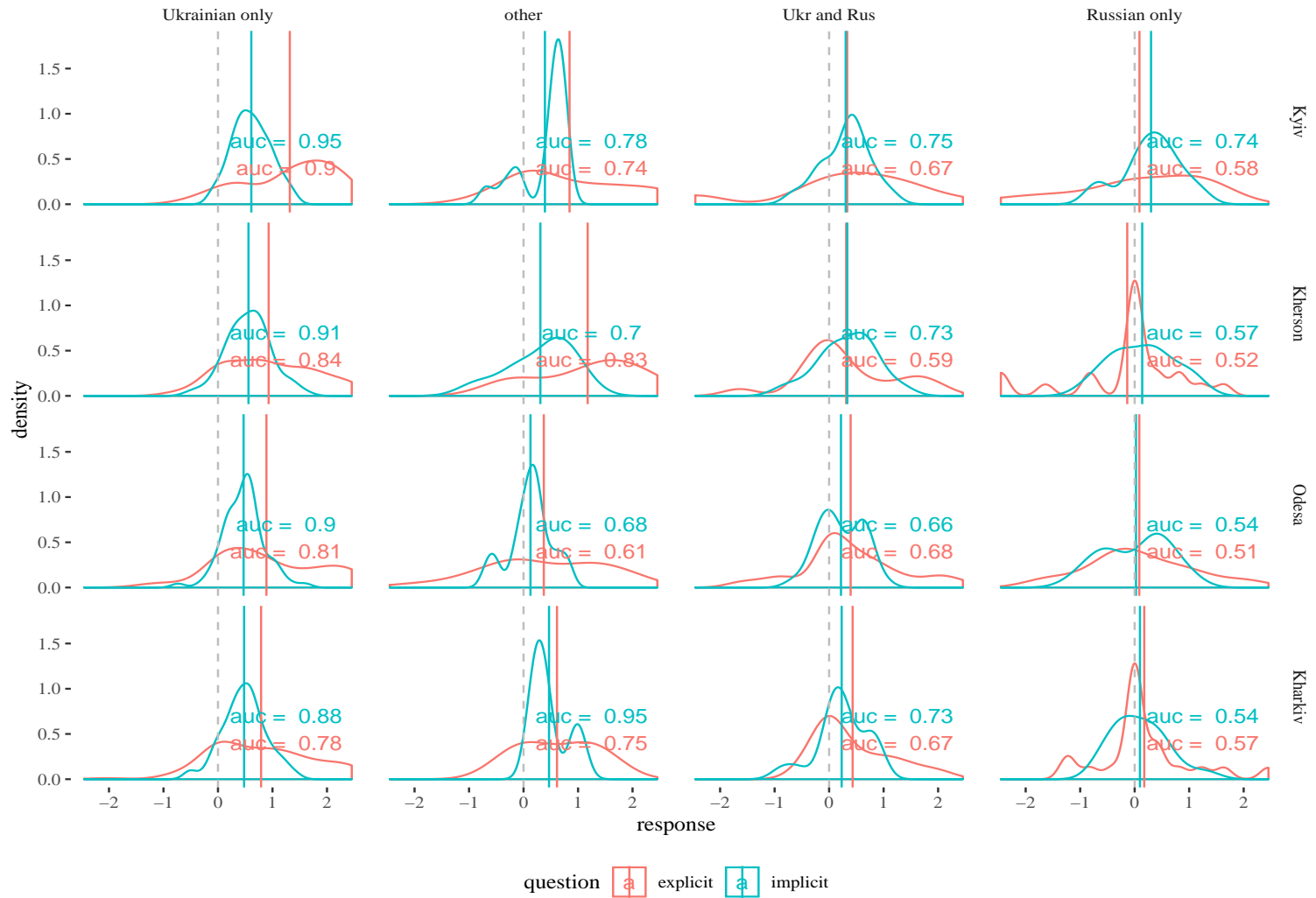


Figure 4.B.1: The density and means of responses to the explicit and implicit bias questions by ethnicity and city (both variables normalized to match the scale of the IAT measurement $[-1.6, 1.6]$ to facilitate comparison). Positive values indicate a bias toward Ukraine. In all cases, the area under the curve greater than zero is greater than 0.50, indicating a bias in favor of Ukraine.

4.C Political Party Categorization

Groups	Parties
Anti-Russia Parties	People’s Front, Bloc of Petro Poroshenko, Samopomich, Civic Position, Batkivshchyna, Radical Party
Russia-leaning Parties	Communist Party, Opposition Bloc, Strong Ukraine
Far Right	Right Sector, Svoboda

4.D IAT procedures

During the IAT data collection process participants were asked a series of questions before administering the IAT. Included in this series were questions that explicitly asked respondents to use a feeling thermometer to report their attitudes toward both Ukraine and Russia (“How do you feel about the following country: Russia/Ukraine”). The order of Russia-Ukraine was randomized to remove order effects. Respondents answered on a seven-point scale that ranged from “strongly negatively” to “strongly positively.” The differences between the responses to the Ukraine and Russia attitude questions produced the explicit bias variable. More positive values indicate stronger explicitly bias in favor of Russia, while more negative responses indicate a stronger explicit bias in favor of Ukraine.

The IAT itself required participants to use a keyboard to associate sets of terms with the qualities “good” or “bad.” To create the list of good and bad terms, we started from the list in Greenwald et al. (1998) and modified with local partners for the Ukrainian/Russian context. The list of words related to Ukraine and Russia were pre-tested in Ukraine to make sure that they were clearly associated with either Ukraine or Russia. Table 4.D.1 presents these words.

Country		Positive		Negative	
Russia	Ukraine	English	Russian	English	Russian
Kremlin (P)	Lviv (C)	Courageous	Мужественный	Crime	Преступление
Moscow (P)	Kyiv (P)	Friend	Друг	Suffering	Страдания
Balalaika (C)	Maidan (P)	Gentle	Нежный	Grief	Горе
Putin (P)	Vyshyvanka (C)	Happy	Счастливый	Hatred	Ненависть
Medvedev (P)	Carpathians (C)	Health	Здоровье	Poverty	Бедность
Urals (C)	Poroshenko (P)	Honest	Честный	Tragedy	Трагедия
Yekaterinburg (C)	Dnipropetrovsk (C)	Love	Любовь	Dirt	Грязь
Vladivostok (C)	Yatseniuk (P)	Devoted	Преданный	Unfortunate	Несчастный
-	-	Vacation	Отпуск	Disease	Болезнь

Table 4.D.1: Sets of words that IAT participants were required to associated with good/bad and the respective categories. P are political words, and C are cultural words. We found no substantive differences between analyses that used all words, just political words, or just cultural words.

Greenwald et al. (2003) suggest an updated scoring algorithm, which Nosek and Greenwald (2013) confirm better controls for age differences in response times — an important consideration in our case. This updated scoring algorithm is similar to Cohen’s d measure, but calculates the pooled standard deviation across blocks rather than within the same type of block. Critically, we use both the practice blocks and test blocks $B3, B4, B6, B7$ as suggested by the literature.

The procedure includes the following steps:

1. Remove all trials (items) for all subjects with response latencies greater than 10,000 ms and eliminate all subjects who respond to more than 10% of trials in under 300 ms.
2. Compute the mean of all of the correct trials for each block.
3. Compute the pooled standard deviation for the shorter (B3, B6) $\text{PooledSD}_{B3, B6}$ and longer blocks (B4, 7) $\text{PooledSD}_{B4, B7}$.
4. For each trial that was incorrect, add 600 ms to the mean for that block.
5. Re-estimate the mean for each block with the all correct and incorrect trials included.
6. Subtract each difference ($B_{\text{diff}1} = \overline{B6} - \overline{B3}$ and ($B_{\text{diff}2} = \overline{B7} - \overline{B4}$)).
7. Divide each difference by its associated pooled standard deviation ($B_{\text{diff}1}/\text{PooledSD}_{B3, B6}$) and ($B_{\text{diff}2}/\text{PooledSD}_{B4, B7}$).
8. Take the mean of the two quotients from the previous step.

4.E Results from statistical models

4.E.1 Regression tables

Table 4.E.1

	Dependent variables: dissociation (dummy variable)		
	(1)	(2)	(3)
Russian (dummy)	0.46*		
	(0.23)		
Eth. (<i>ref Ukr (only)</i>): Other		0.57	0.15
		(0.33)	(0.77)
: Ukr. and Rus.		0.73*	0.85
		(0.29)	(0.52)
: Russian (only)		0.45	0.83
		(0.31)	(0.58)
City (<i>ref Kharkiv</i>): Kyiv	-0.68*	-0.73*	-0.86*
	(0.26)	(0.26)	(0.32)
: Kherson	-0.34	-0.35	-0.36
	(0.25)	(0.25)	(0.30)
: Odesa	-0.46	-0.53*	-0.17
	(0.26)	(0.27)	(0.32)
Home lang. Russian (dummy)	0.26	0.27	0.26
	(0.21)	(0.21)	(0.22)
2014 Vote (<i>ref Anti-Russia</i>) : Pro-Russia	0.24	0.24	0.30
	(0.23)	(0.23)	(0.24)
: Abstain	0.20	0.18	0.25
	(0.23)	(0.23)	(0.23)
IAT	-0.08	-0.03	-0.03
	(0.30)	(0.30)	(0.30)
Implicitly pro-Russia	1.46*	1.47*	1.49*
	(0.35)	(0.36)	(0.36)
Other. eth. × Kyiv			0.83
			(0.98)
Ukr. and Rus. × Kyiv			0.19
			(0.81)
Russian (only). × Kyiv			0.24
			(0.86)
Other eth. × City: Kherson			0.53
			(1.02)
Ukr. and Rus. × Kherson			0.22
			(0.80)
Russian (only) × Kherson			-0.36
			(0.83)
Other eth. × Odesa			0.09
			(0.95)
Ukr. and Rus. × Odesa			-0.75
			(0.72)
Russian (only) × Odesa			-1.27
			(0.84)
Constant	-0.94*	-1.02*	-1.13*
	(0.34)	(0.35)	(0.37)
Demographics	YES	YES	YES
Observations	590	590	590
Log Likelihood	-346.74	-344.59	-341.48
Akaike Inf. Crit.	721.48	721.19	732.95

Note:

Robust SEs in parentheses. * $p < 0.05$.
Demographics include **family in Russia**, **female**, and **age*** (positive in models 1, 2, and 3).

4.E.2 Predicted bias

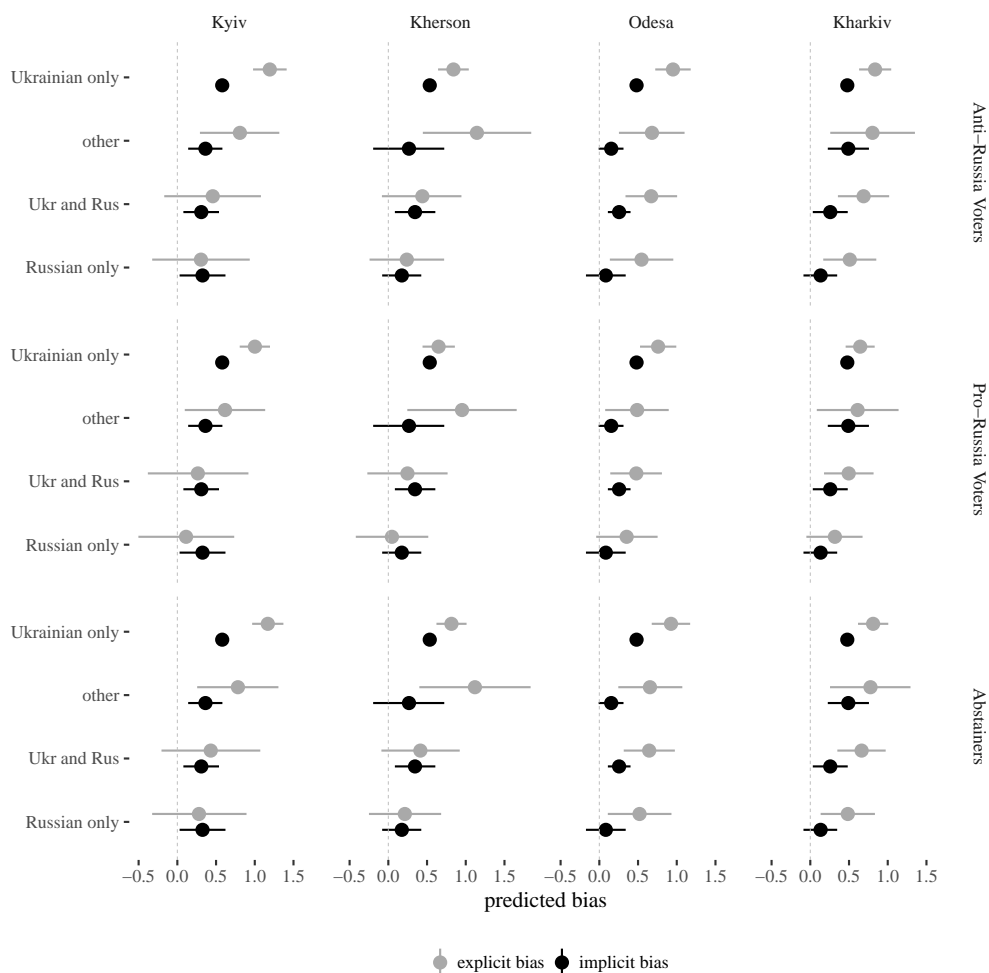


Figure 4.E.1: Predicted bias for all ethnicity, city, and recruitment categories from main empirical models.

The models predict congruent (i.e., implicit and explicit in agreement) pro-Ukraine sentiment, although sometimes at levels that are not statistically distinct from 0 at the 95% confidence level. Predicted values for explicit bias are normalized to standard deviation units to make viewing easier, but explicit and implicit values are not directly comparable.

4.F Analysis weighted by recruitment category

4.F.1 Regression tables

Table 4.F.1

	Dependent variables: explicit attitudes or IAT <i>d</i> -score					
	IAT		EXP		IAT	EXP
	(1)	(2)	(3)	(4)	(5)	(6)
Russian (dummy)	-0.30*		-0.98			
	(0.02)		(2.98)			
Eth. (<i>ref Ukrainian (only)</i>): Other		-0.18*		-0.69	-0.07*	-0.58
		(0.02)		(4.17)	(0.01)	(1.89)
: Ukr. and Rus.		-0.27*		-0.75	-0.28*	-0.48
		(0.02)		(2.73)	(0.02)	(2.04)
: Russian (only)		-0.39*		-1.54	-0.39*	-0.71
		(0.04)		(4.61)	(0.03)	(2.31)
City (<i>ref Kharkiv</i>): Kyiv	0.09*	0.10*	0.51	0.56	0.11*	0.80
	(0.02)	(0.02)	(3.49)	(3.35)	(0.02)	(3.42)
: Kherson	0.04*	0.05*	-0.28	-0.26	0.04*	-0.06
	(0.01)	(0.01)	(1.32)	(1.28)	(0.01)	(1.61)
: Odesa	-0.07*	-0.05*	0.12	0.16	-0.01	0.16
	(0.02)	(0.02)	(1.94)	(1.89)	(0.01)	(2.68)
Home lang. Russian (dummy)	-0.09*	-0.07*	-0.64	-0.55	-0.08*	-0.54
	(0.02)	(0.01)	(3.12)	(3.02)	(0.02)	(3.11)
2014 Vote (<i>ref Anti-Russia</i>): Pro-Russia			-0.36	-0.35		-0.33
			(2.71)	(2.67)		(2.91)
: Abstain			0.33	0.35		0.35
			(3.76)	(3.68)		(3.91)
IAT			1.06	0.97		0.99
			(2.23)	(2.31)		(2.68)
Other. eth. × Kyiv					-0.10	-0.23
					(0.06)	(9.48)
Ukr. and Rus. × Kyiv					-0.06	-0.46
					(0.07)	(7.90)
Russian (only). × Kyiv					0.08	-1.62
					(0.13)	(13.71)
Other eth. × City: Kherson					-0.27*	0.99
					(0.03)	(2.91)
Ukr. and Rus. × Kherson					0.16*	-0.70
					(0.02)	(2.30)
Russian (only) × Kherson					-0.004	-1.27
					(0.03)	(2.38)
Other eth. × Odesa					-0.24*	-0.01
					(0.02)	(4.08)
Ukr. and Rus. × Odesa					0.03	0.01
					(0.03)	(3.39)
Russian (only) × Odesa					-0.14	-0.41
					(0.12)	(5.20)
Constant	0.59*	0.59*	2.32	2.37	0.59*	2.20
	(0.02)	(0.02)	(3.71)	(3.63)	(0.02)	(3.63)
Demographics	YES	YES	YES	YES	YES	YES
Observations	590	590	590	590	590	590
Log Likelihood	-456.40	-449.91	-1,427.49	-1,422.71	-446.93	-1,419.08
Akaike Inf. Crit.	932.80	923.82	2,880.97	2,875.43	935.85	2,886.15

Note:

Robust SEs in parentheses. * $p < 0.05$.Demographics include **age**, **female**, and**family in Russia*** (negative in models 3, 4, and 6).

Observations weighted by raking to voting group.

4.F.2 Predicted bias (weighted data)

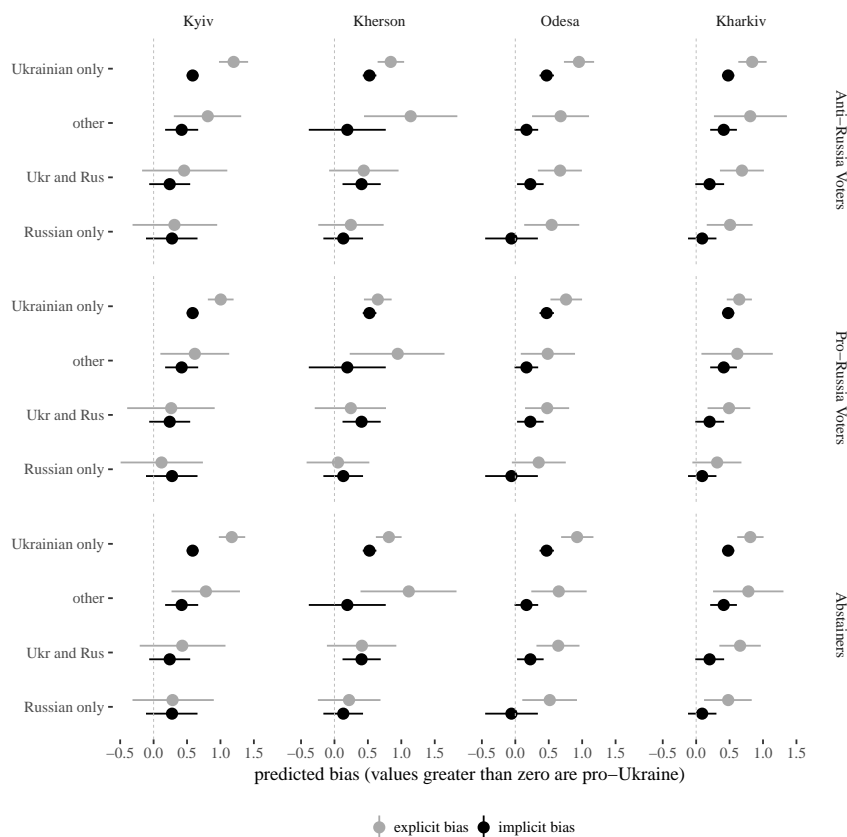


Figure 4.F.1: Predicted preference from models interacting city and ethnicity, weighted using survey data on frequency of recruitment groups in the population of each city.

Both explicit and implicit models predict pro-Ukraine sentiment across all combinations of region and ethnicity, although sometimes at levels that are not statistically distinct from 0 at the 95% confidence level. Predicted values for explicit bias are normalized to standard deviation units to make viewing easier, but explicit and implicit values are not directly comparable. For more on calculation of implicit value scale, see Appendix 4.D.

4.F.3 Weighted IAT data

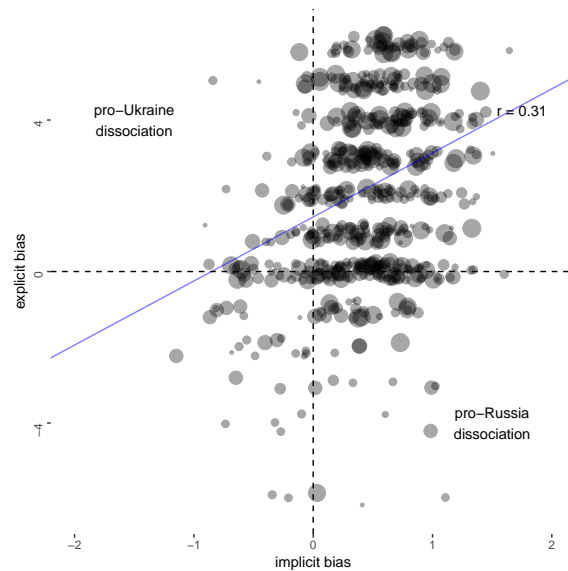


Figure 4.F.2: Explicit and Implicit preferences for Ukraine (greater than zero) or Russia (less than zero).

Size indicates weight based on nationally representative survey data of political attitudes in each of the four cities in which IATs were run. Observations were weighted according to actual population estimates as to whether they abstained in the 2014 parliamentary election, supported an opposition party, or supported the current (at the time of survey enumeration) coalition. These categories responded to those used in the recruitment quotas for the IAT.

Chapter 5

Conclusion

The three chapters presented in this dissertation have engaged with two big questions in comparative and international politics. First, in chapters two and three, I explore how we can explain the ubiquity of multi-party authoritarian regimes. Second, in chapter four, I ask how citizens react when their country is attacked, providing evidence that the “rally ’round the flag” effect applies to ethnic minorities, even those that have ties to the external aggressor state. Both of these topics are, sadly, highly relevant today. They are also questions with a rich history of theory and empirical work, but in both cases I find puzzles and unanswered questions by pushing the boundaries of existing theory and empirical research.

In chapter two I explored the shortcomings of existing formal models of authoritarian politics, which usually assume a unitary opposition, and which consider only the material returns to political engagement. I present a new model—the authoritarian dictator as a discriminating monopsonist—to address both of these shortcomings. This theory argues that by offering multiple parties, the dictator forces political elites to reveal their ideological preferences. In so doing, the dictator can offer less in rents to all members of the political elite because they derive some expressive benefit from joining parties that they agree with. This allows the dictator the ability to reduce the rents that must be sacrificed to buying off any given fraction of the elite, leaving more left over to ensure stability and weather any future

challenges. This is an important insight for the study of authoritarian politics, showing that multi-party authoritarianism may not represent a concession to the opposition, or an inherent weakness of the regime. Rather than being seen as unable to eliminate all opposition parties, the dictator in the multi-party authoritarian regime should be recognized as potentially in as strong a position as the leader of a single-party state. In addition to explaining the utility of systemic opposition parties, the model can generate predictions about elite behavior. It specifies how variation in the rents available for the dominant or systemic opposition parties, or changes in the ideological distance between dominant and opposition parties can cause opposition elites to defect to the dominant party or leave politics altogether.

My presentation of this model focuses on explaining how political actors' consideration of both material and expressive benefits leads to stable multi-partyism in authoritarian regimes. Its implications, however, are broader. The same insight about political actors joining parties can be applied to voters. In this way, and by modeling expressive benefits, the model is in conversation with research on expressive voting in democratic systems (Brennan & Lomasky, 1997), as well as studies of clientelism and vote-buying in weak democracies (Stokes, 2005; Stokes et al., 2013). Similarly, it can explain policy choices by the dictator, especially when those choices seem to outside observers to reduce expected rents. History, and indeed, the contemporary political world, is rife with examples of leaders doing things that seem irrational from an expected utility standpoint, unless we assume that their utility function includes the expressive benefits they get from taking actions that are aligned with their beliefs and values.

While not directly specified in Chapter 2, the model can be extended to make much broader predictions, including about the institutional choices of the dictator. Specifically, changes to regime resources that lower the cost of repression can be incorporated in this model as a downward shift in the Marginal Protective Value curve, which would lead to the regime buying less political participation. The model therefore provides a theoretical framework to explain the empirical finding that repressive capacity is associated with greater levels

of autocracy, even within authoritarian regimes (Albertus & Menaldo, 2012). At a certain point, this could mean that a second political party would not substantially lower the rents paid out, making a single-party state as efficient, from a rent-maximization perspective, as a multi-party one. The model, then, can begin to engage in broader conversations in comparative politics of authoritarianism about the “type” of the authoritarian regime—personalist, military, single-party, multi-party.

While a step forward, this chapter leaves some questions unanswered. For one, the model makes assumptions about the mechanisms that lead to the behavior that we see, and the language in chapter 2 largely assumes that individuals are free to choose between the dominant party and the systemic opposition. It is not clear, however, how much choice individuals actually have in moving to parties that maximize their expressive benefits. It is possible, for example, that the regime has a much heavier hand in this process than my theory implicitly assumes. The regime directing certain individuals to switch parties would not invalidate the empirical findings in chapter 3, however, as it would show the degree to which the dictator understands that ideological preferences matter in switching politicians between parties. Future work could identify these mechanisms through deeper qualitative study of party joining in authoritarian regimes. Additionally, further theoretical development could specify whether such a situation would challenge the core contribution that my theory makes, i.e., that dictators can be modeled as discriminating monopsonists, and that expressive benefits matter in authoritarian regimes.

In chapter three I take the theory developed in chapter 2, and apply it to a single case, the creation of a new, pro-regime party in Russia in 2006. The model makes predictions about which legislators in the Duma will join the new party, and which will not. Following the model’s logic, it specifies that those who would see a material benefit from joining the party and who are ideologically closest to the new party are the most likely to join. I assemble a new data set of Duma deputy roll call votes in the Fourth Duma (2003-2007) to measure deputy ideology as proximate to or distant from both their own parties and AJR. This chapter finds

modest support for the theory; shifting from low to high values on ideological distance from their own party and proximity to AJR is associated with a statistically significant increase in the probability of defecting to AJR. That said, this finding is a substantively small effect from a single case. It should be read as an example of how to test the theory as much as a finding about the theory's ability to explain multi-party authoritarianism.

This chapter also highlighted some of the challenges in taking formal models of authoritarian politics to data and generating meaningful tests. The operationalization of the two key concepts (rents and ideology) are reasonable first attempts, but both could be refined and improved upon in future studies. Additional data collection could potentially offer a way to control for individual-level preferences for rents vs. expressive benefits (the λ parameter in the model). Moreover, given the importance of informal social connections in Russian politics, an approach that modeled the dependency between deputies could offer greater insights.

Finally, chapter four turns to a long-established empirical phenomenon in international politics: the “rally ’round the flag” effect. This chapter was co-authored with Aaron Erlich and published in *International Studies Quarterly* (Erlich & Garner, 2020). This “rally effect” describes the surge of support that citizens report for their state and political leadership when their country is at war, particularly when it is attacked. We explore the rally effect in Ukraine after the start of the war with Russian-led proxies in the country's east in 2014. Ukraine presented an interesting test case in which to study the rally effect because it has a large ethnolinguistic Russian minority, allowing us to examine the degree to which there was a pro-Ukraine sentiment among parts of the population who could also identify with the aggressor state based on ethnolinguistic ties. This question had not been addressed in previous literature on the topic. However, the long and tragic history of the persecution of ethnic minorities who have ties (real or imagined) to the aggressor underscore the importance of the question. Particularly relevant to the case at hand, scholars have previously noted how ethnolinguistic Russian minority communities that were “beached” outside Russia after

the fall of the Soviet Union have been alleged to be potentially disloyal Fifth Columnists (Radnitz & Mylonas, 2022).

Studying this topic presents a major methodological challenge: survey respondents who are co-ethnics with a wartime adversary could conceivably face strong pressures to understate any positive feeling towards their ethnic homeland, as well as strong pressures to overstate their positive feelings towards the state in which they are citizens. This chapter overcomes this challenge by comparing implicit biases revealed on implicit association tests (IATs) with explicit statements of preferences for Ukraine and Russia on traditional surveys. Contrary to the fears of “fifth columns,” our findings indicate that Ukrainian citizens who self-identify as ethno-linguistically Russian 1) expressed pro-Ukraine sentiments, 2) did not harbor secretly held pro-Russia implicit biases, and 3) if anything were more likely to overstate in explicit terms their implicit preferences for Russia. The findings show that pro-state feeling found during a conflict extends beyond the ethno-linguistic lines of the titular majority population. Additionally, it offers mechanisms by which supra-ethnic identities can cohere and be strengthened in multi-ethnic polities. This innovative application of the IAT offers researchers in other areas another example to follow when confronting problems of preference falsification in survey work.

This chapter also leaves several questions unanswered, however, and future researchers could continue to investigate in this area. It is unclear, for example, whether wartime persecution by the state could reverse the pro-state feelings of ethnic minority groups. With respect to our application of an IAT to measure bias between countries, we are unaware of similar work in wartime settings or otherwise, so we are forced to speculate on what a “normal” or baseline level of preference for the home state with respect to any neighboring state might be. Indeed, there has been no research previously on implicit preferences for the home state during wartime, so future work could investigate whether the decline in the rally effect as measured by explicitly stated attitudes on survey measures is mirrored by a decline in the implicit bias for the home state. As for the study of an ethnic minorities’ preference

for a home state over an ethnic homeland, we lacked a clear baseline measure on our implicit association test, so we could not measure the degree to which different ethnic groups rally to different extents. However, the tragedy of Russia's invasion of Ukraine in 2022 could provide future researchers an opportunity to continue this work.

References

- Abdelal, R. (2005). *National Purpose in the World Economy: Post-Soviet States in Comparative Perspective*. Cornell University Press. Retrieved December 20, 2018, from <http://hdl.handle.net/2027/heb.32419>
- Albertus, M., & Menaldo, V. (2012). Coercive capacity and the prospects for democratization [Publisher: Comparative Politics, Ph.D. Programs in Political Science, City University of New York]. *Comparative Politics*, 44(2), 151–169. Retrieved April 27, 2024, from <https://www.jstor.org/stable/23211808>
- Albertus, M., & Menaldo, V. (2017, December 31). *Authoritarianism and the elite origins of democracy* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108185950>
- Anderson, B. R. (1991). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso.
- Andrews, J. T. (2009). *When majorities fail: The russian parliament, 1990-1993*. Cambridge University Press.
- Arcuri, L., Castelli, L., Galdi, S., Zogmaister, C., & Amadori, A. (2008). Predicting the vote: Implicit attitudes as predictors of the future behavior of decided and undecided voters. *Political Psychology*, 29(3), 369–387. Retrieved November 13, 2014, from <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9221.2008.00635.x/full>
- Bakke, K. M., O'Loughlin, J., & Ward, M. D. (2009). Reconciliation in Conflict-Affected Societies: Multilevel Modeling of Individual and Contextual Factors in the North

- Caucasus of Russia. *Annals of the Association of American Geographers*, 99(5), 1012–1021. <https://doi.org/10.1080/00045600903260622>
- Balcells, L. (2012). The consequences of victimization on political identities: Evidence from Spain. *Politics & Society*, 40(3), 311–347. Retrieved July 12, 2017, from <http://journals.sagepub.com/doi/abs/10.1177/0032329211424721>
- Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *The Quarterly Journal of Economics*, 124(4), 1497–1540. <https://doi.org/10.1162/qjec.2009.124.4.1497>
- Blair, G., Imai, K., & Lyall, J. (2014). Comparing and combining list and endorsement experiments: Evidence from Afghanistan. *American Journal of Political Science*, 58(4), 1043–1063. Retrieved July 17, 2017, from <http://onlinelibrary.wiley.com/doi/10.1111/ajps.12086/full>
- Blaydes, L. (2018). *State of Repression: Iraq Under Saddam Hussein*. Princeton University Press.
- Brancati, D. (2014). Democratic authoritarianism: Origins and effects. *Annual Review of Political Science*, 17(1), 313–326. <https://doi.org/10.1146/annurev-polisci-052013-115248>
- Brennan, G., & Lomasky, L. (Eds.). (1997, March 13). *Democracy and decision: The pure theory of electoral preference*. Cambridge University Press.
- Brierley, S. (2020). Unprincipled Principals: Co-opted Bureaucrats and Corruption in Ghana. *American Journal of Political Science*, 64(2), 209–222. <https://doi.org/10.1111/ajps.12495>
- Bueno de Mesquita, B., Siverson, R. M., Morrow, J. D., & Smith, A. (2003). *The logic of political survival*. MIT Press.
- Buhaug, H., Cederman, L.-E., & Gleditsch, K. S. (2014). Square Pegs in Round Holes: Inequalities, Grievances, and Civil War. *International Studies Quarterly*, 58(2), 418–431. <https://doi.org/10.1111/isqu.12068>

- Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4), 563–595. <https://doi.org/10.1093/oep/gpf064>
- Colton, T. J., & Hale, H. E. (2009). The putin vote: Presidential electorates in a hybrid regime. *Slavic Review*, 68(3), 473–503. Retrieved April 26, 2018, from <http://www.jstor.org/stable/25621652>
- Cortina, J. (2020). From a Distance: Geographic Proximity, Partisanship, and Public Attitudes toward the U.S.–Mexico Border Wall [Publisher: SAGE Publications Inc]. *Political Research Quarterly*, 73(3), 740–754. <https://doi.org/10.1177/1065912919854135>
- Cox, G. W. (1997, March 28). *Making votes count: Strategic coordination in the world's electoral systems*. Cambridge University Press.
- D'Anieri, P. (2007). *Understanding Ukrainian Politics: Power, Politics, and Institutional Design*. M. E. Sharpe. Retrieved October 30, 2018, from <http://ebookcentral.proquest.com/lib/washington/detail.action?docID=1968799>
- Davies, G. A. M. (2002). Domestic Strife and the Initiation of International Conflicts: A Directed Dyad Analysis, 1950-1982. *Journal of Conflict Resolution*, 46(5), 672–692. <https://doi.org/10.1177/002200202236169>
- Denzau, A. T., & North, D. C. (2000). Shared mental models: Ideologies and institutions. In A. Lupia, M. D. McCubbins, & S. L. Popkin (Eds.), *Elements of reason: Cognition, choice, and the bounds of rationality* (pp. 23–46). Cambridge University Press.
- Erlich, A., & Garner, C. (2020). Interpersonal incomparability in citizens' views of democracy: Survey evidence from Ukraine. *Journal of Elections, Public Opinion and Parties*, 30(2), 139–160. <https://doi.org/10.1080/17457289.2018.1554664>
- Erlich, A., & Garner, C. (2021). Subgroup differences in implicit associations and explicit attitudes during wartime. *International Studies Quarterly*, 65(2), 528–541.
- Fearon, J. D. (2004). Why Do Some Civil Wars Last So Much Longer than Others? *Journal of Peace Research*, 41(3), 275–301. <https://doi.org/10.1177/0022343304043770>

- Fournier, A. (2002). Mapping Identities: Russian Resistance to Linguistic Ukrainisation in Central and Eastern Ukraine. *Europe-Asia Studies*, 54(3), 415–433. <https://doi.org/10.1080/09668130220129542>
- Freedom in the world 2005*. (2005). Freedom House. Retrieved July 29, 2023, from https://freedomhouse.org/sites/default/files/2020-02/Freedom_in_the_World_2005_complete_book.pdf
- Frye, T. (2010). *Building states and markets after communism: The perils of polarized democracy*. Cambridge University Press. Retrieved March 23, 2024, from <https://hdl.handle.net/2027/heb32391.0001.001>
- Frye, T. (2015). What Do Voters in Ukraine Want?: A Survey Experiment on Candidate Ethnicity, Language, and Policy Orientation. *Problems of Post-Communism*, 62(5), 247–257. <https://doi.org/10.1080/10758216.2015.1026200>
- Frye, T., Gehlbach, S., Marquardt, K. L., & Reuter, O. J. (2017). Is putin’s popularity real? *Post-Soviet Affairs*, 33(1), 1–15. <https://doi.org/10.1080/1060586X.2016.1144334>
- Frye, T., Reuter, O. J., & Szakonyi, D. (2014). Political machines at work voter mobilization and electoral subversion in the workplace. *World Politics*, 66(2), 195–228. <https://doi.org/10.1017/S004388711400001X>
- Frye, T., Reuter, O. J., & Szakonyi, D. (2018). Hitting them with carrots: Voter intimidation and vote buying in russia. *British Journal of Political Science*, 1–25. <https://doi.org/10.1017/S0007123416000752>
- Gandhi, J. (2008). *Political institutions under dictatorship*. Cambridge University Press.
- Gandhi, J., & Lust-Okar, E. (2009). Elections under authoritarianism. *Annual Review of Political Science*, 12(1), 403–422. <https://doi.org/10.1146/annurev.polisci.11.060106.095434>
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132(5), 692–731. <https://doi.org/10.1037/0033-2909.132.5.692>

- Geddes, B. (1999). What do we know about democratization after twenty years? *Annual review of political science*, 2(1), 115–144. Retrieved May 24, 2016, from <http://www.annualreviews.org/doi/abs/10.1146/annurev.polisci.2.1.115>
- Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic breakdown and regime transitions: A new data set. *Perspectives on Politics*, 12(2), 313–331. <https://doi.org/10.1017/S1537592714000851>
- Gibler, D. M., Hutchison, M. L., & Miller, S. V. (2012). Individual Identity Attachments and International Conflict: The Importance of Territorial Threat. *Comparative Political Studies*, 45(12), 1655–1683. <https://doi.org/10.1177/0010414012463899>
- Gilson, J. M. (1968). Soviet elections as a measure of dissent: The missing one percent. *The American Political Science Review*, 62(3), 814–826. <https://doi.org/10.2307/1953432>
- Giuliano, E. (2015). The Social Bases of Support for Self-determination in East Ukraine. *Ethnopolitics*, 14(5), 513–522.
- Giuliano, E. (2018). Who supported separatism in Donbas? Ethnicity and popular opinion at the start of the Ukraine crisis. *Post-Soviet Affairs*, 34(2-3), 158–178.
- Goble, P. (2016). Russian national identity and the Ukrainian crisis. *Communist and Post-Communist Studies*, 49(1), 37–43. <https://doi.org/10.1016/j.postcomstud.2015.12.006>
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of personality and social psychology*, 74(6), 1464. Retrieved December 29, 2014, from <http://psycnet.apa.org/psycinfo/1998-02892-004>
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. <https://doi.org/10.1037/0022-3514.85.2.197>

- Greenwald, A. G., Nosek, B. A., & Sriram, N. (2006). Consequential validity of the implicit association test: Comment on Blanton and Jaccard (2006). *American Psychologist*, *61*, 56–61.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. <https://doi.org/10.1037/a0015575>
- Greifer, N., Worthington, S., Iacus, S., & King, G. (2023). *Clarify: Simulation-based inference for regression models*. <https://iqss.github.io/clarify/>
- Hadenius, A., & Teorell, J. (2007). Pathways from authoritarianism. *Journal of Democracy*, *18*(1), 143–157. <https://doi.org/10.1353/jod.2007.0009>
- Hale, H. (2003). Explaining machine politics in russia's regions: Economy, ethnicity, and legacy. *Post-Soviet Affairs*, *19*(3), 228–263. <https://doi.org/10.2747/1060-586X.19.3.228>
- Hale, H. (2014). *Patronal politics: Eurasian regime dynamics in comparative perspective*. Cambridge University Press.
- Hale, H. E. (2007, October 29). *Why not parties in russia?: Democracy, federalism, and the state* (1 edition). Cambridge University Press.
- Hale, H. E. (2011). Formal Constitutions in Informal Politics: Institutions and Democratization in Post-Soviet Eurasia. *World Politics*, *63*(04), 581–617. <https://doi.org/10.1017/S0043887111000189>
- Hale, H. E. (2017). Russian patronal politics beyond putin. *Daedalus*, *146*(2), 30–40. Retrieved June 8, 2017, from http://www.mitpressjournals.org/doi/abs/10.1162/DAED_a_00432
- Hale, H. E. (2018). How Crimea Pays: Media, Rallying 'Round the Flag, and Authoritarian Support. *Comparative Politics*, *50*(3), 369–391. <https://doi.org/10.5129/001041518822704953>

- Hale, H. E., & Colton, T. J. (2017). Who defects? unpacking a defection cascade from russia's dominant party 2008–12. *American Political Science Review*, *111*(2), 322–337. <https://doi.org/10.1017/S0003055416000642>
- Hamlin, A., & Jennings, C. (2011). Expressive political behaviour: Foundations, scope and implications [Publisher: Cambridge University Press]. *British Journal of Political Science*, *41*(3), 645–670. Retrieved January 13, 2024, from <https://www.jstor.org/stable/41241800>
- Herbst, J. (1990). War and the State in Africa. *International Security*, *14*(4), 117. <https://doi.org/10.2307/2538753>
- Hobbes, T. (1651). *Leviathan; or, the matter, forme, & power of a common-wealth, ecclesiasticall and civill*. London, Printed for A. Crooke. Retrieved April 28, 2024, from <http://archive.org/details/leviathan00hobba>
- Horowitz, J., & Long, J. (2016). Strategic voting, information, and ethnicity in emerging democracies: Evidence from kenya. *Electoral Studies*, *44*, 351–361. <https://doi.org/10.1016/j.electstud.2016.08.009>
- Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307–337. <https://doi.org/10.1146/annurev.psych.60.110707.163600>
- Kahn, J. (2002, June 13). Inter-governmental relations under yeltsin's new federalism. In J. Kahn (Ed.), *Federalism, democratization, and the rule of law in russia* (p. 0). Oxford University Press. <https://doi.org/10.1093/0199246998.003.0006>
- Kalyvas, S. (2006). *The Logic of Violence in Civil War*. Cambridge University Press.
ch. 6 Indiscriminate violence happens for two reasons Cost Institutional distortions
p.167 – put bookmark.
- Kam, C. D., & Ramos, J. M. (2008). Joining and Leaving the Rally Understanding the Surge and Decline in Presidential Approval Following 9/11. *Public Opinion Quarterly*, *72*(4), 619–650. <https://doi.org/10.1093/poq/nfn055>

- Karklins, R. (1986). Soviet elections revisited: Voter abstention in noncompetitive voting. *The American Political Science Review*, 80(2), 449–470. <https://doi.org/10.2307/1958268>
- Kaufmann, C. (1996). Possible and impossible solutions to ethnic civil wars. *International security*, 20(4), 136–175.
- King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science*, 44(2), 341–355.
- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163.
- Kobayashi, T., & Katagiri, A. (2018). The “rally ’round the flag effect” in territorial disputes: Experimental evidence from japan-china relations. *Journal of East Asian Studies*, 18(3), 299–319. <https://doi.org/10.1017/jea.2018.21>
- Korostelina, K. V. (2013). Mapping national identity narratives in Ukraine [Publisher: Cambridge University Press]. *Nationalities Papers*, 41(2), 293–315. <https://doi.org/10.1080/00905992.2012.747498>
- Korovkin, V., & Makarin, A. (2019). Conflict and Inter-Group Trade: Evidence from the 2014 Russia-Ukraine Crisis. *Available at SSRN 3397276*.
- Krol, G. (2017). Legislative performance of the russian state дума: The role of parliament in an authoritarian regime. *East European Politics*, 33(4), 450–471. <https://doi.org/10.1080/21599165.2017.1346504>
- Kulyk, V. (2011). Language identity, linguistic diversity and political cleavages: Evidence from Ukraine. *Nations and Nationalism*, 17(3), 627–648. <https://doi.org/10.1111/j.1469-8129.2011.00493.x>
- Kulyk, V. (2016). National identity in Ukraine: Impact of Euromaidan and the war. *Europe-Asia Studies*, 68(4), 588–608.

- Kuran, T. (1991). Now out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politics*, 44(01), 7–48. <https://doi.org/10.2307/2010422>
- Kuran, T. (1998). Ethnic Norms And Their Transformation Through Reputational Cascades. *The Journal of Legal Studies*, 27(S2), 623–659. <https://doi.org/10.1086/468038>
- Lai, B., & Reiter, D. (2005). Rally ‘Round the Union Jack? Public Opinion and the Use of Force in the United Kingdom, 1948–2001. *International Studies Quarterly*, 49(2), 255–272. <https://doi.org/10.1111/j.0020-8833.2005.00344.x>
- Lambert, A. J., Scherer, L. D., Schott, J. P., Olson, K. R., Andrews, R. K., O’Brien, T. C., & Zisser, A. R. (2010). Rally effects, threat, and attitude change: An integrative approach to understanding the role of emotion. *Journal of Personality and Social Psychology*, 98(6), 886–903. <https://doi.org/10.1037/a0019086>
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and Using the Implicit Association Test: IV What We Know (So Far) about the Method. In B. Wittenbrink & N. Schwarz (Eds.), *Implicit Measures of Attitudes* (pp. 59–102). Guilford Press.
- Lankina, T., & Watanabe, K. (2017). ‘Russian Spring’ or ‘Spring Betrayal’? The Media as a Mirror of Putin’s Evolving Strategy in Ukraine. *Europe-Asia Studies*, 69(10), 1526–1556.
- Laruelle, M. (2016). The Three Colors of Novorossiia, or the Russian Nationalist Mythmaking of the Ukrainian Crisis. *Post-Soviet Affairs*, 32(1), 55–74. <https://doi.org/10.1080/1060586X.2015.1023004>
- Lipman, M., & McFaul, M. (2001). “managed democracy” in russia: Putin and the press [Publisher: SAGE Publications]. *Harvard International Journal of Press/Politics*, 6(3), 116–127. <https://doi.org/10.1177/108118001129172260>
- Lust-Okar, E. (2005). *Structuring conflict in the arab world: Incumbents, opponents, and institutions*. Cambridge University Press.

- Lyall, J., Blair, G., & Imai, K. (2013). Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan. *American Political Science Review*, 107(04), 679–705. <https://doi.org/10.1017/S0003055413000403>
- Magaloni, B. (2008). Credible power-sharing and the longevity of authoritarian rule. *Comparative Political Studies*, 41(4), 715–741. <https://doi.org/10.1177/0010414007313124>
- Magaloni, B. (2006). *Voting for autocracy: Hegemonic party survival and its demise in Mexico*. Cambridge University Press.
- Malesky, E. J., Gueorguiev, D. D., & Jensen, N. M. (2015). Monopoly Money: Foreign Investment and Bribery in Vietnam, a Survey Experiment. *American Journal of Political Science*, 59(2), 419–439. <https://doi.org/10.1111/ajps.12126>
- Mansfield, E. D., & Mutz, D. C. (2009). Support for free trade: Self-interest, sociotropic politics, and out-group anxiety. *International Organization*, 63(3), 425. <https://doi.org/10.1017/S0020818309090158>
- March, L. (2009). Managing opposition in a hybrid regime: Just Russia and parastatal opposition. *Slavic Review*, 504–527. Retrieved July 25, 2016, from <http://www.jstor.org/stable/25621653>
- Marten, K. (2015). Putin's Choices: Explaining Russian Foreign Policy and Intervention in Ukraine. *The Washington Quarterly*, 38(2), 189–204. <https://doi.org/10.1080/0163660X.2015.1064717>
- Martin, T. (1998). The Origins of Soviet Ethnic Cleansing. *The Journal of Modern History*, 70(4), 813–861. <https://doi.org/10.1086/235168>
- McAllister, I., & White, S. (2008). Voting 'against all' in postcommunist Russia. *Europe-Asia Studies*, 60(1), 67–87. Retrieved March 28, 2018, from <http://www.jstor.org/stable/20451473>
- McFaul, M. (1997, October 30). *The election of '96* [Hoover institution]. Retrieved July 29, 2023, from <https://www.hoover.org/research/election-96>

- McNamee, L., & Zhang, A. (2019). Demographic Engineering and International Conflict: Evidence from China and the Former USSR. *International Organization*, 73, 291–327.
- Menaldo, V. (2016). *The institutions curse: Natural resources, politics, and development*. Cambridge University Press.
- Merolla, J. L., & Zechmeister, E. J. (2013). Evaluating Political Leaders in Times of Terror and Economic Threat: The Conditioning Influence of Politician Partisanship [Publisher: The University of Chicago Press]. *The Journal of Politics*, 75(3), 599–612. <https://doi.org/10.1017/S002238161300039X>
- Mo, C. H. (2015). The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters. *Political Behavior*, 37(2), 357–395. <https://doi.org/10.1007/s11109-014-9274-4>
- Mueller, J. (1973). *War, Presidents & Public Opinion*. John Wiley & Sons.
- Mylonas, H. (2012). *The politics of nation-building: Making Co-nationals, refugees, and minorities*. Cambridge University Press.
- Nichols, T. M. (2002). Putin's first two years: Democracy or authoritarianism? [Publisher: University of California Press]. *Current History*, 101(657), 307–312. Retrieved July 27, 2023, from <https://www.jstor.org/stable/45317878>
- Nosek, B. A., Banaji, M., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101–115. <https://doi.org/10.1037//1089-2699.6.1.101>
- Nosek, B. A., & Greenwald, A. G. (2013). The Implicit Association Test at age 7: A methodological and conceptual review. In M. R. Banaji & J. A. Bargh (Eds.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (pp. 265–292). Taylor & Francis.
- Ojeda, C. (2016). The Effect of 9/11 on the Heritability of Political Trust. *Political Psychology*, 37(1), 73–88. <https://doi.org/10.1111/pops.12223>

- O'Loughlin, J., Toal, G., & Kolosov, V. (2017). The rise and fall of "Novorossiia": Examining support for a separatist geopolitical imaginary in southeast Ukraine. *Post-Soviet Affairs*, 33(2), 124–144. <https://doi.org/10.1080/1060586X.2016.1146452>
- Olson, M. (1993). Dictatorship, democracy, and development. *The American Political Science Review*, 87(3), 567. <https://doi.org/10.2307/2938736>
- Onuch, O., & Hale, H. E. (2018). Capturing ethnicity: The case of Ukraine. *Post-Soviet Affairs*, 34(2-3), 84–106. <https://doi.org/10.1080/1060586X.2018.1452247>
- Pavlovsky, G. (2016). Russian politics under putin. *Foreign Affairs*, 95(3). Retrieved May 19, 2016, from <https://www.foreignaffairs.org/articles/russia-fsu/2016-04-18/russian-politics-under-putin>
- Pérez, E. O. (2010). Explicit Evidence on the Import of Implicit Attitudes: The IAT and Immigration Policy Judgments. *Political Behavior*, 32(4), 517–545. <https://doi.org/10.1007/s11109-010-9115-z>
- Perrin, A. J., & Smolek, S. J. (2009). Who trusts? Race, gender, and the September 11 rally effect among young adults. *Social Science Research*, 38(1), 134–145. <https://doi.org/10.1016/j.ssresearch.2008.09.001>
- Petersen, R. D. (2002). *Understanding ethnic violence: Fear, hatred, and resentment in twentieth-century Eastern Europe*. Cambridge University Press.
- Petrov, N. (2011). Who is running russia's regions? In V. Kononenko & A. Moshes (Eds.), *Russia as a network state: What works in russia when state institutions do not?* (pp. 81–112). Palgrave Macmillan ; Ulkopoliittinen Instituutti.
- Posen, B. R. (1993). The security dilemma and ethnic conflict. *Survival*, 35(1), 27–47. <https://doi.org/10.1080/00396339308442672>
- Puhr, R., Heinze, G., Nold, M., Lusa, L., & Geroldinger, A. (2017). Firth's logistic regression with rare events: Accurate effect estimates and predictions? [eprint: <https://onlinelibrary.wiley.com>]. *Statistics in Medicine*, 36(14), 2302–2317. <https://doi.org/10.1002/sim.7273>

- Radnitz, S. (2010). *Weapons of the wealthy: Predatory regimes and elite-led protests in central asia*. Cornell University Press. Retrieved April 27, 2024, from <http://ebookcentral.proquest.com/lib/washington/detail.action?docID=3138365>
- Radnitz, S., & Mylonas, H. (2022, June 21). Introduction: Theorizing fifth-column politics. In H. Mylonas & S. Radnitz (Eds.), *Enemies within: The global politics of fifth columns* (p. 0). Oxford University Press. <https://doi.org/10.1093/oso/9780197627938.003.0001>
- Reisinger, W. M., Miller, A. H., & Hesli, V. L. (2007). Russians and the legal system: Mass views and behaviour in the 1990s [Publisher: Taylor & Francis Group]. *The Journal of Communist Studies and Transition Politics*. <https://doi.org/10.1080/13523279708415352>
- Remington, T. (2008). Patronage and the party of power: President: Parliament relations under vladimir putin [Publisher: [Taylor & Francis, Ltd., University of Glasgow]]. *Europe-Asia Studies*, 60(6), 959–987. Retrieved March 23, 2024, from <https://www.jstor.org/stable/20451568>
- Reuter, O. J. (2017, April 30). *The origins of dominant parties: Building authoritarian institutions in post-soviet russia* [Google-Books-ID: 09IkDwAAQBAJ]. Cambridge University Press.
- Reuter, O. J., & Robertson, G. B. (2015). Legislatures, cooptation, and social protest in contemporary authoritarian regimes. *The Journal of Politics*, 77(1), 235–248. <https://doi.org/10.1086/678390>
- Roccatto, M., & Zogmaister, C. (2010). Predicting the Vote through Implicit and Explicit Attitudes: A Field Research. *Political Psychology*, 31(2), 249–274. Retrieved November 20, 2014, from <http://onlinelibrary.wiley.com.offcampus.lib.washington.edu/doi/10.1111/j.1467-9221.2009.00751.x/abstract>
- Russett, B. (1990). Economic Decline, Electoral Pressure, and the Initiation of Interstate Conflict. In C. S. Gochman & A. N. Sabrosky (Eds.), *Prisoners of War? Nation-States in the Modern Era* (pp. 123–140). Lexington Books.

- Ryan, T. J. (2017). How Do Indifferent Voters Decide? The Political Importance of Implicit Attitudes. *American Journal of Political Science*, 61(4), 892–907.
- Saideman, S. M., & Ayres, R. W. (2008). *For Kin or Country: Xenophobia, Nationalism, and War*. Columbia University Press,
- Sakwa, R. (2002). *Russian politics and society* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203465660>
- Sasse, G., & Lackner, A. (2018). War and identity: The case of the Donbas in Ukraine. *Post-Soviet Affairs*, 34(2-3), 139–157.
- Shevel, O. (2002). Nationality in Ukraine: Some Rules of Engagement: *East European Politics and Societies*, 16(2), 386–413.
- Sigelman, L., & Conover, P. J. (1981). The dynamics of presidential support during international conflict situations: The Iranian hostage crisis. *Political Behavior*, 3(4), 303–318.
- Simpser, A. (2013). *Why governments and parties manipulate elections: Theory, practice, and implications*. Cambridge University Press. Retrieved April 18, 2018, from <http://dx.doi.org/10.1017/CBO9781139343824>
- Snegovaya, M. (2023). Why russia’s democracy never began [Publisher: Johns Hopkins University Press]. *Journal of Democracy*, 34(3), 105–118. <https://doi.org/10.1353/jod.2023.a900436>
- Somer, M. (2001). Cascades of Ethnic Polarization: Lessons from Yugoslavia. *The Annals of the American Academy of Political and Social Science*, 573, 127–151. Retrieved April 18, 2018, from <http://www.jstor.org/stable/1049018>
- Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from argentina. *American Political Science Review*, 99(3), 315–325. <https://doi.org/10.1017/S0003055405051683>
- Stokes, S. C., Dunning, T., & Nazareno, M. (2013). *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge University Press.

- Svolik, M. W. (2009). Power sharing and leadership dynamics in authoritarian regimes. *American Journal of Political Science*, 53(2), 477–494. <https://doi.org/10.1111/j.1540-5907.2009.00382.x>
- Szakonyi, D. (2017). Businesspeople in elected office: Identifying private benefits from firm-level returns. *American Political Science Review*, 1–17. <https://doi.org/10.1017/S0003055417000600>
- Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. In William G. Austin & S. Worchel (Eds.), *The Social psychology of intergroup relations* (pp. 33–47). Brooks/Cole PubCo.
- Taylor, A. (2014). ‘Novorossiya,’ the latest historical concept to worry about in Ukraine [<https://www.washingtonpost.com/news/worldviews/wp/2014/04/18/understanding-novorossiya-the-latest-historical-concept-to-get-worried-about-in-ukraine/>]. Retrieved December 12, 2018, from <https://www.washingtonpost.com/news/worldviews/wp/2014/04/18/understanding-novorossiya-the-latest-historical-concept-to-get-worried-about-in-ukraine/>
- Tellez, J. F. (2019). Worlds Apart: Conflict Exposure and Preferences for Peace. *Journal of Conflict Resolution*, 63(4), 1053–1076. <https://doi.org/10.1177/0022002718775825>
- Theiler, T. (2017). The Microfoundations of Diversionary Conflict. *Security Studies*, 27(2), 318–343. <https://doi.org/10.1080/09636412.2017.1386941>
- Tilly, C. (1990). *Coercion, capital, and European states, AD 990-1990*. BBlackwell.
- Toal, G. (2017). *Near Abroad: Putin, the West and the Contest Over Ukraine and the Caucasus*. Oxford University Press.
- Toft, M. D. (2010). *The geography of ethnic violence: Identity, interests, and the indivisibility of territory*. Princeton University Press.
- Tolz, V., & Busygina, I. (1997). Regional governors and the kremlin: The ongoing battle for power [Publisher: University of California Press]. *Communist and Post-Communist*

- Studies*, 30(4), 401–426. Retrieved August 6, 2023, from <https://www.jstor.org/stable/45302045>
- Treisman, D. (2011). Presidential popularity in a hybrid regime: Russia under yeltsin and putin. *American Journal of Political Science*, 55(3), 590–609. <http://www.jstor.org/stable/23024939>
- Treisman, D. (2014). Putin's popularity since 2010: Why did support for the kremlin plunge, then stabilize? *Post-Soviet Affairs*, 30(5), 370–388. <https://doi.org/10.1080/1060586X.2014.904541>
- Truex, R., & Tavana, D. L. (2019). Implicit Attitudes toward an Authoritarian Regime. *The Journal of Politics*, 81(3), 1014–1027. <https://doi.org/10.1086/703209>
- Tyler, P. E. (2000). RUSSIANS WONDER IF PUTIN ACCEPTS LIMITS TO POWER. *The New York Times*. Retrieved July 29, 2023, from <https://www.nytimes.com/2000/08/14/world/russians-wonder-if-putin-accepts-limits-to-power.html>
- Wahman, M., Teorell, J., & Hadenius, A. (2013). Authoritarian regime types revisited: Updated data in comparative perspective. *Contemporary Politics*, 19(1), 19–34. <https://doi.org/10.1080/13569775.2013.773200>
- Waldner, D., & Lust, E. (2018). Unwelcome change: Coming to terms with democratic backsliding. *Annual Review of Political Science*, (0).
- Wallace, G. P. (2015). *Life and death in captivity: The abuse of prisoners during war*. Cornell University Press.
- Weber, M. (1946). Politics as a vocation. In H. Gerth & C. W. Mills (Eds.), *From max weber: Essays in sociology* (pp. 77–128). Oxford University Press. <http://fs2.american.edu/dfagel/www/class%20readings/weber/politicsasavocation.pdf>
- Wimmer, A. (2013). *Waves of war: Nationalism, state formation, and ethnic exclusion in the modern world*. Cambridge University Press. Retrieved August 17, 2018, from <http://dx.doi.org/10.1017/CBO9781139198318>

- Wood, E. J. (2008). The Social Processes of Civil War: The Wartime Transformation of Social Networks. *Annual Review of Political Science*, 11(1), 539–561. <https://doi.org/10.1146/annurev.polisci.8.082103.104832>
- Zhuravskaya, E., et al. (2010). Federalism in russia. *Russia After the Global Economic Crisis*. Peterson Institute for International Economics. Washington, DC, 59–77.