Summer 8-15-2018

Bankrolled Elites, Controlled State Television, and Payroll Trolls on Wikipedia: Examining Russia's Strategies in Manipulation of Opinions in the context of Ukrainian Revolution in 2014

Elena Labzina
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST.LOUIS

Graduate School of Arts & Sciences
Department of Political Science

Dissertation Examination Committee:
Norman Schofield, Chair
Michael Berchel
Sanmay Das
Justin Fox
Betsy Sinclair

Bankrolled Elites, Controlled State Television, and Payroll Trolls on Wikipedia: Examining Russia's Strategies in Manipulation of Opinions in the context of Ukrainian Revolution in 2014

by

Elena Labzina

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2018
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am grateful to my dissertation advisers, Justin Fox, Betsy Sinclair, Sanmay Das, and Norman Schofield, for being a knowledgeable source of information and advice. I am thankful to Michael Berchel for joining the committee at the very end and for his insightful comments.

I hugely thank my great friends and co-authors, Olga Chyzh and Mark Nieman, for their warmness, knowledge, and support. I learned so much from them about our profession and the field.

Summer schools and workshops I was lucky to have attended contributed a lot to my personal and academic development. I thank the organizers and participants of summer schools in University of Tampere and Central European University, EITM Institute, and Summer School in Computational Social Sciences.

Finally, I am grateful to my parents, Olya and Pavel, and my brother, Jenya, for their continuous love and support. Without them, nothing in my life could have been possible.

Elena Labzina

Washington University in Saint Louis

August 2018

Dedicated to my family.

ABSTRACT OF THE DISSERTATION

Bankrolled Elites, Controlled State Television, and Payroll Trolls on Wikipedia: Examining Russia's Strategies in Manipulation of Opinions in the context of Ukrainian Revolution in 2014

by

Elena Labzina

Doctor of Philosophy in Political Science

Washington University in St. Louis, 2018

Professor Norman Schofield, Chair

Regionally influential powers are likely to pursue not just straightforward, rational policy goals but also sophisticated long-term, possibly ideological, "milieu" goals. These objectives may be far from obvious for an external observer and often require manipulation of public and elites' opinions on both domestic and international level. In our extremely digitalized era of Big Data, opinions have to compete with the facts that is why understanding better how they can be manipulated is crucial for both pundits and practitioners. While many approaches exist to address the topic, this work examines three sound cases from modern Russia in the exceptionally politically salient context of Ukrainian revolution in 2014. This country is an excellent choice for this task being recently highly active domestically and internationally in attempting to influence opinions and attitudes. Chapter I addresses the hidden manipulation targeted at the Ukrainian elites that resulted in an unexpected victory of the street protests in Kyiv. The following chapter investigates the state-controlled TV news coverage around this time. Finally, the last part analyzes the behavior of state-political trolls on the nationally significant Russian-language Wikipedia.

Chapter 1

Introduction

Broadly, my dissertation studies how current global and regional powers may influence public and elite opinions. I perform this by looking at Russia in the context of the Ukrainian Revolution in 2014. I chose this country as a vivid example of an advanced authoritarian regime that mainly rules not by force but by persuasion of its competence through media and diplomatic channels. Hence, studying its recent opinion manipulation strategies is highly illuminating in general, and potential research conclusions extend beyond authoritarianism.

Why is "Ukrainian context" perfect to study Russia from this perspective? First, while it might be up to a discussion what Russia perceives as "its sphere of interest", there is no question in the enormous importance of Ukraine and Ukrainian politics within Russian national politics. As Brzezinski (2012) said, "...without Ukraine, Russia ceases to be an empire, but with Ukraine ... Russia automatically becomes an empire." It is possible to agree or disagree with Brzezinski's views on Russia in general, but it is impossible to avoid the theme of the empire and, hence, Ukraine, in the context of the Russian politics (Kivelson and Suny, 2016) . Curiously, the timing of both recent Ukrainian pro-liberal revolutions in 2004 and 2014 are the same with two major "tightening the screws" on freedoms in modern Russia. Furthermore, multiple journalist investigations on "Russian troll factories" including

several from New York Times (e.g MacFarquhar (2018)) revealed that their residents were required to address the topic of Ukraine in their online comments.

The chapters in the dissertation are arranged chronologically according to the time of their original completion. The first draft for the first part was completed in 2014, for the second part – in 2016, and for the last part – in 2017. Another way to perceive their order is the incremental empirical complexity. From the data-analysis view, each chapter presents a more sophisticated data methodology than the previous one. In particular, Chapter 3 and Chapter 4 employ original data sets that I collected. Each chapter has its own hallmarks and contributions that I outline below.

Chapter 2, which is published in *American Journal of Political Science*, is co-authored work with Olga Chyzh. It studies how a third-party, a major regional power, may 'bankroll' repression against protesters by influencing the local political elites. The core finding is that it may happen even at the risk of the removal of its protégé leader, with the goal of deterring future protests within its sphere of interest. This work presents a game theoretical study with the empirical applications inspired by the 2014 Ukrainian Revolution and the possible Russian involvement in its events. Despite the original focus on the recent revolution in Ukraine, the formal findings provide a general contribution to IR, and their logic is supported by the empirical evidence from Chenoweth and Stephan (2011) on protest-campaigns between 1899-2006. Apart from the collaborative discussion on all aspects of the paper, my contribution to the article was developing model and characterizing its equilibria. I also calibrated the model's parameters, so as to apply it to case of Russia and Ukraine. Finally, I brought my region specific expertise to bear on every aspect of the paper.

Chapter 3 is co-authored work with Mark David Nieman. This work looks at the relationship between state-owned news coverage and military intervention in authoritarian countries. Importantly, the state-owned television is the primary tool to influence public opinion in such political regimes. To explore changes in coverage, the work applies a Bayesian

change-point model that would show shifts in the coverage patterns corresponding to the political events. The research finds that Russian state-owned media significantly increased its coverage of Georgia and Ukraine, in the months preceding Russia's military interventions. This increased coverage was often predicated with an increased discussion of traditional Russian geopolitical rivals, such as the US. Besides collaborative discussions on the direction of the paper, first, my contribution included the discovery and justification of the data sources: the web pages of state-controlled *Russia-24* and the privately held *TVRain*. Second, I was responsible for web-scraping and data post-processing. Finally, I proposed the central method of the paper: studying the perceived salience of the topic by looking at the frequencies of proper names in the headlines of news channels.

In Chapter 4, I examine how online political astroturfing to promote its ideology and self-perceptions of the regime. This phenomenon is explored by looking at Wikipedia during the Russian-Ukrainian crisis of 2013-2016. Based upon the most recent journalist and academic investigations, the paper proposes an algorithm to identify and classify political astroturfing activity. It finds that political astroturfing in the Russian-Ukrainian context involves "edit wars" and the selective deletion of information , with much of this activity occurring on Wikipedia articles that are explicitly related to the Russian-Ukrainian crisis in 2013-2016. Finally, the discourse analysis concludes that identified instances of political astroturfing vividly illustrate the ever-present *imperial* ideology of the regime in Russia.

To sum up, from the technical perspective, parts of this dissertation present drastically different approaches to studying how a regime can influence public and elite opinions. Chapter 2 is predominantly formal, and the empirical qualitative element serves to illuminate and illustrate the key strategic insights that follow from the model. Chapter 3 presents an essential web data task along with classical statistics: the Bayesian change-point model is applied to the processed — translated and stemmed – original data in the Russian language scraped from media websites. Finally, Chapter 3 provides a mixture of modern data science

approaches that combines, following the terminology of Salganik (2017) in his recent book on "social research in the digital age", *readymades* of terabytes of wiki-data and *custommades* of geographical data on "troll factories" together with traditional discourse analysis.

Chapter 2

Bankrolling Repression? Modeling Third-Party Influence on Protests and Repression

[1]

2.1 Introduction

The 2013-2014 anti-government protests in Maidan Nezalezhnosti in central Kyiv, Ukraine, and the ensuing removal of President Yanukovich, raised much speculation within the academic and policy communities. Despite media reports of Russian involvement, the Russian government never officially acknowledged providing assistance in repressing the protesters. The general opacity of the Russian government and their notoriously poor record-keeping suggest that the details of the interactions between Yanukovich and Russia may never become

[1] Olga V. Chyzh and Elena Labzina (n.d.). "Bankrolling Repression? Modeling Third-Party Influence on Protests and Repression." In: *American Journal of Political Science* 0.0. DOI: 10.1111/ajps.12341. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12341

publicly available. As a result, even several years later, much of the international community and researchers are left with more questions than answers. What were possible incentives for Russia's involvement? What types of third parties may try to influence domestic crises of other states? What are the consequences of such influence? Does Yanukovich's removal indicate that Russia's involvement was a failure? Or did Russia gain from involvement, despite Yanukovich's removal?

There is ample evidence that similar third party attempts at influencing domestic crises are not rare. US and Soviet direct and indirect involvement in their respective spheres of interest throughout the Cold War, or Russia's interest in the outcomes of popular protests and revolutions in Eastern Europe and Central Asia, constitute only a few prominent examples. Patrick M Regan and Meachum (2014) find evidence of some level of third party involvement in approximately one third of country-years identified as "at-risk" for experiencing armed conflict in the reasonable future between 1955-2003.[2] Third parties have also provided overt government support in approximately one-third of protest campaigns (Chenoweth and Stephan, 2011) .

In many of these cases, third party involvement was expressed in the form of consultation, economic and military aid, and weapons sales. Some instances of third party involvement, however, happen in secrecy, behind closed doors. In countries with stricter records-keeping practices and greater government accountability, such as the US, some information is usually eventually revealed to the public (Forsythe, 1992) . Other cases may never move beyond the realm of speculation. Despite the pervasiveness of third party involvement, this lack of systematic data presents a considerable challenge for empirical research. We overcome this challenge by using available information from the literature and news coverage to develop

[2]Of these, third-party involvements by foreign governments make up 91% of the cases, while involvements supporting the government make up 31% of the total.

a game-theoretic model that allows for analyzing the interaction between the government, protesters, and an interested third party.

The resulting model is a generalization of the traditional two-player game between the government and the protesters through an introduction of a third player—an interested third party.[3] In our model, the onset of conflict between the government and the protesters is possible as part of a pure strategy equilibrium, in contrast to traditional two-player games that treat such conflict as either an off-equilibrium outcome or part of a mixed strategy equilibrium. The pure strategy explanation for protests and repression, provided by our model, is more intuitive than mixed-strategy equilibria, which occur under rather restrictive parameter conditions. These parameter restrictions are also somewhat incompatible with the empirical prevalence of government repression against protesters. A pure strategy explanation provided by our model, in contrast, holds for a much larger parameter space, which is more consistent with the commonality of protests and instances of state repression.

Our core finding is that third party involvement may result in repression of the protesters, and even the removal of the protégé leader in cases, which could have ended peacefully in the absence of a third party. The key to this finding is that, unlike much of the literature which attributes third parties with *neutral goals* (e.g., mediation, concern for human rights) or at least *neutral means* (e.g., avoid repression), we relax these assumptions and explore the resulting variation. Our model allows for neutral third parties, interested third parties that are averse to repression, *and* third parties that are explicitly interested in the use of repression. Building a reputation as a coercive third party, as in the latter case, may have some benefits, e.g., using repression to create a "scarecrow" for future protests. For example, the images of repression at Maidan are frequently invoked as arguments against civil disobedience within Russia and its sphere of influence (Peterson; Whitmore, 2015; 2017) . Other instances of "scarecrow" tactics include the 1956 Soviet-sponsored repression against workers protests in

[3]The two-player game is a special case of our model.

Poznań, Poland, and student demonstrations in Budapest, Hungary, as well as the massacres of the Eritrean-led opposition by the Soviet-supported Mengistu regime in Ethiopia.

Relaxing the assumption that the third party is motivated by a specific set of goals allows us to derive a set of very general predictions, some counter-intuitive insights, as well as explain a variety of third-party involvement/non-involvement scenarios, such as pressure from a "neutral" third party to avoid repression, an interested third party's decision to stay out due to anticipated costs, or a repressive third party's decision to bankroll repression for the sake of deterring future protesters. The latter scenario, in particular, has received little scholarly attention, despite its rather common occurrence, especially within the Russian sphere of interest. We use Chenoweth and Stephan (2011) data on protest campaigns, as well as some original data, to evaluate some of the predictions.

2.2 Domestic Politics and Third Party Involvement

Repression research has focused on domestic political institutions. Proponents of the domestic democratic peace argue that democratic governments are less likely to repress protesters than authoritarian regimes (Davenport; Richards, Webb, and Clay, 2007; 2015) , although repression by democracies, especially in response to violent dissent, is also not uncommon (Courtenay Ryals Conrad and Will H Moore; Davenport and Armstrong; Davenport, Armstrong, and Will H. Moore; D. W. Hill and Jones; Ritter; Ritter and Courtenay R Conrad, 2010; 2004; 2008; 2014; 2014; 2016) . Others have suggested the murder-in-the-middle hypothesis, which points to semi-democracies or semi-autocracies as the most likely hot spots for observing violent repression of protests (Pierskalla; Patrick M. Regan and Bell, 2010; 2010) .

This literature rarely focuses on the international determinants of domestic repression (Gartner and Patrick M Regan; Will H Moore, 1996; 1995) . Leaders of political regimes,

however, rarely act independent from influences of outside third parties. Minor powers often rely for support on regional or great powers (Lake; M. D. Nieman, 2009; 2016) ; in other cases, third parties claim (explicitly or implicitly) regional "spheres of influence," acting as de facto arbitrators within those areas (Lemke, 2002) . Lake (2009) , for example, argues many governments make a conscious strategic choice to enter an informal (hierarchical) relationship with a (usually like-minded) major/regional power, essentially as a form of an "insurance" arrangement: a smaller power supports the major power's international policy agenda, while the major power enhances the smaller power's economic or military security (Martinez Machain and Morgan, 2013) . The degree of smaller power's dependence on this relationship may range from a military alliance or an economic partnership to a full-on puppet regime that has no power beyond that given to it by the major power. While such language invokes images of the US, USSR/Russia, or China (and their traditional spheres of interest), one can also think of less obvious examples, such as Saudi Arabia's or Iran's roles in Yemen and Syria. Such informal arrangements between domestic governments and international third parties may have important effects on domestic policy outcomes of the protégé state.

The role of international third parties in domestic crises is better understood within the related civil war literature (Aydin; Bapat; Findley and Teo; Gent, 2012; 2006; 2006; 2008) . Involvement of a third party with an independent stake in the issue may affect conflict duration (Aydin and Patrick M Regan; Cunningham, 2012; 2010) . The probability of conflict, for instance, may decrease as a result of involvement by third parties with strong ties to one of the conflict participants (Patrick M. Regan, 2002) . Conflict outcome may also depend on whether the third party supports the government, the challenging group, or has a separate interest in the issue (Andrew H Kydd; Patrick M. Regan, 2006; 2002) .

This paper extends the literature in two ways. First, we extend research on third party involvement in civil wars to explore the effects of observed and unobserved third party influences on the onset of protests and the government's decision to use repression. Second,

unlike much of prior literature, our model relaxes the (implicit) assumption of the legitimacy or neutrality of third party's goals and means. Rather than assuming any specific goals and means, third party's broader involvement incentives are modeled via an unrestricted parameter (β), which may represent a variety of goals and considerations, such as ending conflict (Patrick M Regan, 1998) , preventing human rights violations (Gartner and Patrick M Regan, 1996) , pursuit of economic and geo-strategic goals (Findley and Teo, 2006) , or even the goal of repression for its own sake. These broader third party goals are referred to as "milieu" goals (Wolfers, 1984) , and are distinct from the policy goals that are immediately at stake in any given domestic crisis. Milieu goals may include establishing a certain reputation or a sphere of influence, promoting international law or particular economic models, etc. Milieu goals are, of course, not equivalent to normatively positive goals and depend on the outlook/aims of the leadership of the third party: thus, milieu goals of a repressive regime may include supporting/legitimizing undemocratic governments in other states. Russian President Putin's support for Assad in Syria, for example, may partially stem from his own domestic insecurities (F. Hill, 2013) .

In addition to the nature of their milieu goals, third parties also differ in their legitimacy of means, i.e. whether they consider coercion as an acceptable means for the pursuit of their milieu goals. Whereas some third parties view coercion as an acceptable if undesirable means, others may find it completely unacceptable (e.g, Gorbachev famously refused to support coercion against protests in East Germany in 1989, whereas his predecessors assisted in suppressing protests in Hungary and Poland in 1956.) We refer to third party's acceptance of coercion as a means of achieving its milieu goals using the terms "marginal" (coercion is an acceptable means) vs. "legitimate" (coercion is *not* an acceptable means) third party. An obvious complication is that the content of milieu goals is not always independent of the means that the third party may view as acceptable. A reputation as a coercive third party may have its benefits: brutality and massacres have been long used as effective tools

for deterring future challenges in authoritarian regimes. And vice versa: a third party may derive an additional benefit from pursuing its goals through non-coercive means.

Third parties pursue their policy and milieu goals within their spheres of interest using "the power of the purse": they may attempt to "sweeten" the government's policy concessions with grants or favorable-term loans, military equipment, expertise, or even personnel to back those policies (Bueno de Mesquita and Smith; Licht, 2007; 2010) . Consider, for example, the agreement between Ukraine's Yanukovich and Russia, in which Russia agreed to lower natural gas prices as well as purchase 15 billion in Ukrainian-issued bonds. Notably, the signing of this agreement coincided with Yanukovich's announcement to withdraw from negotiations of Ukraine's association status with the EU (Interfax-Ukraine, 2013) . Similar patterns are observed in Russia's efforts to legitimize the long-time rule of Belarus' Lukashenka through regular diplomatic missions, generous foreign aid, oil subsidies, and debt forgiveness (Ambrosio, 2006) .

2.3 Theoretical Model of Protests and Repression

Our game focuses on the interaction between three actors: *Government* **(G)**, *Protesters* **(P)**, and a *Third Party* **(T)**. The government is the ruling leader(s) that has the executive decision-making power in the state. The protesters are made up of activists within the society. Finally, the third party is a foreign entity, such as a major/regional power (e.g., France in West Africa) or a neighboring state. The third-party and government have the same preference ordering regarding the policy outcome, disputed by the protesters.

We assume that the government and third party are in a quid-pro-quo relationship, such that the government trades some of its policy autonomy in exchange for the third party's resources, security guarantees, or other benefits. In the extreme, such a policy dependency between the governments of asymmetrically empowered states results in a "puppet" regime in

the weaker state. A more general example of such a relationship, however, is that in which the state leader fulfills the role of a protégé of third party, i.e. has significant autonomy over her country's domestic policies, yet consults/defers to the third party on particular issues. Many developing states defer to the United States on economic policies. Similarly, many post-Soviet states, such as Belarus or Kazakhstan, while exercising relative autonomy on their domestic policies, tend to consult Russia on foreign policy.

We further assume that, while the third party and the ruling government are in agreement regarding the policy outcome on the particular issue disputed by the protesters, the broader interests of the third party and the government are not in perfect alignment (Lake; M. D. Nieman, 2009; 2016) . Specifically, we assume that while, all else equal, the third party prefers that its protégé-leader remain in power (i.e., is not overthrown by the protesters), this preference may be outweighed by third party's milieu goals. Throughout the 20th century, for example, the United States had to balance its relationships with a number of corrupt governments in Latin America and the Middle East (e.g., the Somozas of Nicaragua, Mubarak of Egypt) with its broader goals (e.g., international law, liberal economic policies).

The status quo distribution of benefits in the society privileges the government and its supporters and disadvantages the group represented by the protesters. These societal cleavages may be based on specific policy disagreements, as well as other factors (e.g., ethnicity, religion, access to resources). For the sake of parsimony, we abstract away from the precise source of these disagreements and simply assume that the protesters and government disagree over policy x. The disputed policy outcome is modeled as an indivisible zero-sum good of value normalized between 0 and 1, so that the party that obtains its preferred policy obtains the benefit of 1, while the party that does not obtain its preferred policy outcome gets the

payoff of 0.[4] The third party is assumed to possess the resources to influence the resolution of the disagreement.

The scope of protesters' demands is thus narrowed to issues that concern broad swaths of the population and/or multiple societal layers—hence the interest/involvement of a third party. Such demands may require major institutional reform (e.g., empowering a disenfranchised group), administrative change (e.g., autonomy), or a major policy orientation (e.g., 1975 Lebanese Civil War started as a movement against the pro-Syrian policy orientation of the government).[5]

The game starts with *Nature*, N, determining the type of protesters, who are *Strong* with probability α and *Weak* with probability $1-\alpha$. The protesters' type is conceptualized as *strong* or *weak* relative to the government. If repressed, *Strong* protesters (SP) will overthrow the government, while *Weak* protesters (WP) will be defeated. Protesters know their own type, but the government and third party have to form a set of beliefs about the type of protesters they are facing. After N moves, the protesters have to decide whether to challenge (C) the government or not ($\neg C$).

If the protesters do not challenge, the game ends with the *Status Quo* outcome. In this outcome, the government receives a payoff of 1, associated with implementing its preferred policy; the protesters receive a payoff of 0, as they do not get their preferred outcome; and the third party, whose preferred policy outcome aligns with that of the government, obtains the payoff of 1.

If the protesters decide to challenge and take to the streets, this sends a signal regarding their type to the third party, and the third party responds by choosing a level of assistance

[4]The Appendix discusses the consequences of relaxing this assumption. As an extension, protesters' demands may also be modeled as a function of their capacity, e.g., Thomas, Reed, and Wolford (2016) .

[5]The equilibrium analysis presented below reveals that, depending on the actions of the government and the third party, these types of demands may lead to an empirical observation of no protests (*Deterrence* equilibrium), as well as small-scale protest events, or even protest campaigns. For instance, the model makes no clear predictions of protest size for the *Accommodation* equilibrium, i.e. the government may accommodate pre-emptively, after observing a single protest event, or as a result of a protest campaign.

$k \geq 0$ that it is willing to give the government to help repress the protesters and compensate it for a possible loss of office. The game ends with the government's choice of whether to repress the protesters (R) or not ($\neg R$). Assume $r > 0$ is the cost of repressing the protests, which may include paying the internal police, buying the necessary weapons, etc. Then the total government expense on repressing equals $r - k$, as the third party pays the cost k.

The government's use of repression against *Weak* protesters results in the *Successful Repression* outcome. Third-party assistance increases the level of repression against the protesters by k (e.g, through availability of superior weapons or adding to the size of the internal police). The protesters fail to obtain their preferred policy and pay the cost of repression, which yields them a payoff of $-r - k$.[6] The third party obtains its preferred outcome minus the assistance amount k. Should the third party get involved in a domestic crisis in its protégé state, its payoff also includes the milieu goals parameter β.[7] The direction (positive/negative) of β depends on third party's legitimacy of means: a marginal third party that views coercion as an acceptable tool has a negative β-parameter ($\beta < 0$), while a legitimate third party, that decries coercion, has a positive β-parameter ($\beta > 0$). The absolute value (size) of β, however, depends on both the importance/benefits of getting involved in a particular case in pursuit of milieu goals minus the expected costs, such as risk of economic sanctions or conflict with other third parties. As a result, the absolute value, and to a lesser extent, the sign of β may vary on a case-by-case basis, as the same third-party may have different milieu-related incentives to intervene in different cases. For example, due to shared history, language, and ethnic origin, Russia may have different values of β for intervening in Ukraine than for intervening in Kyrgyzstan. Although likely negative in both cases (as

[6] A more precise way to model the effect of k would be to allow the government to spend a proportion of k on additional repression and the leftover amount as compensation for loss of office, i.e. *Strong* protesters' payoff would equal $1 - r - \theta k$. We formally explored the effects of such a complication and concluded they are not consequential for the predictions of interest.

[7] As a simplification, we equate the third party's preference that its protégé remain in power with its preference for a particular policy outcome, i.e. third party's benefit from preserving its protégé is part of the policy benefit that equals to 1.

Russia has long demonstrated its acceptance of coercive means), Russia's β may be larger in absolute value in the case of Ukraine than in the case of Kyrgyzstan. Analogously, Britain, France, and the United States may have all had negative β during the height of the Cold War (as evidenced by their general acceptance of coercive means to fight communism), but the size of their β may have varied from case to case, depending on geo-strategic importance, risk of Soviet involvement, economic and cultural ties to the location, etc. (Carment and Rowlands, 1998) .

Finally, if a third party is uninterested/indifferent towards an outcome of the case, then its β-parameter is 0. This would happen, for example, if a state is unimportant to the pursuit of the third party's milieu goals (a state without geo-strategic or economic importance) or the benefits of getting involved do not outweigh the costs (e.g., Russia's benefits from intervening in a NATO member, like the Baltic states, may not be worth the possible costs of conflict with its NATO allies). In either of these scenarios, $\beta = 0$, which means that the third party would have nothing to gain, on balance, from helping the government repress the protesters.[8]

As a result, the use of repression against *Weak* protesters leads to a payoff of $1 - k - \beta\,\mathtt{I}[k > 0]$ to the third party. Since third party's ability to promote its milieu goals is conditional on its involvement in the crisis, the parameter β is multiplied by an indicator variable $\mathtt{I}$, which takes on the value of 1 when $k > 0$, and 0 otherwise. If the third party does not get involved ($k = 0$), then it derives no cost or benefit related to its milieu goals. Finally, the government obtains its preferred outcome minus the costs of repression plus the third party assistance, for the payoff of $1 - r + k$.

The use of repression against *Strong* protesters results in the *Removal from Office* outcome. In this case, the protesters obtain their preferred policy outcome minus the cost of

[8]Note that, within our game, an indifferent third party is not equivalent to a *neutral* third-party. As long as a neutral third-party has an interest in helping resolve the conflict, it must have a non-zero value of β.

repression $1-r-k$.[9] The third party's payoff from this outcome equals to $-k-\beta\,\mathtt{I}[k>0]$. As earlier, β is multiplied by an indicator variable $\mathtt{I}[k>0]$, so that the third party's reputation is only affected, should it provide non-zero assistance k. Finally, the government pays the cost of repression, r, as well as that of removal from office, y ($y>0$), obtaining the payoff of $-y-r+k$.

If the government does not repress ($\neg R$) (whether against *Weak* or *Strong* protesters), the game ends in the *Accommodation* outcome. In this case, the protesters obtain their preferred policy outcome for a payoff of 1, while both the third party and the government obtain the payoffs of 0. If the government chooses $\neg R$, then the third party's assistance k is not disbursed (i.e. the third party keeps it). The government, in other words, cannot choose to accept k and forgo repression. For example, when, during the Orange revolution, Ukraine overturned the results of the fraudulent 2004 election and installed a pro-Western leader, Russia responded by removing Ukraine's natural gas subsidies (Nygren, 2008) . The structure and payoffs of the game are presented in Figure 2.1. To help keep track of notation, Table 2.1 provides a summary of all the parameters and their constraints.

`[Figure 2.1, and Table 2.1]`

2.3.1 Equilibria

The full solution to the game is presented in Appendix. The game has three pure strategy equilibria: *Leader Removal*, *Deterrence*, and *Accommodation*.

Removal Equilibrium is summarized as:

$$\begin{cases} S_G = R, b = 1;\ S_T = k^* = y + r, a = 1;\ S_{SP} = C; S_{WP} = \neg\mathrm{C}; \\ 1 - 2r - y > 0,\ \ r + y < -\beta \end{cases}$$

[9] Granting of the particular policy demand is likely be a pre-condition for the next leader. Protesters, however, gain no additional utility from regime change, i.e. our model is agnostic on whether the new leader will be more or less favorable to other protesters' grievances beyond the policy that led to the regime change.

with payoffs:

$$\begin{cases} U_{SP}(EQ1) = 1 - 2r - y \\ U_{WP}(EQ1) = 0 \\ U_{G|SP}(EQ1) = 0 \\ U_{G|WP}(EQ1) = 1 \\ U_{TP|SP}(EQ1) = -y - r - \beta \\ U_{TP|WP}(EQ1) = 1, \end{cases}$$

where a and b denote government and third party's beliefs that the protesters are of the *Strong* type.

In this equilibrium, we observe the onset of protests only when protesters are of the *Strong* type (*Weak* protesters are deterred); the government uses repression, irrespective of the protesters' type; and the third party provides the government with an assistance $k^* = y + r$ to help offset the expenses associated with repression and/or the cost of losing office.

This equilibrium only exists is the presence of a marginal third party $\beta < 0$, and does not exist in the two-player version of the game (see Appendix for Lemma 2). The dynamics of the game that correspond to this equilibrium, therefore, constitute the core of the paper's contribution and help understand the previously unexplored effects of involvement by a third party. In less technical terms, the *Removal* equilibrium occurs when the third party is of the marginal type: it stands a lot to gain from involvement (β is large in absolute value) and is not averse to bankrolling repression in pursuit of its goals (β is negative). An additional condition is that the costs of repression, r, and leader removal, y, are low to moderate. Third parties that derive a benefit from building a reputation as coercive powers that are not to be

challenged, such as the USSR/Russia, fit these requirements especially well.[10] When these conditions are met, a marginal third party would provide sufficient resources to repress the protesters.

In the game, protesters can anticipate this outcome, and hence, only *Strong* protesters—those that are prepared to fight—challenge the government. The use of major repression against *Strong* protesters is, of course, risky from the perspective of the leader. For example, while the initial protester demands in Maidan related to the country's pro-EU orientation, the first use of repression resulted in cries for Yanukovich's resignation and prosecution. Live images of government brutality against the protesters in Maidan attracted international attention, which further increased the stakes for the Yanukovich regime. Importantly, within the game, these increasing stakes for the leader are recognized by both the leader and the third party: the third party chooses to bankroll repression *despite* the increasing risk of the removal of their protégé leader. In this equilibrium, brutal repression, which helps accomplish its milieu goals, is more important for the third party than the policy at stake and keeping their protégé in power. The protégé leader also recognizes the risks, but since her tenure in office *and* post-tenure fate depend on third party's continued support, she uses repression if such is the preference of the third party, even despite the risks.

This logic is supported by empirical evidence. Using Chenoweth and Stephan (2011) data on protest campaigns between 1899-2006, Table 2.2 provides a cross-tabulation of government use of major repression, based on whether it received overt third-party support.[11] It shows that third-party involvement substantially increases the probability of major

[10] In contrast, if the third party viewed the use of repression as a (unavoidable) cost, the absolute value for β will be smaller, which would make such a third party less likely to meet the condition.

[11] A protest campaign—defined as "as a series of observable, continuous, purposive mass tactics or events in pursuit of a political objective" (Chenoweth and Stephan, 2011) —is a distinct type of a protest event from a small-scale riot or a localized protest. Data on protest campaigns is appropriate for analyzing the *Removal* equilibrium, as this equilibrium predicts that the protests are sufficiently continuous and purposive to pose a threat to the leader's security in office.

repression—a 7 percentage-point difference. And these results are likely to be conservative, due to a lack of data on *covert* third party support.

For further empirical evaluation, we expanded the subset of cases of major repression in the presence of overt third party support (cases in the bottom right cell of Table 2.2) from Chenoweth and Stephan (2011) data to include several additional variables, such as whether the protests resulted in the removal of the leader, the post-tenure fate of such leaders, and the country name of the third party.[12] First, we find that, in the presence of third party support, engaging in major repression results in a rather high rate of leader removal from office (approximately 70% of the cases).[13] Next, Table 2.3 summarizes the post-tenure fate of protégé leaders who were removed as a direct result of using repression against protesters. Consistent with the model, the third party seems to compensate their protégés for a loss of office: a large majority of such leaders—76%—enjoys a safe retirement in their own country, the third-party state, or another friendly state.[14]

`[Table 2.2, Table 2.3, and Table 2.4 here]`

Of course, this equilibrium only occurs for marginal third-parties (Lemma 2), so a more accurate evaluation of empirical evidence would account for third-party type. As preliminary evidence, Table 2.4 breaks up the cases based on the (primary) sponsoring third party and temporal period. Although it is admittedly difficult to devise an accurate measure of third-party type β, the temporal period may serve as a proxy for an increase in β for the United States and Britain, due to their increased post-1990 emphasis on human rights. Consistent with the model expectations, we see a substantial decrease in the number of

[12]See Data section of Appendix for data and coding rules.

[13]We may speculate that, in the rest of the cases, the leader is able to remain in power, despite the use of repression, by using third party's resources to pay off supporters. Moving from a deterministic theoretical model to a probabilistic empirical one, we may think of y as a function of the probability of removal q and a cost of removal ζ, (i.e., $y = q\zeta$). Third party's assistance that compensates the leader for possible loss of power y then may represent resources that leader may use to remain in power and/or spent to ensure his/her post-tenure safety.

[14]Unfortunately, data on leaders removed (regularly or irregularly) as a result of using repression in the absence of a third party are not currently available for comparison.

repressive leaders sponsored by the United States in the post-1990 time period, while the corresponding number for Russia is still high: there are 7 cases in which repressive leaders, sponsored by Russia, found a safe retirement in the post-1990 years, in contrast to only 2 cases for the US. In the pre-1990 period, both superpowers tended to engage in much more sponsorship of repression in third-party states, guaranteeing a safe retirement to 11 (USSR) and 12 (US) leaders.

Deterrence Equilibrium is summarized as:

$$\begin{cases} S_G = R, b = \alpha \text{ ; } \mathrm{S}_T = k^* = y + r, a = \alpha \text{ ; } \mathrm{S}_{SP} = \neg \text{ C; } \mathrm{S}_{WP} = \neg \text{ C} \\ 1 - 2r - y < 0, \;\; r + y < -\beta \end{cases}$$

with payoffs:

$$\begin{cases} U_{SP}(EQ2) = 0 \\ U_{WP}(EQ2) = 0 \\ U_G(EQ2) = 1 \\ U_T(EQ2) = 1 \end{cases}$$

This equilibrium can only occur under conditions that are characteristic of repressive regimes that are sponsored by a marginal third party. Neither protests nor repression constitute part of this expected equilibrium outcome, i.e. the government is able to deter any protests (weak or strong) (Ritter and Courtenay R Conrad, 2016) . It is of interest that, were the protests to occur, the cost of repressing them is fully covered by the third party assistance k^*, which compensates the government's expenses, associated with repression, as well as the costs of removal from office, if necessary. The third party's guarantee of economic help, in other words, is what enables the government to effectively deter the protesters.

Accommodation Equilibrium is summarized as:

$$\begin{cases} S_G = \neg \text{ R},\ b = \alpha;\ \text{S}_T = k^* = 0, a = \alpha; S_{SP} = C; S_{WP} = C; \\ \alpha > \frac{1-r}{1+y},\ \ r + y > -\beta \end{cases}$$

with payoffs:

$$\begin{cases} U_{SP}(EQ3) = 1 \\ U_{WP}(EQ3) = 1 \\ U_G(EQ3) = 0 \\ U_{TP}(EQ3) = 0 \end{cases}$$

In this equilibrium, referred to as the *Accommodation Equilibrium*, the government allows protests and, rather than repressing, tends to find a sustainable accommodation outcome.[15] It is noteworthy that rather than reflecting of the government's tolerance, attributed to liberal democratic regimes, this equilibrium is merely a function of third party's decision against "bankrolling" repression. The government's use of repression, or lack thereof, in other words, is solely determined by the third party. In this case, the third party does not provide k, because it has a high prior belief that the protesters are strong enough to overthrow the government. As long as repression is costly ($r > 0$), the government does not repress any protesters without the help of a third party.

Taken together, the three pure strategies equilibria help explain the known empirical regularity that protest campaigns are overwhelmingly successful. The overall success rate of protest campaigns is 54%, which *increases* to 60%, when the government receives overt support from a third party (a summary of Chenoweth and Stephan (2011) data).[16] The

[15] *Accommodation* equilibrium, however, does not rule out leader removal, i.e. leader removal may itself be the demand of the protesters.

[16] Successful cases are defined as those that resulted in at least partial concessions.

intuition is that potentially unsuccessful cases of *Weak* protesters are deterred from challenging the government in both of the equilibria, in which the government uses repression (*Removal* and *Deterrence*). *Weak* protesters, therefore, only challenge the government as part of the *Accommodation* equilibrium, which results in government concessions. Protest failures, in other words, are less likely to be observed and recorded in the data than protest successes.

2.4 Implications

The game clarifies the causal mechanisms behind the observed outcomes of repression, removal, and accommodation, as well as the frequently unobservable outcome of deterrence. In what follows, we zero in on two types of insights: those that link the outcome with the type of third party, and those that explore the effects of domestic institutional variation.

2.4.1 Third Party Type

We now turn to T's decision regarding the amount k that it allocates to G. Since T is an uninformed actor, selecting which type of equilibrium will occur in the game boils down to the exogenous parameter values in T's expected utilities. There exist ranges of β that allow for each of the pure strategy equilibria, or preclude the existence of certain equilibria. The first insight is that a negative β is a necessary but not sufficient condition for either the *Removal* or the *Deterrence* equilibria: third party's acceptance of coercive means does not, by itself, guarantee that it is willing to bankroll repression in any particular case.

To illustrate this, Figure 2.2 displays the parameter spaces for each equilibrium as a function of the cost of repression, r, on the horizontal axis, and third party type and interest, β, on the vertical axis, while holding the cost of leader removal at a moderate value ($y = 0.5$). Figure 2.2 shows that the game provides unequivocal *pure strategy* predictions regarding the outcome of the interaction between the three players, for any set of parameter values.

The parameter space that allows for *Accommodation* exists under all, even very negative values of β, although its area increases/decreases with changes in β. If we think of the costs of repression as proportional to the spontaneity and size of the protests, this insight may explain why even coercive third parties are known to back down when faced with spontaneous massive protests (Chenoweth and Stephan, 2011) . For example, the United States promptly withdrew its support for Marcos' regime in Philippines, faced with the overwhelming size of the protests in 1986. The value of US's β, although negative (as they continued to prop Marcos despite instances of previous repression), was not large enough in absolute size to justify the enormous costs of repressing a massive protest. Instead, the United States stayed out of the conflict and the protesters obtained their demand of Marcos' resignation. In contrast, had the protest been less spontaneous, Marcos might have been able to use US aid to prevent its occurrence in the first place.

`[Figure 2.2 here]`

A similar logic may also explain why Russia did not assist with repression during the Orange Revolution in 2004, but did during the Maidan protests in 2013-2014. Although Ukraine has consistently been at the center of Russia's milieu goals (large absolute value of β), and Russia is generally accepting of coercion in its domestic and foreign policies ($\beta < 0$), the Orange Revolution, which was the first large-scale and spontaneous protest campaign in post-Soviet Ukraine, took Russia by surprise (Beissinger, 2013) . Having learned from the experience of the Orange Revolution, Russia was more prepared for a popular protest, when its long-term protégé, Yanukovich, backed out of signing the EU Association treaty, at Russia's insistence. Maidan protests, in other words, may have been as large in size, but were less unexpected by the Russian government: since the cost of repression on longer notice is lower than that on shorter notice, the resulting equilibrium outcome shifted from *Accommodation* in 2004 to *Removal* in 2014.

A second insight is that the threshold value of β that rules out the two coercive equilibria (*Deterrence* and *Removal*) actually falls *below* zero. This insight challenges the basic intuition that marginal third parties ($\beta < 0$) will always bankroll repression. This finding also provides an intuition of what type of third parties may act as *neutral*; why even marginal third parties may act neutral in some cases; and why the same third party may act as neutral in some, but not all cases. In particular, as long as β exceeds a certain negative threshold β^*, third party's milieu benefits do not outweigh its costs of bankrolling repression, in which case it prefers to act as an indifferent bystander (do nothing), or even a neutral mediator (e.g., provide non-coercive assistance).

One implication, in particular, is that third-party's decision to stay out of a domestic crisis is not necessarily indicative of its lack of interest ($\beta = 0$) or unacceptability of coercive means ($\beta > 0$). If we were to draw a horizontal line at $\beta = 0$, the area below that line and above the diagonal line β^* corresponds to the parameter space, in which a marginal third party with an interest in the interaction ($\beta < 0$) will act indistinguishable from a neutral or a legitimate third party, i.e. will provide no assistance with repression. This happens when the cost of bankrolling repression does not outweigh the milieu benefits: e.g., Egypt was forced to withdraw its aid to Yemen's al-Sallal regime after its devastating losses in the Six Day War, which moved the outcome from a possible *Removal* equilibrium to the *Accommodation* equilibrium, in which al-Sallal was removed in a "bloodless coup" (Bidwell, 1994) . Another example is the United States' decision to withdraw support from their long-term protégé, the Somozas of Nicaragua: although not fully averse to repression, the United States did not view a repressive outcome as a milieu benefit in itself (negative, but small absolute value of β), and hence, gave up Somoza as his regime's brutality started attracting international and domestic attention. US' withdrawal of support for Mubarak in response to 2011 mass protests serves as an example of a legitimate third party with an important and not selfless stake in an interaction acting indistinguishable from a neutral third party. This insight fits nicely

within the existing research on mediation that argues that "neutral" third parties frequently have ulterior motives (Findley and Teo, 2006) .

2.4.2 Institutional Features and Repression

The theoretical model also speaks to institutional explanations for protest–repression interactions. In particular, our model sheds light on the effect of two important institutional features—the cost of repression r and the cost of a removal from office y. Both of these parameters are proxies for institutional features known to affect the probability of protests and repression. The costs of repression may serve as a proxy for regime transparency; Bel-letal:2012, for example, find that the presence of human rights organizations may decrease repression by increasing the government's cost through publicizing the abuses. Courtenay R Conrad (2014) , similarly, shows that a leader's cost of repression increases in the presence of independent judiciaries.

The second institutional feature—leader's cost/punishment as a result of a removal—has been previously linked to a leader's level of institutional constraints. Leaders of less institutionally constrained regimes, such as personalist autocrats, face higher costs of removal than more constrained leaders. For example, while removal from office is rarely accompanied by additional punishments in democracies, deposed leaders of nondemocratic regimes frequently face additional penalties, such as exile, imprisonment, or even execution (Debs and Hein E Goemans, 2010) .

The model helps understand the interplay between these two parameters and the outcome of the protests. To highlight this aspect of the model, Figure 2.3 provides a visualization. The x-axis displays a range of possible repression costs r, and the y-axis displays a range of costs of removal from office y, while β is constrained to the range associated with a marginal

third party,[17] and the probability that the protesters are strong, α, increases as we move from left to right between subfigures.

[Figure 2.3 here]

First, the figure highlights that, as long as the costs of repression are high, otherwise repressive regimes may *appear* non-repressive and accommodate rather than repress protesters. The *Accommodation* equilibrium exists for any value of the cost of removal from office. If we think of the costs of removal from office as a proxy of whether a regime is democratic, this prediction would suggest that both authoritarian and democratic regimes may accommodate protesters if the cost of repression is sufficiently high (e.g., the regime is monitored by human rights organizations).

Next, the figure shows the combination of parameters for the *Deterrence* equilibrium, in which the presence of a marginal third party helps deter protests against the regime through the threat of repression. Holding third party's preference for repression constant, we see that whether the parameter space is conducive to the *Accommodation* or the *Deterrence* equilibrium largely depends on the belief that the protesters are strong, α: as this belief increases (move from the left subfigure to the right subfigure), so does the area associated with the equilibrium in which the protesters obtain the concessions, *despite* the governments' relationship with a marginal third party. When the belief that protesters are strong is moderate or high (e.g., $\alpha = 0.6$ in the right subfigure), both authoritarian and democratic regimes accommodate the protesters' demands.[18] When the belief that protesters are strong is low (e.g., $\alpha = 0.2$ in the left subfigure), a regime may act in a repressive manner (i.e. deter protests through the threat of repression).

[17] We focus on the marginal third party equilibria, as legitimate third parties never bankroll repression.

[18] The threshold is $\alpha > .5$. When $\alpha = 0.5$, $1 - 2r = (1 - r)(\alpha) - 1$, i.e. the two equilibria conditions, represented by solid lines in Figure 2.3 overlap. To enhance visualization, we therefore hold α at 0.6 rather than 0.5 in the right subfigure.

Third, the model speaks to the literature on individual leader outcomes (Hein E Goemans; Henk E Goemans, Gleditsch, and Chiozza, 2008; 2009) . As the separating equilibrium is the only equilibrium, in which protests, repression, and leader removal are part of the observed outcome, our model helps identify the parameter space that increases the risk of a leader's removal. According to the logic of the model, the separating equilibrium is observed when (1) strong protesters have a positive expected utility, even after accounting for the costs of repression, $1 - 2r - y > 0$, (2) when the third party is of the marginal type , $\beta < -r - y$, and (3) when the third-party and the government expect challenges only from the strong protesters, $a = b = 1$. As highlighted in Figure 2.3 and conditions (1)-(2), this equilibrium is possible when both the cost of leader removal, y, and the cost repression, r, are relatively low, i.e. are justified by the milieu benefits for the third party.

The second of the above conditions implies that leaders are most likely to be removed from office when the costs of repression are low for the third party. Costs of bankrolling repression, for example, may be lowered by close collaboration/interconnectedness between the internal police and security apparati of the third party and its protégé governments (e.g., close cooperation between Russian and Ukrainian internal security forces). Greater degrees of dependence between the third party and its protégé make it easier (less costly) for the third party to bankroll repression, and also install a new protégé leader in the future, should the current leader be removed as a result of repression. This prediction sheds light on the motivating example of Ukraine, as well as on a number of other cases of removal of leaders propped up by third parties.

2.5 Conclusion

State responses to popular protest have traditionally been modeled as a domestic phenomenon. Such studies often treat international influences in structural rather than strategic terms

(e.g., control for proximity to a major power). Growing empirical evidence, however, suggests that, in a significant number of cases, the domestic-level interaction between the government and the protesters may be affected by involvement of an outside third party with its own stakes in the matter. The current study zeroes in on third-party involvement with the goal to affect the protester's decision to challenge the government and the government's response to such a challenge.

The study contributes to the understanding of the relationship between state leaders and their domestic audiences in cases, where the leader herself depends on support from an outside third party. Approaching the interaction from a game-theoretic perspective, we identify the conditions, under which the leader may choose to engage in inefficient repression against the protests, even at the risk of her own removal from office. We show that this puzzling outcome is possible, when the leader's response to protests is influenced by an outside third party with broader regional or systemic goals. In the motivating example of Ukraine, Yanukovich is the protégé leader, and Russia acts as his third party sponsor. Yanukovich's abrupt withdrawal from the EU Association treaty—a policy highly sought by a significant part of the population—triggered the Maidan protests. Despite every indication of the protesters' strength—their large numbers that counted in hundreds of thousands, their willingness to brave the cold of Ukrainian winter, and the threat of police brutality—Yanukovich made no serious attempts at accommodation, choosing instead to repress, despite the high risk of removal. We explain this outcome by zeroing in on Russia's role in the crisis, in light of its chief concern with creating a "scarecrow" to help prevent other "Maidans" within its sphere of influence. Russia provided Yanukovich with both the means to carry out the repression against the protesters, and the safe option to flee the country into a comfortable retirement.

Our study also contributes to the game-theoretic literature on the onset of protests and repression, much of which has treated inefficient repression as off-equilibrium behavior ([[)]Ritter:2014, Pierskalla:2009. We can certainly treat Yanukovich's decision to turn

Kyiv's Independence Square into a bloodbath as a terrible blunder. While plausible, such an explanation, however, is rather unsatisfying. Moreover, given what we know about the close relationship between Yanukovich and Russia, a more satisfying explanation would help provide insights regarding the reasons for, and implications of, possible Russian involvement.

Some research treats repression as part of a mixed strategy equilibrium, in which *Strong* protesters challenge the government with some positive probability, and the government mixes between opting for repression and accommodation. One way to think of the mixed strategy equilibrium is that, from time to time, governments accept some risk of losing office and incurs the costs of repression, in order to deter some future protests. E.g., Maidan was a calculated risk on the part of Yanukovich, in an attempt to intimidate future protesters. Mixed-strategy equilibrium explanations, however, are rather unintuitive. For cases with evidence of third-party involvement, repression as a part of pure strategy equilibria, as in our model, is both more intuitive and satisfying.

Tables and Figures

Figure 2.1: An Extensive-Form Three-Player Game of Protests and Repression with Incomplete Information about Protesters' Type

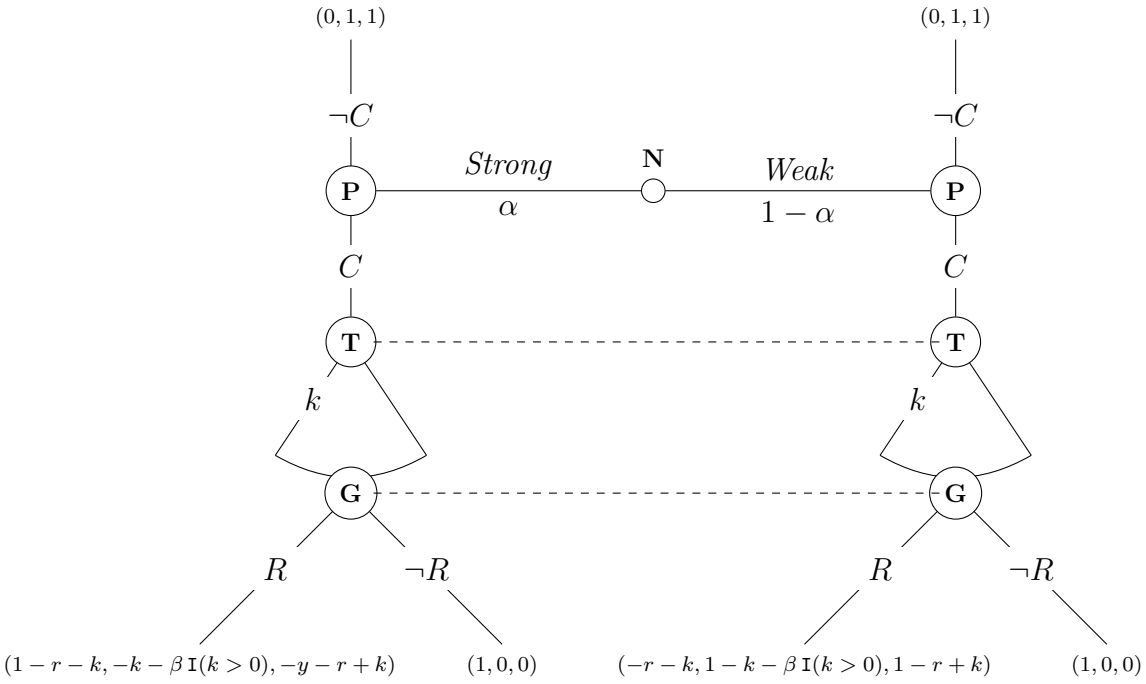


Figure 2.2: Effect of Third-Party Milieu Goals and Means (β), and the Cost of Repression (r) on Equilibria Outcomes.

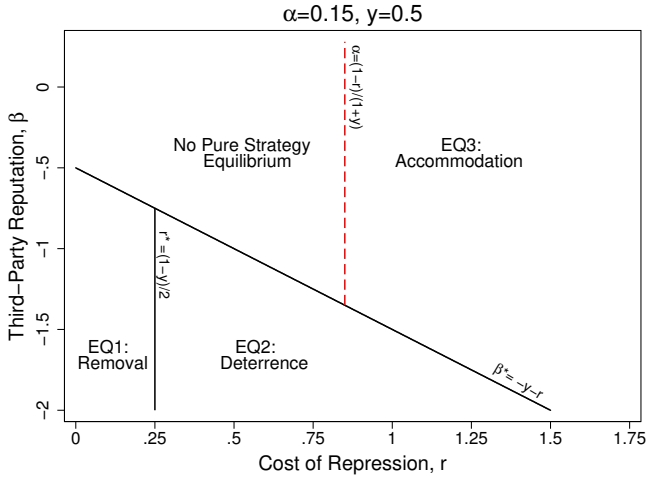


Figure 2.3: Effect of Regime Type Parameters (y and r) on Equilibria Outcomes

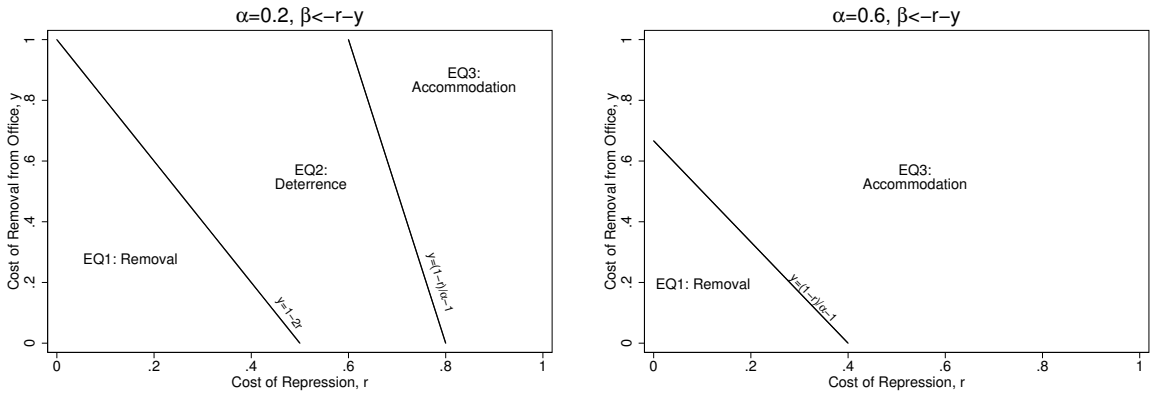


Table 2.1: Game Parameters

Parameter	*Description*	*Constraints*
α	Probability that the protester's are strong	$0 < \alpha < 1$
r	Cost of repression	$r > 0$
β	Third party's broader cost/benefit from achieving milieu goals	$-\infty < \beta < +\infty$
y	Cost of being removed from office	$y > 0$

Table 2.2: Overt Support for the Government by a Third-Party State During a Protest Campaign, 1899-2006

	Major Repression		
	No	Yes	Total
No Overt 3^{rd} Party Support for Government	31 (14.16)	188 (85.84)	219 (100)
Overt 3^{rd} Party Support for Government	7 (6.73)	97 (93.27)	104 (100)
Total	38 (11.76)	285 (88.24)	323 (100)

Source: Chenoweth and Stephan (2011) . Numbers in parentheses represent percent, by row. $\chi^2_1 = 3.74$, $p = 0.053$

Table 2.3: Post-Tenure Fate of Protégé Leaders, 1899-2006

Safe Retirement			*No Safe Retirement*			*Indeterminable*		
Stayed in Country	22	43%	Imprisoned	6	12%	Natural Death	2	4%
Exile	17	33%	Executed	2	4%	Assassinated	2	4%
Total:	39	76%	*Total:*	8	16%	*Total:*	4	8%

Table 2.4: Third Party Type and Post-Tenure Fate of Protégé Leaders, 1899-2006

	Before 1990		*1990-2006*	
Third Party	Not Safe	Safe	Not Safe	Safe
USSR/Russia	2	11 (9)	0	7 (4)
US	4	12 (5)	0	2 (0)
Britain	0	2 (1)	1	0 (0)
Syria	0	1 (1)	0	1 (1)
China	0	1 (1)		
France	0	2 (0)		
Iran	1	0 (0)		

Note: Numbers in parenthesis indicate the number of cases in which leader was able to stay in his/her own country vs. go into exile to a different country.

2.6 Appendix: Game Solution

2.6.1 Pure Strategy

We solve the game using the Perfect Bayesian equilibrium (PBE) solution concept. The first insight that is revealed by the model relates to the government's decision to use repression. The government's decision reduces to choosing the strategy that yields the highest expected payoff. By subgame perfection, the government would repress the protesters at the final decision node if, and only if, its expected payoff from repressing is at least as good as from accommodating. Let us denote the government's belief that the protesters are strong as b. Then the government's payoff from choosing to repress can be expressed as $U_G(R) = b(-y-r+k)+(1-b)(1-r+k) = k-b(y+1)+1-r$ and the government's payoff from choosing to forgo repression is $U_G(\neg R) = 0$. Therefore, the government will choose to repress if $U_G(R) \geq U_G(\neg R)$, or $k-b(y+1)+1-r \geq 0$. Solving for b, we obtain that, ignoring knife-edge cases, G will repress when:

$$b < \frac{1-r+k}{y+1}. \tag{2.1}$$

Note that there exists a range of conditions, for which the government's decision to use repression against protesters depends, in large part, on the size of the assistance k provided by the third party.[19] This means that, should it so wish, the third party can single-handedly induce the government to use repression simply by offering sufficiently large assistance k.

Lemma 1. *When $b(y+1)+r > 1$, government's decision to use repression depends solely on the size of third party assistance k. Government uses repression against protesters, when the third party assistance $k \geq b(y+1)+r-1$. When $k < b(y+1)+r-1$, the government does not use repression.*

[19]The specific conditions are $b(y+1)+r > 1$, which would hold, for example when the government has a non-zero belief that the protesters are strong, and the cost of repression is relatively high.

Suppose that k^* is the minimum assistance by T that ensures that G uses repression. There can be at most one such *assured repression* level in equilibrium. To see that, suppose that there were more than one. But then the third party could always benefit by switching to the lower of the possible amounts, i.e., k^* is the lower bound on the level of assistance by the third party in any equilibria, in which the government uses repression. Furthermore, the third party would only allocate k^* if its payoff from an outcome, in which the government uses repression outweighed its payoff from an outcome, in which the government does not use repression.

To derive the conditions, under which the third party prefers the outcome, in which the government uses repression, we denote the third party's belief that protesters are strong by a. We further assume that $a = b$, since neither T or G has any additional information to enhance their prior beliefs regarding α. Empirically, this assumption is grounded in each party's lack of an incentive to withhold relevant information, given their shared preferences regarding the policy at hand: i.e. the government and the third party have an incentive to work together and share intelligence regarding the type of the protesters.

The third party's payoff, conditional on protesters challenging and the government repressing, is $U_T(k|R) = 1 - a - k - \beta\,\mathtt{I}[k > 0]$. If the government does not use repression, the third party obtains the payoff of $U_T(k|\neg R) = 0$. As long as $U_T(k|R) \geq 0$, the third party prefers that the government use repression and provides assistance k, such that:

$$\begin{aligned} \max_{k} \quad & 1 - a - k - \beta\,\mathtt{I}[k > 0] \geq 0 \\ \text{s.t.} \quad & k - b(y+1) + 1 - r \geq 0, \quad U_T(k|R) \geq 0 \end{aligned} \tag{2.2}$$

This problem has a non-zero binding solution:

$$\mathrm{k}^* = b(y+1) + r - 1 \tag{2.3}$$

$$\text{s.t.} \quad 0 < b(y+1) + r - 1 \leq \quad k \leq 1 - a - \beta.$$

The final step is to look at the behavior of the protesters. If government's strategy is $S_G = R$, then $U_{WP}(C) = -r - k$, which is negative for any k, and therefore worse than $U_{WP}(\neg C) = 0$. Hence, weak protesters never challenge in the equilibria, in which the government uses repression. Therefore, if G observes C, it believes that the protesters are of the *Strong* type, or $b = 1$. Substituting $b = a = 1$ into the utility of T, we can derive T's strategy—the optimal amount k that would assure that G does not deviate from R:

$$\mathrm{k}^* = b(y+1) + r - 1 = y + r \tag{2.4}$$

$$\text{s.t.} \qquad\qquad y + r \leq -\beta. \tag{2.5}$$

Now that we know k^*, we can evaluate SP's utilities to see that SP plays C as long as $U_{SP}(C) \geq U_{SP}(\neg C)$, or $1 - 2r - y \geq 0$. The separating equilibrium, EQ1, derived above, can be summarized as:

$$\begin{cases} S_G = R, b = 1;\ S_T = k^* = y + r, a = 1;\ S_{SP} = C; S_{WP} = \neg\ \mathrm{C}; \\ 1 - 2r - y > 0,\ \ r + y < -\beta \end{cases}$$

The payoffs:

$$\begin{cases} U_{SP}(EQ1) = 1 - 2r - y \\ U_{WP}(EQ1) = 0 \\ U_{G|SP}(EQ1) = 0 \\ U_{G|WP}(EQ1) = 1 \\ U_{TP|SP}(EQ1) = -y - r - \beta \\ U_{TP|WP}(EQ1) = 1 \end{cases}$$

Notably, Inequality 2.5 indicates that this equilibrium only exists for marginal third parties ($\beta < 0$, or more specifically $\beta \leq -r - y$) who derive a benefit (e.g., international reputation for "toughness") from a repressive outcome.

Lemma 2. *A separating equilibrium only exists for marginal third parties, or* $\beta < 0$.

Next, let us determine whether there exists a pooling equilibrium, such that G always represses. The pooling equilibrium in which both types of protesters challenge is impossible (since, as shown above, if G represses, weak protesters do not challenge). Hence, the only possible pooling equilibrium, given the government represses, is the one in which SP and WP do not protest. Hence, ignoring the knife-edge cases, the game has the following pooling equilibrium, EQ2:

$$\begin{cases} S_G = R, b = \alpha \text{ ; } \mathrm{S}_T = k^* = y + r, a = \alpha \text{ ; } \mathrm{S}_{SP} = \neg \text{ C; } \mathrm{S}_{WP} = \neg \text{ C} \\ 1 - 2r - y < 0, \;\; r + y < -\beta \end{cases}$$

The payoffs:

$$\begin{cases} U_{SP}(EQ2) = 0 \\ U_{WP}(EQ2) = 0 \\ U_G(EQ2) = 1 \\ U_T(EQ2) = 1 \end{cases}$$

We denote this equilibrium as the *Deterrence Equilibrium.*

Next, consider the situation when G never represses, which happens when $a > \frac{1-r+k}{y+1}$ (by Equation 2.1). In that case, T does not provide k. Substituting $k^* = 0$ into (1), we obtain $a > \frac{1-r}{y+1}$, which is stable as long as T does not have an incentive to deviate from $k^* = 0$, or $y + r > -\beta$. This gives us another pooling equilibrium, in which G never represses, T provides assistance of $k = 0$, and both types of P choose to challenge. This equilibrium satisfies the definition of PBE, when $a = b = \alpha$. Then, the game has a second pooling equilibrium, EQ3:

$$\begin{cases} S_G = \neg \text{ R}, b = \alpha; \text{S}_T = k^* = 0, a = \alpha; S_{SP} = C; S_{WP} = C; \\ \alpha > \frac{1-r}{1+y}, \ \ r + y > -\beta \end{cases}$$

The payoffs:

$$\begin{cases} U_{SP}(EQ3) = 1 \\ U_{WP}(EQ3) = 1 \\ U_G(EQ3) = 0 \\ U_{TP}(EQ3) = 0 \end{cases}$$

Finally, PBE solution concept requires that we define strategies and beliefs even for non-equilibrium paths. The game is structured in such a way that such off-equilibrium

paths exist for G after some $k \neq 0$. Since $S_G = R$ as long as $\alpha < \frac{1-r+k}{1+y} \ \forall \ k$, we define $a = b = 0 \ \forall \ k \neq 0$, which implies that $S_G = R \ \ \forall \ k \neq 0$.

Sequential rationality and the intuitive criterion

Another critical question is whether these PBEs are sequentially rational. The strongest critique of the notion of the PBE is its possible irrationality from the perspective of the forward induction (e.g., see Mas-Colell, Whinston, Green, et al. 1995). The intuition beyond this concern is based on that in reality actors take actions sequentially. Hence, if they discover themselves on an off-equilibrium path, they must act rationally given that they are *already* there. However, the *ex-ante* approach to the derivation of PBEs may lead to the existence of such off-equilibria strategies (supporting the behavior on the equilibrium path) that are irrational if the actor happens to find herself off the equilibrium path.

These critiques lead to the development of the literature proposing possible PBE refinements (Selten 1975, Kohlberg and Mertens 1986, Kohlberg 1990). One of the ways to capture sequential *rationality* is imposing *the intuitive criterion* (for more information see Cho and Kreps 1987). The idea of this criterion is that no player (the sender) should yield a higher utility off-path than in the equilibrium while there exists another player (the receiver) that assigns a zero probability to this off-path strategy.

Do the equilibria in our game survive the IC? In the context of our game, the violation of the IC would mean that T and G base their actions in the lower part of the game on the assumption that they respond to the weak type of the protesters. The reason for this is that the strong type is more likely to end up in the lower part of the game, since there is a larger parameter space, for which it is rational to contest for the strong type. For any possible action of T and G, the utility of the strong type is not less than of the weak type. Hence, if an equilibrium survives the IC, the strategies of T and G must stay rational even it is only the strong type that contests.

EQ1 is separating and only the strong type contests, hence the IC does not apply at all. In EQ2, the game does not reach the lower part of the game graph, and the IC does not apply as well. Finally, the only equilibrium in that the IC might be potentially violated is EQ3. Does it happen? No, it does not since the choice of k and, then, whether or not to repress the protesters is the result of the relation, $r + y > \beta$. It is not based on the assumption that the protesters are weak. Contrary to that, if the protesters are indeed strong given, $r + y > \beta$; even if $\alpha = 1$ and so are the beliefs, the rational strategies of G and T stay the same and EQ3 holds.

2.6.2 Mixed Strategy

The game also has a set of semi-separating equilibrium, in which $S_{WP} = \neg C$, but SP plays C with a positive probability. First, remember that $S_{WP} = \neg C$, as long as $a = b > \frac{1-r-k}{1+y}$. Second, in order to mix C and $\neg C$, SP must be indifferent between these two strategies, i.e. G must mix R and $\neg R$ with such a probability that leaves SP indifferent between C and $\neg C$. More formally, if we denote G's probability of playing R as γ, then:

$$\begin{aligned} \gamma(1 - r - k) + 1 - \gamma = 0; \\ \gamma(1 - r - k - 1) = -1; \\ \gamma = \frac{1}{r + k}. \end{aligned} \tag{2.6}$$

In an equilibrium, condition 2.6 must be consistent with SP's strategy. If we denote SP's probability of playing C as w, then:

$$w\alpha = a \tag{2.7}$$

or

$$w = \frac{1 - r + k}{\alpha(1 + y)}. \tag{2.8}$$

These strategies must be stable. Since T's strategy k is continuous, we define a semi-separating equilibrium for our three-player game as an equilibrium in which T's strategy cannot affect the outcome of the game: in that case T will provide $k = 0$. Hence, in semi-pooling equilibria, the game effectively reduces to a two-player signaling game between P and G.

Then, the semi-pooling equilibrium can be summarized as:

$$\begin{cases} S_G = (\min(\frac{1}{r}, 1), 1 - \min(\frac{1}{r}, 1), \alpha = a = b = \max(\frac{1-r}{1+y}, 0)); \\ S_T = k^* = 0; \\ S_{SP} = (\max(\frac{1-r}{\alpha(1+y)}, 0), 1 - \max(\frac{1-r}{\alpha(1+y)}, 0); \\ S_{WP} = \neg C. \end{cases}$$

The concept of the Perfect Bayesian Equilibrium (PBE) requires that we also define the strategies for non-equilibrium paths. Since the semi-pooling equilibrium defined above holds for all k, we assume that if T provides any amount other than $k^* = 0$, the strategies of the government and the protesters remain the same.

This semi-separating equilibrium allows for two types of observable outcome: (1) when $r > 1$, the government represses with a probability $\frac{1}{r}$, and neither type of protesters challenge ($S_{SP} = S_{WP} = \neg C$); and (2) when $0 < r < 1$, then the government's and the third-party's believe that the protesters are of the *Weak* type ($a = b = 0$), in which case the government always represses, and the *Strong* protesters challenge with any probability. Note that the outcome, in which we observe both protests and repression is not part of the semi-pooling

equilibrium. Hence, given the parameter restrictions adopted in the current version of the game, the outcome of observed conflict between the government and the protesters as part of an equilibrium is only possible in the separating equilibrium (i.e., *Removal* equilibrium).

Conflict between the government and the protesters is, of course, obtained as apart of the semi-separating equilibrium in a more general form of the game that relaxes the assumption that the policy benefit is zero-sum for the three players. Suppose, instead, that instead of a policy benefit of 1 for each player, the protesters' policy benefit is m_p, the third party's policy benefit is m_t, and = the government's policy benefit is m_g. Then, the semi-separating equilibrium may be re-written as:

$$\begin{cases} S_G = (\frac{m_p}{r}, 1 - \frac{m_p}{r}), \alpha = a = b = \frac{m_g - r}{m_g + y}; \\ S_T = k^* = 0; \\ S_{SP} = (\frac{m_g - r}{\alpha(m_g + y)}, 1 - \frac{m_g - r}{\alpha(m_g + y)}); \\ S_{WP} = \neg C. \end{cases}$$

In this more general case, we observe conflict between the government and the protesters as part of a semi-separating equilibrium, as long as $m_g > r$.

2.6.3 The Special Case of Third-Party with No Interest in the Interaction

The case $\beta = 0 \implies k = 0$, hence the third-party has no interest in the game and becomes insignificant. What does the game look like in that case? Then, G plays R given the belief b: $u_G(R) = (-y - r)b + (1 - b)(1 - r) \geq 0 = u_G(\neg R)$.

1. Check for a separating equilibrium, such that $S_G = R$, which requires $b < \frac{1-r}{1+y}$, which, in turn implies that $S_{SP} = C$ and $S_{WP} = \neg C$. However, then b must be 1, hence this equilibrium is inconsistent and does not exist.

2. Check for a pooling equilibrium, such that $S_G = \neg R$, which requires that $b > \frac{1-r}{1+y}$, which implies that $S_{SP} = C$ and $S_{WP} = C$. Since both types of the protesters are equally likely to protest in this case $b = \alpha$. Hence, this equilibrium exists if $\alpha > \frac{1-r}{1+y}$

3. As shown above, these conditions also allow for a semi-separating equilibrium.

2.7 Appendix: Data

In order to conduct an empirical evaluation of the model's predictions, we collect original data on the post-tenure fate of leaders who received overt support from a third-party. We started with the cases that are likely to result in the separating equilibrium of our game: those identified by Chenoweth:2011 as (1) characterized by major instances of government repression, and (2) in which leader received overt third-party support. Cases of foreign occupation during international wars, such as German occupation during WWII, are excluded, as the third party is the same as domestic government (e.g., German occupation of Poland). The resulting dataset consists of 77 cases. The unit of analysis is the protest campaign. Our dataset extends GoemansGleditschChiozza:2009 dataset, as those do not include leaders whose tenure lasted less than 3 years. Our data also extend the name of the third party, which is not included in Chenoweth:2011 dataset.

For each case, we collected information on the following variables:

- *Leader*—the name of the leader who faced the protesters and was supported by the third party. If a protest campaign took place during tenure of several leaders, we collect information on the leader who was in power in the beginning of the protests and can

thus be directly linked to the issue of protester's grievance. Most protest campaigns end or decline in scale within the tenure of a single leader.

- *Third Party*—the name of the country that operated as the third party;
- *Removal*—takes on the value of 1 if the leader was removed (as a result of a regular election, a special election, a coup, a resignation, etc.) as a direct consequence of protests;
- *Fate*—post-tenure fate of the leader in the first several years after removal. *Fate* takes on the following values:
    1. *Stayed*—leader stayed in their own country;
    2. *Exile*—leader went into exile, "medical treatment," or moved to a different country for other reasons;
    3. *Prison*—leader was imprisoned, including cases of house arrest;
    4. *Execution*—leader was executed as a result of a trial;
    5. *Assassinated*—leader was assassinated;
    6. *Natural Death*—leader died a natural death of old age or illness.

The data, along with a brief description of third-party's role in each case, are presented in Table 2.5. All the data were collected by two independently-working coders. A small number of the cases of disagreement between the coders were reviewed/corrected by the authors.

Table 2.5: Post-Tenure Fate of Repressive Leaders, 1899–2006

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
Cuba	1956	1959	Cuban Revolution	Batista	1	Exile	US	Financial, military, and logistical support from the United States government (Leonard, 2004, p. 20) .
Haiti	1918	1920	Caco Revolt	Dartiguenave	0		US	
Dominican Republic	1965	1965	leftists	Cabral	1	Stayed	US	The US stationed about 23000 troops (Paul, 2008, p. 15) .
Mexico	1923	1924	Huerta led rebels	Obregon	1	Stayed	US	The US sold $1.28 million worth of weapons to Obregon (Gonzales, 2002, p. 200) .
Honduras	1924	1924	Conservative movement	Gutiérrez	1	Nat. Death	US	Involvement of US marines (Daugherty, 2015) .
El Salvador	1979	1981		Ramos	1	Exile	US	US economic and military aid (Byrne, 1996, p. 92) .
El Salvador	1979	1991	Farabundo Marti National Liberation Front (FMLN)	Duarte	1	Exile	US	US military aid peaked in 1984, but decreased to $81.5 million in 1988. US economic aid is estimated between $300-$500 million. (Carothers, 1991, p. 32)
Panama	1987	1989		Noriega	1	Prison	US	Noriega worked closely with the CIA in the beginning of his term. In 1988, US suspended economic and military aid (Scranton, 1991) .
Venezuela	1958	1958		Jiménez	1	Exile	US	The U.S. had a commercial relationship with Jimenez dictatorship. Eisenhower awarded him the Legion of Merit Medal in 1954. After he was removed from the office, Jimenez fled to Miami (Galván, 2012, p. 69) .
Chile	1983	1989		Pinochet	1	Stayed	US	Both Nixon and Kissinger provided covert assistance (Kornbluh, 2013, pp. 211, 214, 432) .
Argentina	1977	1981	pro-democracy movement	Videla, Viola	1	Prison	US	US military assistance ended during the Carter administration as a result of human rights abuses (Sheinin, 2010, p. 164) .
Uruguay	1904	1904	Blancos Rebellion	Ordóñez	1	Stayed	US	The US sent USS Tacoma (Knarr, 2012, p. 22) .
Spain	1936	1939	Fascists	Negrín	1	Exile	France, Soviet Union	Purchased weapons from the Soviet Union and France (Beevor, 2006, p. 413) .
East Germany	1953	1953		Ulbricht	0		Soviet Union	Soviet troops (Granville, 2006) .

Continued on next page

Table 2.5 – *Continued from previous page*

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
East Germany	1956	1956		Ulbricht	0		Soviet Union	Soviet troops (Granville, 2006) .
Poland	1956	1956		Ochab	1	Stayed	Soviet Union	Ochab had ties to the Soviet military, fought on the Soviet side in World War II (Cook, 2001, p. 947) .
Poland	1968	1970		Gomulka	1	Stayed	Soviet Union	Soviet troops (Babiracki, 2015, p. 413)
Poland	1981	1989	Solidarity	Jaruzelski	1	Stayed	Soviet Union	Military and economic help from the Soviet Union (J. Gallagher, 1982) .
Austria	1934	1934	socialists	Dollfuss	1	Assass.	Soviet Union	Dollfuss cut ties with the Soviet Union and outlawed communism (Bradberry, 2012, p. 279) .
Hungary	1989	1989	pro-democracy movement	Grósz	1	Stayed	Soviet Union	
Czechoslovakia	1989	1989	Velvet Revolution	Husák	1	Stayed	Soviet Union	Soviet Union helped Gustav Husak to obtain power in 1968, although there is no evidence that the Soviet Union provided assistance during the Velvet Revolution (Roman, 2003, p. 624) .
Italy	1943	1945	Italian resistance	Badoglio	1	Stayed	Germany, Britain	Britain saw Badoglio as an ally against communism (S. C. Tucker, 2014, p. 197) .
Albania	1989	1989		Alia	0		Soviet Union	Threatened by protests, Alia transitioned from Communism to a multi-party system, thus losing Soviet support (Bideleux and Jeffries, 2007) .
Yugoslavia	1970	1971	Croatian nationalists	Tito	0		Soviet Union	
Yugoslavia	1981	1981	Kosovo Albanian nationalist movement	Mijatovic	1	Stayed	Soviet Union	
Bosnia-Herzegovina	1991	1995	Serb militias	Milosevic	0		Russia	
Slovenia	1989	1990		Milosevic	1	Stayed	Soviet Union	
Greece	1944	1949	Greek communists	Sofoulis	1	Nat. Death	US	US military aid (Couloumbis, 1966, pp. 34-36) .
Greece	1974	1974		Ioannidis	1	Prison	US	US military aid (Renouard, 2015, p. 56) .
Bulgaria	1989	1989		Zhivkov	1	Prison	Soviet Union	Zhivkov lost his support from the Soviet Union (Dimitrov, 2013) .

Continued on next page

Table 2.5 – *Continued from previous page*

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
Romania	1987	1989		Nicolae Ceasescu	1	executed	Soviet Union	Ceasescu tried to play the Soviet Union off of the US and took aid from West (W. C. Thompson, 2012, p. 527) .
Estonia	1989	1989	Singing Revolution	Sillari			Soviet Union	
Ukraine	2001	2004	Orange Revolution	Yanukovich	1	Stayed	Russia	Financial assistance from Russia (Moser, 2014, p. 137) .
Belarus	1989	1989		Sokolov	1	Stayed	Soviet Union	
Georgia	1991	1994	Gamsakurdia & Abkaz	Shevardnadze	0		Russia	A treaty with Russia, Russian military aid (Bohlen, 1994) .
Azerbaijan	1991	1994	Armenians in Nagorno-Karabakh	Mutallibov	1	Exile	Russia	Mutallibov lost Russian support after failing to negotiate an energy deal (Vanderhill et al., 2013, p. 127) .
Azerbaijan	1991	1994	Armenians in Nagorno-Karabakh	Mutallibov	1	Exile	Russia	Mutallibov lost Russian support after failing to negotiate an energy deal (Vanderhill et al., 2013, p. 127) .
Benin	1989	1990		Kérékou	1	Stayed	Russia	Faced with protests, leader abandoned communism.
Liberia	2003	2003	LURD	Taylor	1	Exile	US	Logistical support from the US military (New York Times, 2003) .
Nigeria	1967	1970	Biafrans	Gowon	0		Britain	British army helped repress the protesters (Leapman, 1998) .
Chad	1966	1990	Frolinat	Tombalbaye; Malloum	1	Assass.	France, Nigeria	Military support from France between 1968-1984. Malloum tried to free himself from economic and political dependence on France, but was not successful. France continued to provide consultation and renewed military aid. Later in the conflict, France negotiated directly with rebels cutting out Malloum (Ciment et al., 1998, p. 391) .
Uganda	1996	2006	LRA	Museveni	0		US, African Union	The U.S gave about $750 million of aid annually and trained their military for 10 years (, 2016) .
Ethiopia	1974	1991	Eritrean-led rebels	Mengistu	1	Exile	Soviet Union	$4 billion in military hardware between 1978-1986, Mengistu signed a 20 year treaty with the Soviet Union that increased the level of military aid helped fight the Somalian guerrillas and Eritrea (Rule, 1988) .
Ethiopia	1976	1983	Somali rebels (Ogaden)	Mengistu	1	Exile	Soviet Union	$4 billion in military hardware between 1978-1986, Mengistu signed a 20 year treaty with the Soviet Union that increased the level of military aid helped fight the Somalian guerrillas and Eritrea (Rule, 1988) .
Ethiopia	1978	1991	Tigrean Liberation Front	Mengistu	1	Exile	Soviet Union	$4 billion in military hardware between 1978-1986, Mengistu signed a 20 year treaty with the Soviet Union that increased the level of military aid helped fight the Somalian guerrillas and Eritrea (Rule, 1988) .

Continued on next page

Table 2.5 – *Continued from previous page*

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
Angola	1975	2001	UNITA	Neto; dos Santos	0		Soviet Union, Cuba	Cuban troops used flame throwers, bulldozers, and planes with napalm to destroy population centers. Eventually Neto lost Soviet support by refusing the construction of Soviet military bases (Abbott, 2011, p. 10)
Mozambique	1979	1992	Renamo	Machel; Joaquim Chissano	0		France, Britain, US	U.S. provided $10 million in economic and humanitarian aid as well as military aid to Machel. Third-parties withdrew support after the Cold War (Pascoe, 1986) .
Zimbabwe	1974	1980	Zimbabwe African People's Union	Smith; Muzorena	1	Stayed	Soviet Union, Cuba	
Morocco	1953	1956	Moroccan Independence War	Aarafa	1	Exile	France, Spain	France helped Arafa come to power (International Business Publications and Ibp, 2004, p. 111) .
Sudan	1962	1973	Anya Nya	Abboud	1	Exile	Britain	Economic and military support from the British: training airforces and supplying civil and combat aircrafts. Sudan also bought armored cars and weapons from Britain (Poggo, 2008, p. 165) .
Sudan	1983	2005	SPLA-Garang faction	Nimeiry	1		US, Iran	U.S.-aligned countries resumed supplying Sudan in the mid-1970s. The United States began selling Sudan equipment around 1976, hoping to counteract Soviet support of Marxist Ethiopians and Libyans. Military sales peaked in 1982 at $101 million. West Germany established an ammunitions factory in Khartoum and supplied quantities of automatic small arms. After the beginning of the second civil war, American assistance dropped, and was eventually cancelled in 1987. In 1993, Iran started providing financial aid. (de Lama, 1985)
Iran	1977	1979	Iranian Revolution	Pahlavi	1	Exile	US	US helped istall Pahlavi in a coup and supported his regime since (Hastedt, 2014) .
Iran	1979	1996	KDPI	Khamenei; Rafsanjani	0		Russia	Military aid (San-Akca, 2016, p. 202) .
Iraq	1959	1959	Shammar Tribe and pro-Western officers	Qasim	0		Soviet Union	Economic and military aid (S. C. Tucker, 2010, p. 622) .
Iraq	1961	1975	Kurdish rebellion	Qasim	0		Soviet Union	Economic and military aid (S. C. Tucker, 2010, p. 622)

Continued on next page

Table 2.5 – *Continued from previous page*

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
Lebanon	1958	1958	leftists	Chamoun	1	Stayed	US	Eisenhower sent 5,000 soldiers and marines who stayed in Beirut for six months (, 1987) .
Lebanon	1975	1975	leftists	Frangieh	1	Stayed	Syria	President Frangieh asked for a Syrian intervention aid in ending the civil war (Salamey, 2013, p. 46) .
Lebanon	1982	2000	Hizballah	Israel occupation	0		Israel	
Lebanon	2005	2005	Cedar Revolution	Karami	1	Stayed	Syria	Peaceful non-sectarian protests against Syrian involvement, including the presence of 14,000 Syrian troops and intelligence officers (Whitaker, 2013)
Yemen Arab Republic	1962	1969	Royalists	al-Sallal	1	Exiled	Egypt, Soviet Union	Egypt had to withdraw aid after 6-day war. Al-Sallal signed a treaty of friendship and economic and technical coorporation, which, in essence, to a $39 million Soviet loan (Abir and Yodfat, 2013, p. 51) .
Afghanistan	1978	1979	Afghans	Taraki	0		Soviet Union	The Soviets provided trucks, artillery, tanks and other weapons and increased the number of military advisers (Oberdorfer, 1980) .
Tajikistan	1992	1997	Popular Democratic Army (UTO)	Rakhmon	0		Russia	Support of Russian and Uzbek military forces (Dagiev, 2013, p. 99)
Kyrgyzstan	1989	1989	Kyrgyzstan Democratic Movement	Masaliev	1	Stayed	Soviet Union	Soviet support; no specific information available (Olson, L. B. Pappas, and N. C. Pappas, 1994)
Mongolia	1989	1990		Batmönkh	1	Stayed	Soviet Union	Batmönkh was against using repression against protesters and resigned (Ellison, 1988) .
India	1921	1922	Moplah Rebellion	Hindi leaders	0		Britain	The British Government used repression to put down the rebellion.
Pakistan	1973	1977	Baluchi rebels	Bhutto (prime minister)	1	executed	Iran	
Pakistan	1983	1983	pro-dem movement	Zia-ul-Haq	0		US, Saudi Arabia	Carter cut off aid to Pakistan in 1979 after the refusal of international nuclear inspectors, but later offered military aid (Hevesi, 1988) .
Burma	1948	2006	Karens	U Nu; Thaik; Ne Win	1	Prison	Britain, China	

Continued on next page

Table 2.5 – *Continued from previous page*

Country	Start	End	Campaign	Leader	Removal	Fate	Third Party	Evidence of Third-Party Assistance (if available)
Burma	1988	1988	pro-dem movement	Lwin	1	Stayed	China	After over-throwing Lwin in a military coup, the junta nonetheless provided him with assistance, including guards, a pension, several four-wheel drive vehicles and monthly food rations. China provided aid (Chenoweth and Stepan 2011, 175).
Sri Lanka	1971	1971	JVP	Bandaranaike	0		Soviet Union, Pakistan, India,Singapore	Soviet military aid (Patrick, 2006, p. 120) .
Sri Lanka	1972	2006	LTTE	Jayewardene; Premadasa; Sirisena	0		US	Military aid (Marks and Singh Brar, 2016) .
Thailand	1970	1973	communist rebels	Kittikachorn	1	Exile	US	Massive financial aid (Tarling, 1999, p. 295) .
Cambodia	1970	1975	Khmer Rouge	Heng; Nol	1	Exile	US	Troops (Semple, 1970) .
Cambodia	1978	1997	Khmer Rouge	Samrin	0		US	US aid (Ciment, 2015, p. 533)
Laos	1960	1975	Pathet Lao	Vatthana	1	Prison	US	Vatthana was the last representative of the royal dynasty.
Philippines	1946	1954	Hukbalahap Rebellion	Roxas; Magsaysay	0			Congress allocated an undetermined amount of grant aid for the government under the Military Defense Assistance Act. (IBP, 2012)
Philippines	1986	1983	People Power	Marcos	1	Exile	US	Military and economic aid amounting to U.S. $400 million, but by 1984, U.S. President Ronald Reagan started distancing himself from the Marcos regime (IBP, 2012) .

Chapter 3

State-controlled Media and Foreign Policy: Analyzing Russian-language News

[20]

3.1 Introduction

State-owned media outlets provide authoritarian regimes with an invaluable resource that helps frame events and policy decisions. Recent research has evaluated authoritarian regimes' strategic use of state-owned media to control dissent (King, Pan, and Roberts, 2013) , or to influence public opinion during political crises (L. Windsor et al., 2017) . We build on this research to evaluate the state-owned media outlets with the goal of framing military interventions.

[20]Elena Labzina and Mark David Nieman (2017). *State-controlled Media and Foreign Policy: Analyzing Russian-language News*

Authoritarian regimes often employ state-owned media to explain and justify their domestic and foreign policies. Policy framing, however, is most effective if it is consistent with the existing conceptions of the state, its place within the larger political structure, and the pre-established social identities (C. G. Thies and M. D. Nieman; Wendt, 2017; 1994) . State-sponsored media accounts of current events generally are consistent with these pre-existing narratives. In other words, autocrats often frame their own actions and the actions of foreign states within a set of familiar tropes, which can often be condensed into characterizations such as whether the state is a *superpower*, *rising power*, *regional power*, *ally*, *rival*, or some combination.

These tropes help contextualize, justify, and make sense of either the regime's, or a rival state's, foreign policy actions. Lo (2004) , for example, contends that Russian leadership views itself as a barrier against a barbarian West, while Lavrov (2006) and MacFarlane (2006) argue that Russia considers itself a great power central to the international order with a hegemony in the post-Soviet space. From this vantage point, Russian military action in Eastern Europe fulfills the conflict management obligations of a regional or great power. These characterizations and attributed identities provide a framework that both clarifies and constrains the ways in which a government can effectively cast friendly or unfriendly foreign governments to their population.

The goal of the paper is to investigate the use of state-media to frame militarized interventions via textual analysis. Advances in textual analysis have created new research opportunities for systematic analysis of mass media coverage, including manipulation of coverage for the purposes of furthering specific goals of the political elite. McManus; McManus (2014; 2016) and Andrew H. Kydd and McManus (2015) , for example, examine statements by US presidents to study how leaders signal resolve. Textual analysis has also been widely used to investigate how closed regimes convey (and sometimes manipulate) information for both domestic and foreign audiences. Lucas et al. (2016) perform a study of

leading Muslim clerics' *fatwas* to explain variation in Jihadist legal thought. L. C. Windsor, N. Dowell, and A. Graesser (2014) and N. M. Dowell, L. C. Windsor, and A. C. Graesser (2015) investigate linguistic patterns of persuasion utilized by authoritarian leaders during crises.

We use these tools to analyze the Russian government's use of state-owned media to frame the public's perception of issues and events. We argue that, in contrast to independent media outlets whose coverage is determined by occurrence/salience of events, state-owned media outlets may be used to shape public response/perception as, or even *before*, a state implements a specific policy. We develop a web-scraping application that allows us to extract and analyze vocabulary structures of both state-owned and independent Russian-language media. The goal is to explore how the Russian government projects both its self-perceptions and its perception of "the West" onto its citizens. By focusing on Russian-language sources, rather than relying on secondary sources or English translations, we directly access government efforts to frame issues for their domestic audience.

In what follows, we analyze the vocabulary structure of *Russia-24*, a state-owned news channel, and *Dozhd*, an independent news channel, in order to separate event-driven coverage from state-directed propaganda. We use a change-point model to identify shifts in the nature of Russian media coverage. We find that Russian state-owned media significantly increased its coverage of Georgia and Ukraine in the months *preceding* Russia's military interventions. This increase in coverage was often supplemented and associated with an increased focus on traditional Russian political rivals, such as the US.

In the next section, we discuss how media, and state-owned media in particular, influence political outcomes. We then develop a foreign policy theory of state behavior and media framing, and apply this account to the case of Russia. The following section discusses our data collection technique and displays preliminary evidence. Finally, we describe and present

the results of our data analysis, and conclude by summarizing the results, implications, and next steps.

3.2 Media Effects and Political Behavior

There are several accounts of how individuals process information provided be media outlets, such as Receive-Accept-Sample (RAS) (Zaller; Zaller, 1992; 1996) and motivated reasoning (Redlawsk; Taber and Lodge, 2002; 2006) . Each of these theories expect that individuals with low politically sophistication may change their attitudes when presented with a counter-attitudinal message, whereas politically sophisticated individuals mediate messages that counter their pre-existing beliefs.[21] An alternative account, selective exposure, holds that individuals selectively choose media sources that fit with their pre-existing views (Iyengar and Hahn; Iyengar, Hahn, et al., 2009; 2008) .[22] Individuals that select like-minded news are less likely to receive counter-messages that may change or mediate their attitudes, and when they do they are more likely to show hostility towards the media source (Arceneaux, Johnson, and Murphy; Vallone, Ross, and Lepper, 2012; 1985) .

State-owned media alters the manner in which individuals can process information and affect political behavior. Countries with state-owned media substantially raise the costs of individuals to engage in selective exposure.[23] Moreover, the subset of the population likely to engage in selective exposure is relatively small (Stroud, 2011) . Thus, much of the population is susceptible to media influences. In the absence of counter-balancing information and limited options for selective exposure, large segments of the population can be influenced by a centralized and orchestrated account of policy or events.

[21] While RAS expects that counter-attitudinal messages are discounted or even neutralized, conditional on the the recipient's level of political sophistication, motivated reasoning permits recipients to attitudes to change in the opposite direction of the message. See Prior (2013) for a review.

[22] Selective exposure may be mitigated by other selection criteria, such as endorsement cues or anxiety (Messing and Westwood; Valentino et al., 2014; 2009) .

[23] As detailed in the next section.

3.2.1 State-owned Media

Autocratic regimes frequently use media to cement their own political power, even as the media environment has become more commercial and less centralized since the 1990s. Despite risks associated with undermining the state capacity to respond to environmental and economic crises by restricting credible information flows of bureaucratic performance to the central regime, autocratic regimes are incentivized to restrict media freedom in order to prevent the public from knowing of policy failures (Egorov, Guriev, and Sonin, 2009) .

In an effort to gain the benefits of information without the risk of fueling discontent, autocratic regimes frequently attempt to permit some media freedom, within a specific boundaries, in order to minimize any collective action against the regime (Egorov, Guriev, and Sonin; Gunitsky; King, Pan, and Roberts, 2009; 2015; 2013) . Rather than directly managing content via government-scripted stories and pre-publication censorship, autocratic regimes have increasingly relied on self-censored media to support regime goals (Hassid; Schimpfossl and Yablokov; Stockmann and M. E. Gallagher, 2008; 2014; 2011) [24], as well as providing a higher-quality product on state-owned and state-affiliated media to maintain viewership (Moehler and Singh; Stockmann and M. E. Gallagher, 2011; 2011) .[25] Media self-censorship relates to the choice, distortion, and emphasis of stories covered (or not covered) (Lee, 1998) .

Self-censorship arises from one, or a combination, of mechanisms. First, self-censorship can result from uncertainty regarding what content may result in punishment of journalist and managers. In the case of China, for instance, reporters temper how aggressively to pursue stories because they are unsure how far they can push a story without facing harsh and arbitrary punishment (Hassid, 2008, p. 422) . Second, the government may make clear that certain topics that are taboo to cover at all. Coverage critical of Vladimir Putin, for instance,

[24]The threat of coercion, of course, is necessary in order to enforce self-censorship.

[25] Oates (2007) argues that state subsidies offer government's leverage over media coverage and content.

is not allowed on state-aligned networks (Schimpfossl and Yablokov, 2014) .[26] Finally, self-censorship can result from journalist responding to a reward structure that avoids raising content that the government considers problematic. From the perspective of the regime, an advantage of self-censorship is that each mechanism should apply to nonstate-affiliated media.

These mechanisms produce media coverage that toes the party line. At the same time, state-owned media are able to take advantage of relatively high programming quality, and subsequent viewer interest, resulting from reporter creativity.[27] Even non-political shows, but ostensibly entertainment programming, such as legal shows, can be used to push the government's agenda (Stockmann and M. E. Gallagher, 2011, pp. 453,456) .[28] Independent, private-owned media, conversely, are often disadvantaged in financing and may have lower programming quality, on average, than state-owned media outlets (Coyne and Leeson; Vartanova, 2009; 2012) .

That state-media reflect their government's positions and policies is important because the public often has higher or equal levels of trust in state-owned rather than independent media (Moehler and Singh, 2011) . One implication is that viewing independent media does not always affect public opinion when it contrasts with state-owned media (Gehlbach and Sonin; Shi, Lu, and Aldrich, 2014; 2011) . Relatively high trust, combined with the lack of well-resourced competition, mean that state-owned media can effectively relay the government's policy positions to the public.

[26]These strategies can serve as substitutes. China has consistently been one of the world's leaders in jailing journalist, but seldom executes journalist. Russia, on the other hand, has consistently seen the most journalists killed (Committee to Protect Journalists, 2017) .

[27]The existence of multiple state-aligned channels in Russia, for instance, ensures that they must compete, to some degree, for viewers. Thus, to keep viewer interest, they cannot aim solely at brainwashing, as viewers will become bored and change the channel (Schimpfossl and Yablokov, 2014, p. 308) .

[28] Huang (2015) argues that, in addition to direct attempts to influence public opinion, the constant presence of government signals strength in effort to maintain political order.

3.2.2 Foreign Policy and State-owned Media

State-owned media can be used to frame and justify a government's policy position. Consistent with China's increasingly cooperative foreign policy towards the US, the state-owned *People's Daily* was more positive in coverage towards the US than independent media (Stockmann, 2011) . State-owned media can also be used by the government to solidify or coordinate policy with domestic elites. Gunitsky (2015) argues that mass media, and more recently social media, is used to coordinate the interests of central and local elites. He notes that Russia, in particular, has used mass media to central power and limit the authority of local politicians.

The ability of a regime to project policies onto its population, however, is not absolute. Even with state-owned casting policies in a favorable light, a government faces some restrictions when presenting formation to the public. For instance, though Chinese labor law has transitioned from a collective to a market-based system since the economy has been gradually liberalized since the 1970s, appeals to work rights remain a part of the Communist government's ethos. Reforms in Chinese labor law, and the related effect on workplace rights, are often cast by state-owned and state-affiliated media as a new "rights protections" to champion the laborer over exploitative employers, even if in practice cases favorable to labor are often not implemented (Stockmann and M. E. Gallagher, 2011, pp. 444-448) . States, in other words, are constrained and framed by the messages and identities that a they have previously invoked.

To account for the constraints imposed by national identities and value systems, we adopt a foreign policy approach to analyzing state behavior towards its neighbors. Foreign policy approaches consider both international and domestic features that serve to structure and constrain the choices available to a government regime (Holsti; C. G. Thies and Breuning, 1970; 2012) . Through interactions with other states, state leaders adopt perceptions of themselves and assign identities to other states (Breuning; S. G. Walker; Wendt, 2011;

1992; 1994) . The specific roles/identities that a state assigns to itself—i.e. its *national role conceptions* (NRCs)—reflect domestic elite preferences and their expectations of appropriate state behavior.[29] These adopted and assigned identities/roles provide social constraints on a state's behavior through limiting the strategies a leader can even conceive of undertaking (M. D. Nieman; C. G. Thies; C. G. Thies and M. D. Nieman, 2016; 2013; 2017) .

Governments can use state-owned media to describe itself, as well as those states identified as foreign friends and foes, as part of an effort to generate a well curated and managed message for the public. Governments use these NRCs to reify their own influence and to manage how foreign events are perceived domestically. Thus, states have an incentive to maintain NRCs and frequently use them as a point of reference for making sense of and understanding unexpected events. By operating within a foreign policy theoretical framework, we are able to understand how authoritarian regimes can raise the salience and context of specific issues or interactions with neighboring states within the broader understanding of a state's position within the international system. For instance, following the Maidan protests that resulted in the ouster of Ukraine's President Yanukovych, the Russian government described their intervention in Ukraine and support of militant forces in Donetsk and Luhansk within the broader context of a geopolitical struggle with the US, rather than simply as an attempt to maintain political control of a traditional ally.

Efforts to unite and control domestic (and sometimes foreign) populations through normative influence via mass media is a critical component of state sovereignty, and is a cheaper alternative to state-building and political control than coercion (Warren; Warren, 2014; 2015) . While independent media tend to be event- rather than policy-driven, i.e. their coverage is driven by event-occurrence and salience, state-owned media outlets may be used to shape public response/perception, even *before* a state implements a policy. State-controlled

[29]The roles that a state assigns to itself may face some domestic contestation from domestic rivals (Brummer and C. G. Thies; Cantir and Kaarbo, 2015; 2012) .

media may thus be used to re-assert identities that a government wants to emphasize. For instance, governments can justify intervention in neighboring state by invoking the role of *great power*, *protector* of a region or co-ethnics, or a regional or global *balancer*.[30] Mass communication can also be used to co-opt potential partisans in neighboring states.

As an example, following the 2013–2014 Maidan protests and subsequent departure of the pro-Russian President Yanukovych from Ukraine, Russian media repeatedly referred to a region in Ukraine—stretching from Donbass in the East, to Crimea and along the coast adjacent to the Black Sea in the South, to Transnistria in Moldova—as *Novorossiya*. In the aftermath of the Maidan uprising, Russia claimed that *Novorossiya* was historically a Russian territory—citing the conquests of Catherine the Great—and referred to the new government in Kiev as a fascist junta, attempting to invoke geographical and political divisions from World War II. Such references could potentially be used to provide justification for intervening in a neighbor in effort to protect co-ethnics, as well as attempt to assert a greater-Russian identify and generate support from those living in southern and eastern Ukraine who may view Russian news broadcasts advocating this perspective.[31]

3.3 Russian Media and Foreign Policy

In this paper, we focus on the Russian government's use of media in an effort to preemptively build support for their international conflicts. Contemporary Russia constitutes a suitable case for our analysis, as the government directly or indirectly controls much of the media landscape.[32] A privately owned independent channel, *Dozhd*, is a notable exception.

[30]While the *great power* role is well known within the international relations literature (⌈ ⌋) ⌋Chibaetal:2014, the regional and co-ethnic protector (Davis and W. Moore; Frazier and Stewart-Ingersoll, 1997; 2010) and regional and global balancer (Levy and W. R. Thompson, 2010) roles have also been invoked to explain conflict behavior.

[31]Russian media is popular in many post-Soviet states, as it frequently provides higher quality news and entertainment programming than domestic offerings (Trenin; Vartanova, 2009; 2012) .

[32]Despite the increase in television channels, news coverage is conducted primarily by three national networks, which are directly controlled by the government (Gehlbach and Sonin, 2014) . Russia's high

We argue that governments are constrained in the manner in which they can portray both themselves and others depending on pre-existing NRCs. Governments can manufacture support in the buildup to potential international conflicts by framing the coverage of the target. Thus, we expect to see an increase in the number of new stories about a target *prior* to military intervention. In the case of regional or rising powers, such as Russia, constructing support for potential conflict abroad can be relatively easy, if prior NRCs have emphasized an explicit sphere of influence or a security complex. Governments can simultaneously build nationalist sentiment and direct this energy against target states. Governments can direct state-owned media to cover issues and run stories on target states to raise the salience of these states and magnify grievances with them.

To remain consistent with a state's NRCs, governments can invoke similarities between the target of their aggression and other enemies of the state, such as other traditional rival powers. This can be done by stating that a target is in danger of potentially moving under the influence of a traditional rival. Doing so can help reduce domestic aversion to the use of force against co-ethnics. As we demonstrate, state-owned Russian media does precisely this in the months prior to intervening in Georgia and Ukraine by increasingly referencing traditional rivals, such as the US.

Russian NRCs, under President Putin, have emphasized the roles of a *regional power*, an *independent great power*, and an *aspirant global power* (Allison; Lo; Trenin, 2008; 2002; 2009) .[33] The creation of the Commonwealth of Independent States (CIS) and Eurasion Economic Union (EEU) have even been described as evidence of *neoimperial* or *hegemonic* role in the post-Soviet space on the part of Russia (Frazier and Stewart-Ingersoll; Rahr,

degree of media control is consistent with theories of information communication and resource allocation ([[)]Egorovetal:2009.

[33] MacFarlane (2006) suggests that Russia, rather than being an emerging power, seeks to reverse its decline and return to *great power* status.

2010; 2007) . In addition, since at least 2006, the Russian government has once again viewed the US as its primary geopolitical rival (Rahr, 2007) .

Given these NRCs, one would expect Russia to have an interventionist foreign policy towards neighboring states (C. G. Thies and M. D. Nieman, 2017) . Russia is directly or indirectly involved in a number of outstanding disputes: the status of Abkhazia and South Ossetia with Georgia, the annexation of Crimea with Ukraine, the status of Transnistria with Moldova, a border dispute with Estonia, and possession of the Southern Kurils with Japan. Russia has also expressed concern about the welfare of Russian-speaking populations in Estonia, Latvia, Lithuania, Moldova, and Ukraine. Russia has been invited to send troops to help defend a neighbor's territorial integrity in the cases of Tajikistan and Uzbekistan (Gibler and Sewell, 2006) . Lastly, Russia has also been involved in managing the conflict in Nagorno-Karabakh region between Armenia and Azerbaijan and, more recently, in Syria.

To justify such an interventionist foreign policy, Russia has framed many of these disputes and actions as part of a larger conflict between itself and the US. NATO expansion, for example, has been argued to serve a destabilizing influence on Russia, especially as Russia is increasingly dissatisfied with its global status and faces economic difficulties (Gibler, 1999) . It is therefore unsurprising that Russia viewed and framed the conflict with Georgia within the broader context of US efforts to undermine Russia's great power status by its involvement in the colored revolutions and NATO expansion (Tsygankov and Tarver-Wahlquist, 2009) . Similarly, Russian officials accused the US of aiding and orchestrating the Maidan protests, and used this claim to justify potential intervention in the crisis (Higgins and Baker, 2014) .

We focus our analysis on a time series of Russian media coverage of Ukraine (and, more specifically, Crimea) and Georgia. We select these cases because of our focus on *changes* in Russian state-owned media coverage. Both the Ukrainian and Georgian cases experience interventions in the time series of Russian news coverage in the form of Russian military involvement. Russian interference in Ukraine began with the appearance of "little green men"

on 27 February 2014. After initial denials, Russian President Putin admitted that the forces were, in fact, Russian troops (S. Walker, 2015) . This military intervention culminated in the formal annexation of Crimea by Russia on 18 March 2014. In addition, Russian military "tourists" entered Eastern Ukraine in March 2014 (S. Walker, 2015) . In the case of Georgia, Russia sent military forces into South Ossetia and Abkhazia, two breakaway regions supported by Russia, on 8 August 2008 and 9 August 2008, respectively. From a methodological perspective, these interventions indicate that there are periods of stability in the relations of both the Ukrainian and Georgian governments and Russia, punctuated by significant animosity. The interventions may lead to variation in Russian media coverage of each neighboring state.

To summarize, we expect to observe increasing new stories involving both the target of Russian aggression, as well as its traditional rivals, by state-owned media *prior* to the onset of military intervention. In contrast, independent media is more events-driven and fluctuation in coverage should coincide more directly with events on the ground. In particular, increases in state-owned media coverage should precede the Russian military interventions in Crimea on 27 February 2014 and Eastern Ukraine in March 2014. We also expect increases in media coverage prior to the Russian intervention into Georgia on 8 August 2008. These expectations lead to the following hypotheses:

Hypothesis 1: Ukraine and Georgia will feature more frequently in Russian state-owned news stories prior *to military intervention.*

Hypothesis 2: The US will feature more frequently in Russian state-owned news stories prior *to Russian intervention in Ukraine and Georgia.*

3.4 Analyzing Russian Media

We focus on Russian-language state-owned media outlets to analyze the Russian government's manipulation of news coverage aimed at maintaining consistency with pre-existing NRCs. By analyzing primary sources in their native language, we can assess the impact of Russian media coverage more directly than if we relied on secondary sources compiled by academics, or English-language state-owned media, such as *Russia Today*. Secondary sources may reflect *ex post* rationalizations or critiques in light of subsequent events, while English-language outlets may present an account intended for foreign, rather than domestic, audiences.

We analyze Russian media texts using a supervised learning method (Grimmer and Stewart, 2013) . We count the frequency of key words in news headlines to identify the salience of that subject (Hopkins and King, 2010) .[34] We assume that the more frequently a key word appears in the headlines of state-owned media, the more salient the government wants the topic to be to the public.

Our corpus of interest are new headlines from Russian media websites. The frequency analysis of the proper names in the headlines of the Russian news channels is based on two sources. The first one is the key state-controlled *Russia-24*. Russia-24, formerly *Vesti1* ("news" in Russian), is a state-owned news channel. Russia-24 promotes the official position of the Russian government and serves as one of the Kremlin's propaganda devices that serves to shape the public opinion. Though banned in Ukraine, Moldova, and Latvia, the network is popular among Russian-speaking minorities in several neighboring formerly Soviet republics.

The second source is the independent private Russian channel *Dozhd* ("Rain" in Russian). As an independent media outlet, Dozhd is not privileged to private government information. Thus Dozhd provides purely events-driven news coverage and functions as a placebo to the state-owned news coverage in our analyzes.

[34] Hopkins and King (2010) note that observing counts of key words have consistently been sufficient to extract substantively meaningful features of texts ([[) 273]GrimmerStewart:2013.

We collected the data with a specifically developed web-application that scraped all news headlines from the news archives available on the web pages of Russia-24.[35] When developing this application, we first identified the HTML-tags referring to the main headlines in the source code of the web pages. Next, the application collected all headlines going back as far temporally as data are available. The application then performed a stemming procedure to identify personal and geographical names: first, we eliminated all non-proper names. Then, we taught our application to treat all inflections of the same proper name as the same word. This was important as the Russian language has a highly inflectional morphology: for example, the inflections of Москва – Москве, Москвы, Москвой, Москву – are translated as the same word *Moscow*(noun) in English. For the case of the adjective meaning of *Moscow*, the inflections of московский – московского, московским, московскому, московская, московской, московское, московские, московскими, московских.

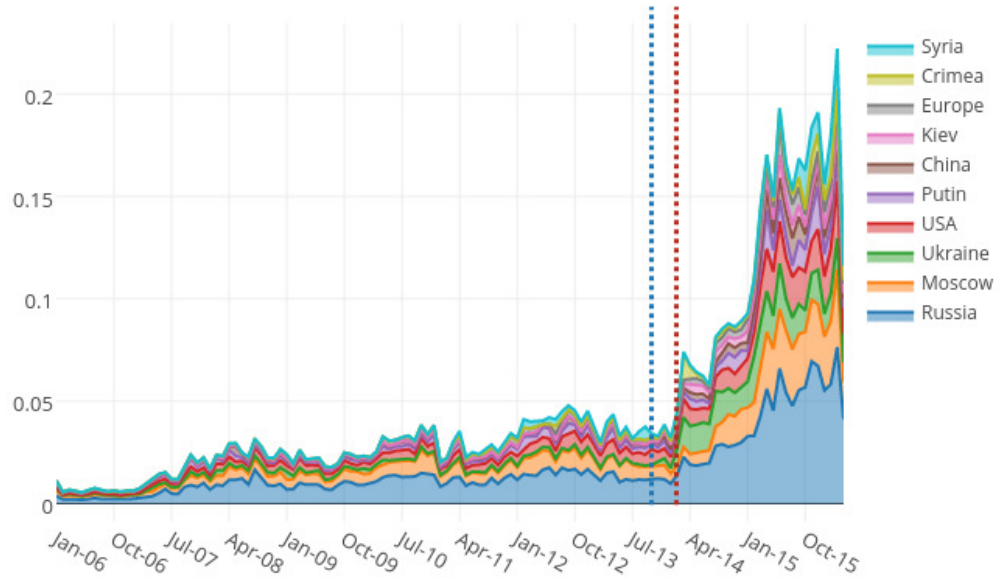
We examine a total of 585,615 headlines from Russia-24 for the period 1 January 2006 to 30 April 2016 (4,355,263 words). We review a total of 116,895 headlines from Dozhd between 24 February 2011 to 20 September 2016 (1,048,362 words). Data were collected at the daily level, but we aggregate the word counts to the monthly level for our analysis in order to include other covariates.

3.4.1 Changes in Russian News Coverage

We begin with a preliminary examination of the raw data. Figure 3.1 displays the top ten topics (proper names) present in Russia-24 headlines from 1 January 2006 to 30 April 2016. The two most common words, unsurprisingly, are *Russia* (light blue) and *Moscow* (orange), with *Putin* also featuring prominently. More interestingly, the use of the third most popular word, *Ukraine*, significantly changes in frequency during the period under review. It increases dramatically in the winter of 2014 during the Maidan protests before gradually declining in

[35]Russia-24: `http://www.vesti.ru/news/` and Dozhd: `https://tvrain.ru/news/`.

Figure 3.1: Word Frequencies–Top 10 from Russia-24

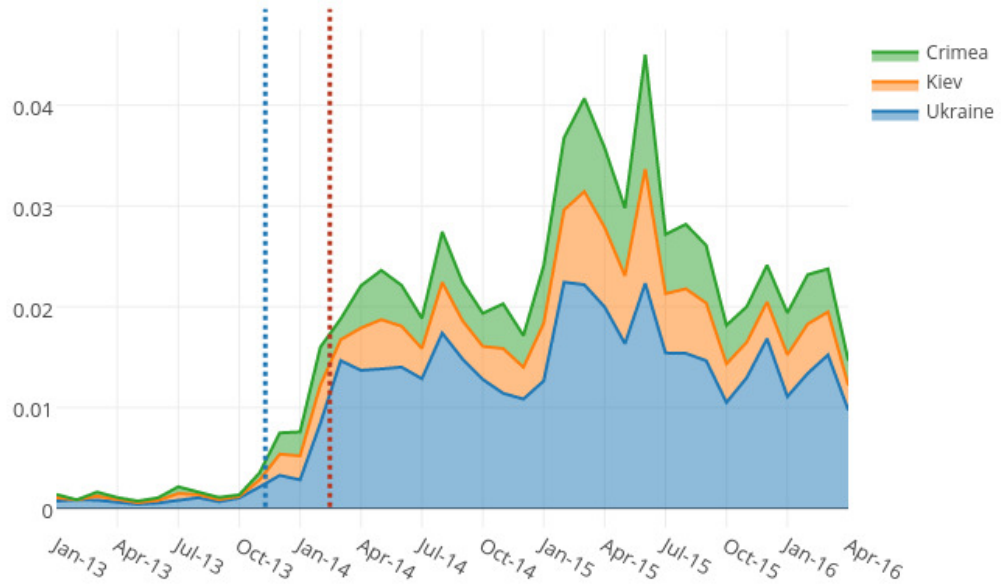


Note: The blue vertical line indicates the start of Maidan and the red vertical line the annexation of Crimea.

2015. Other notable topics include states with a sometimes adversarial role, such as the *US* and the *EU*.

Figure 3.2 focuses on frequencies of the use of the words *Ukraine*, *Kiev*, and *Crimea*. We focus on *Ukraine* and *Kiev*, as the the frequency of these words may be expected to increase in response to the Maidan protests, given Russia's long-standing interest in, and proximity to, Ukraine. We focus on *Crimea* because it was occupied and annexed by Russia. Examining the frequency of these specific words also let us compare media focus on *unexpected* compared to *expected* events. The Maidan protests and the eventual overthrow of the pro-Russian Yanukovych regime, for example, would qualify as unexpected events, as the Russian government had no hand encouraging either event. The invasion of Crimea, on the other hand, is an expected event, as a military invasion requires premeditation on the part of the

Figure 3.2: Word Frequencies–Ukraine from Russia-24



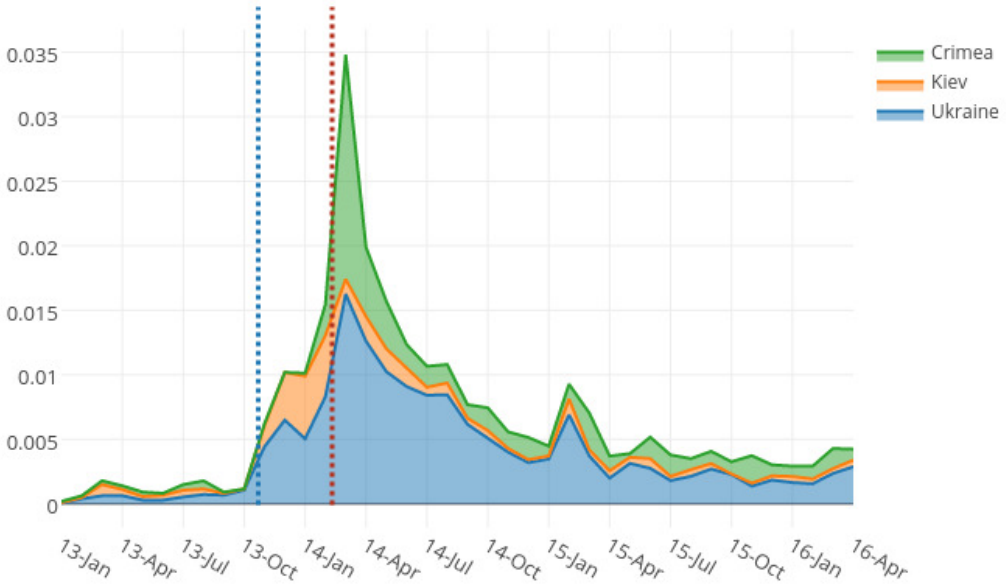
Note: The blue vertical line indicates the start of Maidan and the red vertical line the annexation of Crimea.

government. In contrasted to unexpected events, state-owned media coverage may presage expected events, as the government attempts to preemptively frame how the public will respond.

A preliminary analysis of the data reveals that frequencies of *Ukraine* and *Kiev* increase at the same time as the start and continuation of the Maidan protests, as would be expected with news coverage of unexpected events. Frequencies of *Crimea*, on the other hand, begin to dramatically increase in February, the month *prior* to the bulk of the Russian invasion and two months prior to the formal acknowledgment of Russian forces and subsequent annexation. Frequencies of the US and EU also increase at this time.

As a point of comparison, Figure 3.3 displays the word frequencies for *Ukraine*, *Kiev*, and *Crimea* from Dozhd. Frequencies of *Ukraine* and *Kiev* increase with events in Maidan

Figure 3.3: Word Frequencies–Ukraine from Dozhd.



Note: The blue vertical line indicates the start of Maidan and the red vertical line the annexation of Crimea.

and the removal of the Yanukovych regime. Frequencies of *Ukraine*, *Kiev*, and *Crimea* all peak in March, reflecting the timing of actual events in the conflict.

These results provide initial support for our theoretical expectation. Yet, a simple time series counting the frequency of our key words does not tell us if their observation is systematic or if other factors can explained by other factors. The counts could result, for example, from events-driven coverage, rather than government-directed coverage. In the next section, we systematically analyze the data using a change-point model to identify sudden changes in its underlying data generating process. In order to separate event-driven coverage from government-directed coverage, we control for Dohzd, an independent news source, when estimating counts of the frequency of key words. By controlling for event-driven news coverage, we are able to identify when state-owned media begin to focus on specific

topics. We expect that state-owned media will increase coverage their coverage of the targets of Russian aggression prior to the onset of military action, as the government attempts to rally support and justify its behavior.

3.5 Data Analysis

In contrast to events-driven coverage of independent media, we expect that the Russian government increases state-owned media coverage of target states prior to military intervention. We also expect that state-owned media increases its coverage, and hence the salience, of auxiliary issues—such as as its rivalry with the US—to frame the imminent militarized conflict within the structure of existing NRCs.

In order to identify sudden changes in the structure of the data, we employ an endogenous Baysian Markov chain Monte Carlo Poisson Change-point model (Chib; Park, 1998; 2010) .[36] The estimator is appropriate because (1) our variable of interest is a count of the frequency of our key words in a given month and (2) it is designed to identify structural breaks in the underlying data generating process of the frequency of word counts. A change-point model is a type of time series model where estimates are generated from multiple temporal regimes (or states) and the primary quantities of interest are the number and timing of changes in temporal regimes (Martin, Quinn, and Park, 2011) . In other words, the goal of the change-point model is to identify the optimal number of sub-periods within a time series, and estimate parameters within each sub-period. One benefit of the Bayesian change-point model is that, in addition to identifying sub-periods, it provides measures of uncertainty regarding

[36]Change-point models have previously been used to identify time-dependent effects for predictors of militarized conflict (M. D. Nieman; Park; C. G. Thies and M. D. Nieman, 2016; 2010; 2017) .

their specific timing, i.e. are the sub-periods distinct or do they gradually change from one sub-period to the next.[37]

Identifying the *timing* of change-points allows us to evaluate our first hypothesis concerning changes in the number of stories about a target prior to Russian military intervention. Parameters on covariates are estimated within each of these temporal regimes and are free to vary in both effect and direction between them.[38] Examining the effect of there covariates allows us to assess our second hypothesis concerning whether mentions of the US are associated with increases in the number of stories of a target state during a temporal regime in which a militarized conflict occurred.

The estimator is a Markov model with hidden states and restricted transition properties (Chib; Chib, 1996; 1998) . Changes from latent states follow a first-order Markov process:

$$p = \begin{pmatrix} \mathrm{p}_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & 0 & p_{m,m} & p_{m,m+1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

where $p_{i,j} = \Pr(s_t = j | s_{t-1} = i)$ is the probability of moving to regime j at time t given that the regime at time $t-1$ is i and m is the number of change-points. We specify our Bayesian Poisson change-point model as

$$\begin{aligned} y_t &\sim \mathcal{P}\mathrm{oisson}\left(\lambda_t\right), t = 1, \ldots, T \\ \lambda &\sim x_t' \beta_m, \quad m = 1, \ldots, M \end{aligned}$$

[37]Parameter estimates are drawn from the posterior distribution of the full state space, which accounts for the precision (or imprecision) of the estimated change-points on the Poisson parameter estimates (M. D. Nieman; Park, 2016; 2010) .

[38]The posterior sampling distribution of a Poisson with covariates does not adhere to a known conditional distribution. Fruhwirth-Schnatter and Wagner (2006) develop a technique taking the logarithm of time between successive events to transform the Poisson regression into a linear regression with log exponential (1) error.

and assume priors for the parameter estimates for the Poisson and for temporal regime transitions of

$$\begin{aligned} \beta_m &\sim \mathcal{N}(0, 10), m = 1, \ldots, M \\ p_{ij} &\sim \mathcal{B}\text{eta}\left(\alpha, \beta\right), m = 1, \ldots, M \end{aligned}$$

To recover estimates for the model, we run 20,000 MCMCs, after discarding a burnin of 10,000 MCMCs.

We use Bayes Factor to assess model fit and identify the optimal number of change-points (Chib; Park, 1996; 2010) . Bayes Factor compares two models, treating one as the baseline model and the other as an alternative model. That is, $BF_{ij} = \frac{m(y|\mathcal{M}_i)}{m(y|\mathcal{M}_j)}$ where $\mathcal{M}_i$ is the baseline model, $\mathcal{M}_j$ is the alternative model, $m(y|\mathcal{M}_i)$ is the marginal likelihood under $\mathcal{M}_i$, $m(y|\mathcal{M}_j)$ is the marginal likelihood under $\mathcal{M}_j$. Taking the log of the Bayes Factor, negative values are interpreted as evidence against the baseline model and positive values are evidence in favor of the baseline (Gill, 2009, p. 209) .[39]

Finally, we include several other variables in the model that may influence the frequency of our variables of interest and that may affect the timing of a structural break. We include counts of the frequency of *US* and *EU* in the Poisson model. We also include the level of inflation that Russia experiences in a month, as this the state of the economy may encourage diversionary stories from the state-owned media. Monthly inflation data are obtained from the OECD (2016) . When appropriate, we include a control variable for the period including the Maidan protests. This accounts of the instantaneous effect of the protests. Any long-term effects will be captured by a structural break leading to a new temporal regime. Lastly, in both the *Ukraine* and *Crimea* time series, we include a control variable accounting for

[39] Kass and Raftery (1995) provide additional guidelines for interpreting Bayes Factor comparisons; values between 3-20 offer some support for the baseline model while values >20 provide strong support for the baseline model.

observations of the dependent variable from the independent media outlet Dozhd.[40] The inclusion of this variable explicitly separate 'events on the ground' from propaganda, in that independent media can only report what is observable while not being privy to government plans. This variable cannot be included in the Georgian time series, however, as the Dozhd data are not available at the time of Russian military intervention.

3.6 Empirical Results

We analyze three different dependent variables: the frequency of *Ukraine*, the frequency of *Crimea*, and the frequency of *Georgia* in Russia-24 headlines. We use a Poisson change-point model to identify when there is structural change in the frequency of our dependent variable. In each case, we begin by using Bayes Factor to identify the number of change-points that best fit the data. Next, we visualize the posterior density for any identified temporal regimes, as well as posterior probability density of a change-point in a given year, to identify the timing of structural breaks. Lastly, we report summaries of the posterior parameter estimates for each endogenous variable within each identified temporal regime.

3.6.1 Ukraine

We begin by identifying the optimal number of change-points/temporal regimes present in the data. We do this using Bayes Factor to compare the marginal likelihood across models with m change-points. These results are reported in Table 3.1. They indicate that $\mathcal{M}_1$ has the best model fit.[41] We now focus on the results from this model.

[40] As Dozhd has a larger number of total headlines on a monthly basis, we include the dependent variable as a proportion of all headline words. An advantage of this approach is that it eases calculations of the marginal likelihood. Parameter estimates are similar regardless of whether the count or proportion measure is used. The correlation of proportion and count measures are greater than $r = 0.9$ for *Ukraine* and *Crimea* for both Russia-24 and Dozhd.

[41] The results from the Bayes Factor suggest that $\mathcal{M}_1$ and $\mathcal{M}_2$ are difficult to separate. Both $\mathcal{M}_1$ and $\mathcal{M}_2$, however, identify a change-point in February 2014 and provide nearly identical parameter estimates for the post-February time regime. See appendix for these results.

Table 3.1: Bayes Factor Comparison of Poisson Change-point Models for Frequency of *Ukraine* in Russia-24 Headlines.

Log(Bayes Factor)	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$
$\mathcal{M}_0$	0	-1833.35	-1832.71	-455.20	-418.8
$\mathcal{M}_1$	1833.35	0.00	0.64	1378.10	1414.53
$\mathcal{M}_2$	1833.35	-0.64	0.00	1377.46	1413.89
$\mathcal{M}_3$	455.20	-1378.10	-1377.46	0.00	36.40
$\mathcal{M}_4$	418.8	-1414.53	-1413.89	-36.40	0.00

Note: $\log\Big(BF_{ij} = \frac{m(y|\mathcal{M}_i)}{m(y|\mathcal{M}_j)}\Big)$ where BF_{ij} is the Bayes Factor comparing model $\mathcal{M}_i$ to model $\mathcal{M}_j$ Columns are $\mathcal{M}_i$ and rows are $\mathcal{M}_j$.

Figure 3.4: Identifying Change-points in the Frequency of *Ukraine* in Russia-24 Headlines.

Change-point in February 2014. Local means: 38.1, 236.4.

Figure 3.4 presents the posterior density for the three temporal regimes, as well as posterior probability density of the change-point in a given year. Figure 3.4 identifies a change-point in February 2014. The results indicate that there is a structural break in the data between January and February 2014. This structural break suggests that the determinants of how frequently *Ukraine* appeared in news headlines changed. Figure 3.4 also indicates that this break occurred quite sharply. The local means between the temporal regimes increase dramatically (38.1 compared to 236.4). A change-point at this location, along with an increase in the frequencies of *Ukraine* appearing in the headlines, is consistent with our expectations expressed in hypothesis 1, that Russian state-owned news stories feature target states more frequently *prior* to military intervention. In other words, Russian state-owned media increased their news coverage before their military intervention occurred.

Lastly, Table 3.2 displays the summaries of the posterior distribution for each estimate associated with the covariates. Of primary interest is the regime following the February 2014 change-point—the second temporal regime. The coefficient on *US* is positive and significant at traditional levels, after being negative in the first temporal regime. It is worth noting

Table 3.2: Parameter Estimates from Poisson Change-point Model for Frequency of *Ukraine* in Russia-24 Headlines.

	Feb 2011– Feb 2014	Feb 2014– Apr 2016
Variable	Mean/SD	Mean/SD
Mean/SD		
Count US	-0.0027	0.0008
	0.0009	0.0004
Count EU	-0.0008	0.0002
	0.0012	0.0006
Inflation	-0.0318	-0.0227
	0.0181	0.0040
Maidan	1.0338	-0.1019
	0.0725	0.0672
Dozhd	0.5309	36.8044
	3.1904	2.5092
Constant	4.1726	5.3283
	0.1954	0.1023

Note: Mean and standard deviation estimates drawn from the posterior distribution. 20,000 MCMC were run after discarding a 10,000 burnin. Parameter estimates of the posterior sample from the full state space, accounting for the precision (or imprecision) of the estimated change-points.

that these results hold after accounting for expected media coverage related to the events taking place in a neighboring state (via the independent news *Dozhd* variable). Overall, these results offer support for hypothesis 2, which expected that stories featuring the US would be associated with increases in headlines of Ukraine.

Table 3.3: Bayes Factor Comparison of Poisson Change-point Models for Frequency of *Crimea* in Russia-24 Headlines.

Log(Bayes Factor)	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$	$\mathcal{M}_4$	$\mathcal{M}_5$
$\mathcal{M}_0$	0	-1282.0	-1749.7	-1894	-1608.6	-1652.1
$\mathcal{M}_1$	1282.0	0	-467.6	-612.0	-326.6	-370.1
$\mathcal{M}_2$	1749.7	467.6	0.0	-144.2	141.1	97.6
$\mathcal{M}_3$	1894.0	612.0	144.2	0.0	285.3	241.8
$\mathcal{M}_4$	1608.6	326.6	-141.1	-285.3	0.0	-43.5
$\mathcal{M}_5$	1652.1	370.1	-97.6	-241.8	43.5	0.0

Note: $\log\Big(BF_{ij} = \frac{m(y|\mathcal{M}_i)}{m(y|\mathcal{M}_j)}\Big)$ is the Bayes Factor comparing model $\mathcal{M}_i$ to model $\mathcal{M}_j$ Columns are $\mathcal{M}_i$ and rows are $\mathcal{M}_j$.

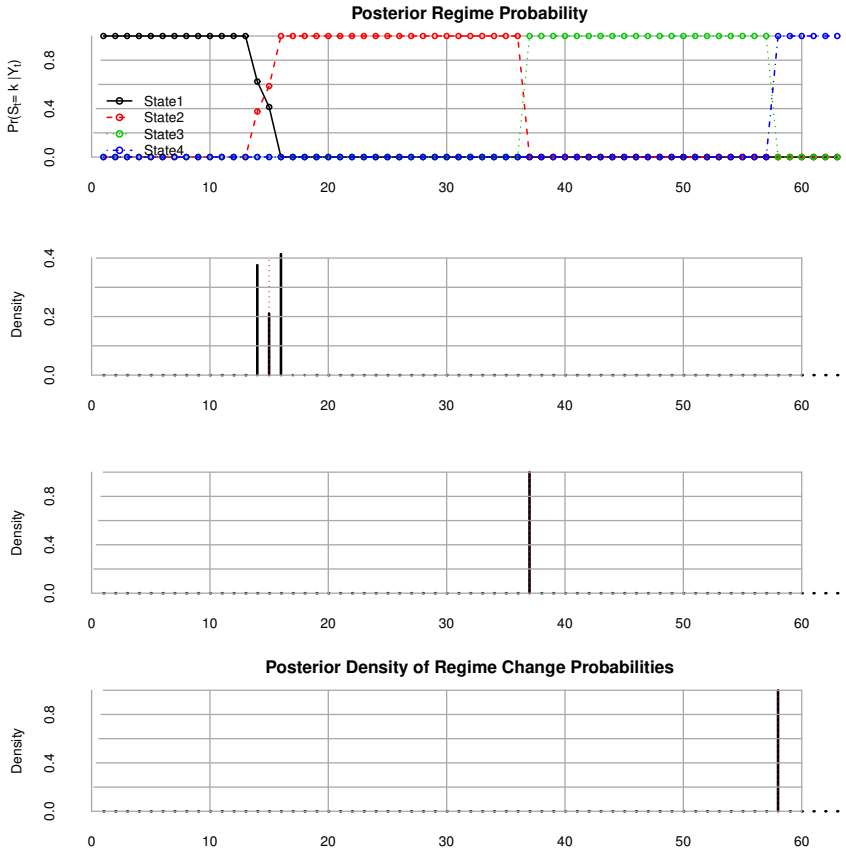
3.6.2 Crimea

Next, we turn our analysis to exploring changes in the frequency of observing *Crimea* in Russian news headlines. We use Bayes Factor to identify the change-point model with the best fit. Table 3.3 reports the model comparisons and indicates that $\mathcal{M}_3$ has the best model fit to the data.

Figure 3.5 presents the posterior density for the four temporal regimes and the posterior probability density of the three change-points. Change-points are identified in April 2012, February 2014, and November 2015. The structural break in February 2014, the increase in local means between the second and third temporal regimes (10.5 and 68.8), and the sharpness of the break, is consistent with hypothesis 1. That is, there appears to be evidence that Russian state-owned news began to feature stories about Crimea in the lead up to military intervention.

Finally, Table 3.4 displays the summaries of the posterior distribution for each covariate. Our focus is on the third temporal regime, which follows the February 2014 change-point. The coefficient on *US* is positive and significant at traditional levels, after being negative in the previous temporal regime. The results control for the expected media coverage from an

Figure 3.5: Identifying Change-points in the Frequency of *Crimea* in Russia-24 Headlines.



Changes in April 2012, February 2014, and November 2015. Local means: 1.0, 10.5, 68.8, 64.5.

independent outlet (via the *Dozhd* variable). The results are consistent with those which focused on *Ukraine*, and again offer support for hypothesis 2.

3.6.3 Georgia

In our last set of analyses, we focus on changes in the frequency of observing *Georgia* in Russian news headlines. We identify the change-point model with the best fit using Bayes Factor. Table 3.5 reports the model comparisons and indicates that $\mathcal{M}_2$ has the best model fit to the data.

Table 3.4: Parameter Estimates from Poisson Change-point Model for Frequency of *Crimea* in Russia-24 Headlines.

	Feb 2011– Apr 2012	Apr 2012– Feb 2014	Feb 2014– Nov 2015	Nov 2015 Apr 2016
Variable	Mean/SD	Mean/SD	Mean/SD	Mean/SD
Mean/SD				
Count US	-0.0135	-0.0205	0.0208	0.0052
	0.0145	0.0059	0.0010	0.0017
Count EU	-0.0388	0.0597	0.0086	0.0078
	0.0155	0.0074	0.0014	0.0042
Inflation	-0.1708	0.2091	-0.0778	0.2119
	0.1857	0.3243	0.0088	0.0252
Dozhd	0.0100	0.3423	6.2267	-0.0127
	3.1584	3.1670	3.0878	3.1597
Constant	6.0511	0.4909	-0.0513	0.2853
	2.4156	1.5503	0.2649	0.6348

Note: Mean and standard deviation estimates drawn from the posterior distribution. 20,000 MCMC were run after discarding a 10,000 burnin. Parameter estimates of the posterior sample the full state space, accounting for the precision (or imprecision) of the estimated change-points.

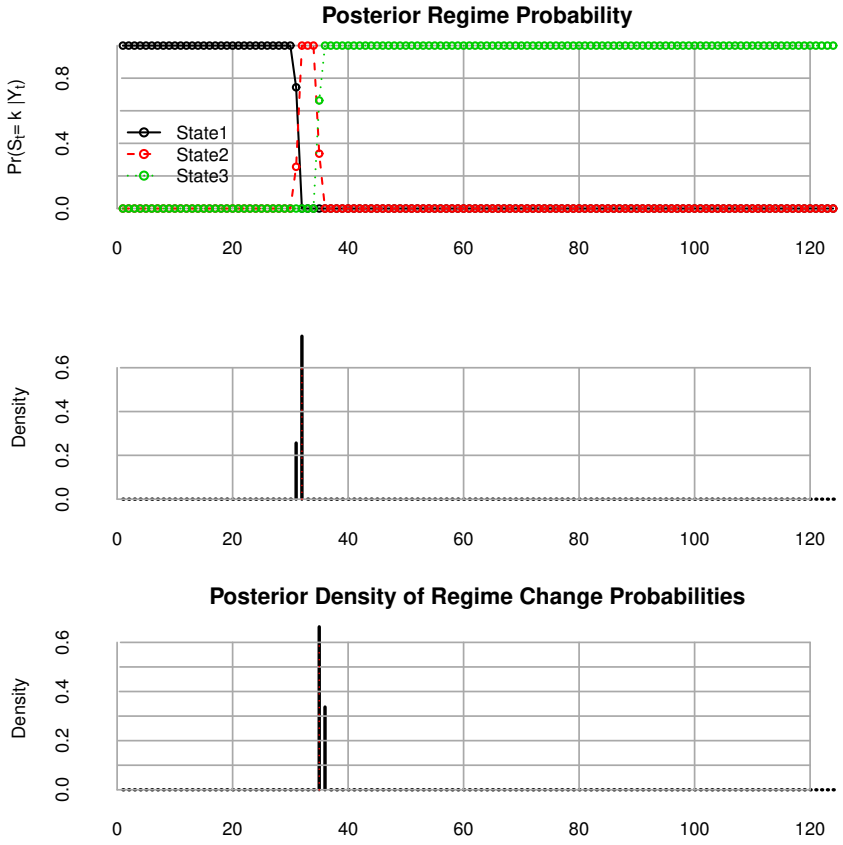
Table 3.5: Bayes Factor Comparison of Poisson Change-point Models for Frequency of *Georgia* in Russia-24 Headlines.

Log(Bayes Factor)	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_3$
$\mathcal{M}_0$	0.0	-387.1	-465.0	-416.4
$\mathcal{M}_1$	387.1	0.0	-77.9	-29.2
$\mathcal{M}_2$	465.0	77.9	0.0	48.7
$\mathcal{M}_3$	416.4	29.2	-48.7	0.0

Note: $\log\Big(BF_{ij} = \frac{m(y|\mathcal{M}_i)}{m(y|\mathcal{M}_j)}\Big)$ where BF_{ij} is the Bayes Factor comparing model $\mathcal{M}_i$ to model $\mathcal{M}_j$ Columns are $\mathcal{M}_i$ and rows are $\mathcal{M}_j$.

Figure 3.6 presents the posterior density for the three temporal regimes and the posterior probability density of the two change-points. Change-points are identified in Aug 2008 and Nov 2008. The structural break in August 2014 is the same month as Russia's intervention.

Figure 3.6: Identifying Change-points in the Frequency of *Georgia* in Russia-24 Headlines.



Changes in Aug 2008 and Nov 2008. Local means: 33.5, 183.3, 14.0.

The increase in local means, from 33.5 to 183.3, and the probability of the break, with a 30% chance that it occurred in July, indicating that increases in Russian news coverage began in mid to late July, is consistent with hypothesis 1. That is, there appears to be evidence that Russian state-owned news began to feature stories about Georgia prior to the military crisis and intervention in the break-away republics of Abkhazia and South Ossetia.

Finally, Table 3.6 displays the summaries of the posterior distribution for each covariate. Our focus is on the second temporal regime, which follows the August 2008 change-point. The coefficient on *US* is positive and significant at traditional levels, after exerting a negative effect during the previous temporal regime. These results again offer support for hypothesis 2.

Table 3.6: Parameter Estimates from Poisson Change-point Model for Frequency of *Georgia* in Russia-24 Headlines.

	Jan 2006– Aug 2008	Aug 2008– Nov 2008	Nov 2008– Feb 2014
Variable	Mean/SD	Mean/SD	Mean/SD
Count US	-0.0027	0.0214	-0.0018
	0.0018	0.0126	0.0008
Count EU	0.0112	-0.0899	-0.0125
	0.0036	0.0214	0.0013
Inflation	0.2065	0.5071	0.0629
	0.0197	0.3704	0.0096
Constant	1.2195	-2.8420	3.1620
	0.1744	4.3175	0.1446

Note: Mean and standard deviation estimates drawn from the posterior distribution. 20,000 MCMC were run after discarding a 10,000 burnin. Parameter estimates of the posterior sample the full state space, accounting for the precision (or imprecision) of the estimated change-points.

3.7 Conclusion

We argue that authoritarian governments use media to build support in the lead up to foreign intervention, beyond what events-driven coverage would entail. Further, we contend that state-owned media will add context to this by linking target states to other, traditional enemies. We develop a web-scraping application to investigate and analyze Russian-languange media coverage using textual analysis (over 700,000 news stories) of Ukraine and Georgia from 1 January 2006 to 30 April 2016. We analyze this data for changes in news coverage using an endogenous Bayesian MCMC Poisson change-point model.

We find evidence that Russian state-owned media does initiate coverage of targets in the months preceding military interventions. Specifically, we find that Russian state-owned media increased the number of headlines with *Ukraine* and *Crimea* prior to the onset of

the March 2014 military invasion of Crimea. We also find that Russian state-owned media increased coverage of *Georgia* before the August 2008 military invasions in South Ossetia and Abkhazia. In both cases, state-owned media coverage intensified *prior* to the onset of actual fighting; this is in contrast to the events-driven coverage from independent media, which increased at the same time as military intervention. Further, there also evidence that Russian state-owned media is more likely to discuss target states as it invokes a traditional rival, the US. These findings are consistent with our expectations. The results indicate that the Russian government directs state-owned media to manufacture support for its foreign military adventures. In addition, it frames these conflicts in a manner consistent with a pre-conceived national roles; namely, as part of a broader geopolitical struggle with the US.

Our analysis contributes the growing literature on foreign policy approaches and role theory by analyzing primary data sources to evaluate hypotheses about a state's national role conceptions. More broadly, our approach offers an innovate way to gather and analyze data from authoritarian regimes. State-owned media can be used to formulate and test prediction about foreign and domestic identities of these states. Future research can refine these tools to zero in on specific policy areas and other authoritarian regimes. For example, while we focus on Russian-language media outlets to explore the effects of state-owned media on domestic audiences, one could analyze English-language media outlets to compare differences in issue salience between intended domestic and foreign audiences.

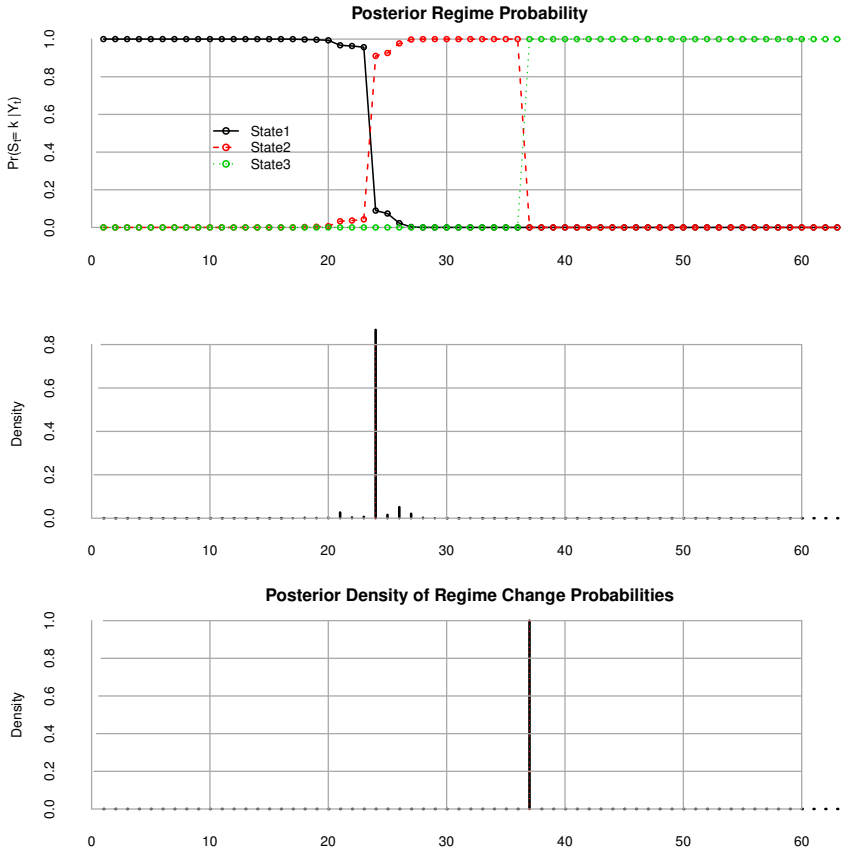
3.8 Appendix—Ukraine: 2 Change-points

Table 3.1 could not definitively say that $\mathcal{M}_1$ is more likely than $\mathcal{M}_2$ (Kass and Raftery, 1995) . We report the results from $\mathcal{M}_2$ below.

Figure 3.7 presents the posterior density for the three temporal regimes, as well as posterior probability density of the change-point in a given year. Figure 3.4 identifies change-points in January 2013 and February 2014. Consistent with the previous results (i.e. Figure 3.4), these results indicate that there is a structural break in the data between January and February 2014. Figure 3.7 also indicates that this break occurred sharply. The local means between the temporal regimes increase dramatically (39.5 compared to 236.4). This change-point and the increase in local means is again consistent with our expectations expressed in hypothesis 1, that Russian state-owned news stories feature target states more frequently prior to military intervention.

Table 3.7 displays the summaries of the posterior distribution for each covariate. Focusing on the third temporal regime—February 2014 to April 2016—it is clear that the coefficient on *US* is positive and significant at traditional levels. *US* is also significant in the second temporal regime—January 2013 to February 2014—suggesting that increasing coverage of the US by Russian state-owned media was associated with increasing coverage of Ukraine during the final year of Yanukovych's government.

Figure 3.7: Identifying Change-points in the Frequency of *Ukraine* in Russia-24 Headlines.



Change-point in January 2013 and February 2014. Local means: 37.3, 39.5, 236.4.

Table 3.7: Parameter Estimates from Poisson Change-point Model for Frequency of *Ukraine* in Russia-24 Headlines.

	Feb 2011– Jan 2013	Jan 2013– Feb 2014	Feb 2014– Apr 2016
Variable	Mean/SD	Mean/SD	Mean/SD
Count US	-0.0006	0.0052	0.0008
	0.0015	0.0024	0.0004
Count EU	-0.0028	0.0086	0.0002
	0.0013	0.0050	0.0006
Inflation	-0.0185	-0.2525	-0.0227
	0.0180	0.1411	0.0040
Maidan	-0.0118	1.3483	-0.1002
	3.1659	0.1410	0.0675
Dozhd	0.1202	0.2458	36.84861
	3.184	3.1636	2.5235
Constant	4.0084	3.3791	5.3286
	0.2288	1.1181	0.1013

Note: Mean and standard deviation estimates drawn from the posterior distribution. 20,000 MCMC were run after discarding a 10,000 burnin. Parameter estimates of the posterior sample the full state space, accounting for the precision (or imprecision) of the estimated change-points.

Chapter 4

Political Astroturfing on Wikipedia as an ideological manifestation

4.1 Introduction

On October 17, 2017, the journalists released a new investigation about the foreign political trolls' efforts to influence US politics (Meduza, 2017) . They argue that an organised group of online hackers associated with Russia not only influenced the American presidential elections in 2016, but also significantly contributed to the organization of civil unrest in Charlottesville in August 2017. Furthermore, a year ago, during the presidential debate on September 26, 2016, the topic of cyber security was already one of the core focuses of the candidates' conversation (Blake, 2016) . Clearly, the importance of the cyber aspect of the political processes and conflicts has recently become as important as probably it had never before. The control over the spread of the information is a vital component in the modern "cyber wars". In this new reality, regimes have to master new ways to deliver their messages to the public in the context of the newly emerged universe of the fluid social media.

This paper looks at a particular form of media campaigns – *online political astroturfing*. In particular, this research focuses at the Internet campaigns performed by a government in order to influence public opinion. It is organized by hiring individuals or programming bots who can mimic regular users that comment on social media or public opinion networks (Keller et al., 2017) . While being a prominent factor of today's political online discourse, *political astroturfing* is vastly under-investigated. In particular, one perspective completely overlooked so far is whether and how major powers manifest their ideology. Especially, the paper is interested in how regimes perceive their self-assigned specific roles and identities. This approach likely captures the social and psychological motivations of the regimes that may be completely neglected in the rational analysis based on power incentives that are projected in current events. As C. S. Thies and M. Nieman (2017) state in their volume, "(self-perceived) role conceptions ... refer to the expectations about a role held by a country (or at least its leaders)". Often these "self-expectations" are far less than obvious. So why are they important to understand? One main reason is that the ideology manifested through the self-perceived identity greatly influences the outcomes of a country's foreign policy, hence they have direct polical and diplomatic implications. This work argues that political astroturfing may carry an essential and unexpected element of ideology and, hence, provides a potential fruitful source for the study.

While many researchers (Hegelich and Janetzko; Sanovich, Stukal, and J. A. Tucker, 2016; 2017) study bot online users, this paper is primarily interested in the human contributors to the astroturfing campaign. In the case of the human contributors, political astroturfing imposes considerable difficulties in terms of investigation and research because of the identification obstacles that make it challenging to differentiate between false and genuine regime supporters. To overcome this challenge, this work further proposes a novel strategy that would employ the available geographical information to investigate the behavior of the political trolls

The paper argues that the salient and ever-changing nature of political discourse can translate to a historical salience of reality. This is a critical implication in terms of the agenda-based perception formation of the general public as well as the role of political trolls in perpetuating it. As Labzina and M. Nieman (2017) show, autocratic regimes portray themselves and others based on their pre-existing self-perceptions with the aid of the state-controlled media. They emphasize that regimes may use controlled media to raise political salience and awareness in the desired direction. This paper argues that political astroturfing can likewise be viewed as a political message placing Labzina and Neiman's argument in more general perspective.

As the case of the analysis, this study focuses on Russia. The reason for choosing this country is its outstanding advances in terms of developing the most institutionalized framework of political trolls in the modern world. Hence, the Russian case provides a fruitful example of the online media campaign of an entirely new type. Importantly, despite their precise empirical focus on a particular region, the results of this paper have broader implications, especially, considering the recent growth of the importance of the cyber aspect of politics.

4.2 Background and literature review

First of all, this paper contributes to the newly emerging literature in political science on *political astroturfing*, or, as it was initially named, *reverse censorship*, started by King, Pan, and Roberts (2013) in their discussion on China. Their paper argues that the Chinese censors are mainly concerned with preventing mobilization but may allow for the critique of the regime. In terms of the subject of the study, this work is closest to that of Sobolev (2017) , who discusses both human and automated cases of astroturfing on Russian *LiveJournal* and establishes that the political trolls have a limited effect on the online political discourse.

In the meanwhile, Sanovich, Stukal, and J. A. Tucker (2017) investigate Russian bots on *Twitter* and argue that it is possible to develop a framework that distinguishes them from genuine regime supporters. Similar to Keller et al. (2017) , who studies political astroturfing in South Korea, this article claims that the meta-data may be applicable for the identification argument.

Meanwhile, to illustrate how the topics and the content of the contributions produced by political trolls and their accomplices manifest the self-perception of the regime, it is foremost essential to carefully place the study within the framework of existing qualitative literature on the current Russian regime. Luckily, *Journal of Democracy* dedicates a large share of its latest October 2017 issue to *Putinism* - the political regime that evolved over the last 18 years since Putin became the Prime Minister of Russia. Six articles in a designated section "*The Kremlin Emboldened*" focus on the topic from various perspectives, such as the innate origins of (Fish, 2017) to the regime's orientation towards the future (Kara-Murza, 2017) . These publications may disagree with one another in certain aspects, but they share two common takeaway points about the modern Russian regime highly relevant to this paper: (1) its *conservatism* along with a uniquely Russian variation of "*populism*" and (2) its vastly developed "propaganda machine", which it "uses more effectively than the Soviet Union or Nazi Germany ever did" (Inozemtsev, 2017) . Likewise, another essential point this paper recognizes is that most of those articles emphasize the importance of the regime using "the Crimea" conflict as a strategy for bolstering the regime's popularity, especially given that there was an overall steep economic decline in Russia since 2014 but support for Putin grew from approximately sixty percent to more than eighty percent. While this paper looks at a very specific, politically-motivated phenomenon – political astroturfing, along with contributions resembling its identifiers on Russian Wikipedia – it is also able to bridge the general theory of National Role Conceptions (NRCs) recently developed by C. S. Thies and M. Nieman (2017) in their upcoming volume and the comprehensive in-depth study of

modern-day Russia presented by Fish (2017) , Aron (2017) , Inozemtsev (2017) , Robertson and Greene (2017) , Shevtsova (2017) , and Kara-Murza (2017) , presented in the latest *Journal of Democracy.*

In light of this, C. S. Thies and M. Nieman (2017) accomplish a tremendous task of translating how the country's history relates to the ideological regime's self-perception (see Figure 4.1). However, their classification is slightly dated and lacks relevance to the state of affairs of the 21st century. Present-day Russia is very different from its 1999 days when Putin was sworn into office. Indeed, the regime has had at least four distinct phases over the last 17 years. The first period ended around 2003-2004 with the imprisonment of the oligarch Mikhail Khodorkovsky, the Beslan massacre, the Orange Revolution in Ukraine, and the eventual elimination of the elected governors. The second interim lasted from 2004 to 2008 and was characterized by extremely high oil prices. It ended in 2008 when Medvedev came to office. During this time, the Georgian War along with the global financial crisis marked the complete end of this peaceful period. The third period occurred during Medvedev's presidency and ended with a wave of massive protests- mainly in Moscow and Saint Petersburg- when Putin declared that he was planning to run for presidency again. The last phase began in 2012, when the regime underwent tremendous transformation to a strong authoritarian leadership that valued patriotic mobilization over economic growth and which is still in effect today (Aron, 2017)). Robertson and Greene (2017) present a detailed and substantive explanation of all of these waves that characterize Russian leadership's recent ideological metamorphosis. This paper examines the very last period of Putin's Russia, particularly at its most authoritative incarnation – the time-period after the annexation of Crimea in 2014.

4.3 Research design

This work employs a mixed-methods approach by combining Big Data analysis and careful substantive evaluation. It begins by introducing the Wikipedia platform and justifying its relevance under the conditions of political salience, especially in the context of recent Russian political affairs. The next section starts with the description of the existing evidence of "the troll centers" in order to establish familiarity with the background of the topic of Russian political astroturfing in its domestic context. Then, the text proceeds to explain how select Wikipedia contributions may be linked to the available geographical information and the location of the political trolls. Further, the research substantively evaluates the identified, suspicious Wikipedia edits, first by looking at their meta-information and then by performing a detailed examination of their content for detected potential instance of political astroturfing. Combining all obtained evidence from both the geographical and substantive analyses, the first section concludes with a compiled list of detected astroturfing interventions. The goal of the following section is to extrapolate the results from the previous section and to uncover a greater number of edits similar in content to the ones already identified as astroturfing content. To do so, the paper also provides a methodology by which to compare the edits. In particular, it proposes a novel measure for the comparison of the texts - *conditional cosine similarity*. Next, the paper evaluates the content of the findings and postulates that despite a detectable level of suspicious contributions, no evidence had been found of any systematic and large-scale political astroturfing operation on Wikipedia. The paper concludes with a discussion of the findings from the perspective of NRCs of present-day Russia.

This research mainly addresses two sets of questions, those concerning the identification of political astroturfing and those dealing with the manifestation of the ideology and self-perceptions of Russia through political astroturfing:

1) Can political astroturfing be identified on Wikipedia? What are its major strategies?

2) Does the regime portray its ideology and self-perceptions though its political astroturfing? Which special insights can be obtained by analyzing the instances of political atroturfing on Wikipedia?

4.4 Wikipedia: Political and Historical Salience

This following assertion is the foundation of the main argument of this paper: that the country's history is a part of its self-perceived agenda and political astroturfing can be seen as a manifestation of its influence. Indeed, if in response to highly salient political events people's curiosity about the events' broad historical background is heightened, the coordinators of political astoturfing that manage these newly acquired perceptions of the public can anticipate the informational bias to work in benefit of their agenda. In more general terms, it simply means that in this specific population, the projected self-perceptions, an integral part of the regime's ideology, must be entrenched with historical relevance, and Wikipedia provides an exceptional chance to observe this trend.

To begin, Wikipedia, the largest collaborative network of knowledge, is the primary source of information for the information generation. Interestingly enough, its impressive coverage of numerous topics is a relatively recent phenomenon. Starting from a few thousand articles on the English language Wikipedia in the mid-2002, in 2016, the cumulative number of its articles spanning all its regional versions (nowadays, it operates in more than 100 languages) was approximately 39,721,659 [42].

Crucially, a number of reasons justify looking particularly at the Russian version of Wikipedia. Above all, the Russian segment of the Internet does not provide much alternative

[42]https://stats.wikimedia.org/EN/Sitemap.htm

to Wikipedia in the sense of obtaining general, supposedly unbiased knowledge on almost any topic. Furthermore, despite the span of more than twenty years since the collapse of the Soviet Union, Russian is still the main language that a large share of the ex-Soviet Republics population revert to in order to get access to a wider range of knowledge than is available in their own languages. The Russian Wikipedia is one of the larger Wikipedias, with the total number of 1,314,556[43]) pages. It has the third highest number of page views per hour, after the English and the Spanish counterparts. This fact is especially impressive, given that Russian language is ninth most spoken language in the world and has the seventh highest number of Wikipedia pages. Regarding editing activity, it holds a high-ranking position as well: it ranks fourth in terms of very active editors who make hundred or more edits per month and the seventh among regular but less active users who make five or more edits per month. Considering the relatively moderate number of Russian speakers in the world, this statistic shows that the usage of Wikipedia in this demographic is wide-spread.

Despite these telling statistics about Russian Wikipedia usage, so far, its relevance as a tool of investigation has been almost completely overlooked in the political science field. Yet, Russian Wikipedia is perceived by the Russian-speaking world as a trusted media source. To test this claim, this paper examines a highly salient case of the Ukrainian Crisis of 2013-present. Having originated as a domestic Ukrainian issue, a revolution as a response to the highly corrupt government of Yanukovich, the crisis slowly escalated after Russia's capture of the Ukrainian province of Crimea, which came to impact Russian citizens as well. The official position of the Kremlin was that the intervention was necessary because the Russians residing in Crimea - the majority of the population in the province - were in danger under the new "highly nationalistic government" in Ukraine (Gentleman; Lally and Englund, 2014; 2014) . According to multiple sources, Russia got involved in a foreign military conflict, often referred to as "a hybrid war" in the Eastern Ukraine, where ethnic Russians constitute

[43]May 16 2016

a significant portion of the population. Ukraine's widespread perception of these events was that Russia (both the Russian government and its citizens) betrayed them by occupying their territory and supporting the military conflict in the Eastern Ukraine (CNN, 2014) .

Figure 3.3 showcases the extent to which the topic of the Ukrainian crises occupied the news feed of the only independent Russian TV channel, *Dozhd*. Labzina and M. Nieman (2017) argue that Dozhd's coverage very accurately reflects the actual events taking place without any bias. The coverage of Ukrainian topics begins to climb steadily at the end of October before the start of the Ukrainian protest (the vertical blue dashed line) and exponentially grows just before the annexation of Crimea (the vertical red dashed line) in March. Now, let us look at how these events paralleled the page viewing activity on Russian Wikipedia.

First of all, Figure 4.2 illustrates that the amount of total page views did not exhibit any exceptional spikes or trends. The dynamics mainly reflect the seasonal trends with detectable traction. Meanwhile, the page views of the pages of the categories *Ukraine* (n=19), *Kiev* (n=3), and *Crimea* (n=11) do emphasize an interesting pattern of engagement. Until November 25th, the number of views fluctuates around its average mean of around 8,000 per day. With the start of the massive protests in Maidan, Kiev's central city square, the average slowly starts increasing and with some fluctuations, 10 days after the *Rada*[44] passes "a harsh anti-protest legislation", reaches a local apex of 26,871 views on January 26th (CSIS, 2016) . Then, after some decrease, it starts its rapid climb on February 17th, just before Yanukovich fled Ukraine. A month later on March 18th, when Putin signed a document confirming that "Crimea including Sevastopol would join Russia", the page views reached a global unprecedented engagement of 362,840. This is approximately 45 times more than before the start of the events in Maidan. This graph clearly indicates that people went online to read articles related to the ongoing political events, but did not exactly contribute to the

[44]Ukrainian Parliament

heated discourse. Hence, the numbers show that people did get curious about the history of Crimea, Ukraine and Kiev as a response to related political events.

Summing up, Wikipedia gives an exceptional opportunity to acquire general insight into the strategies and tactics of political astroturfing. Foremost, the ideological perception of Russian Wikipedia is that it houses unbiased content. Secondly, the personal component of the interaction is virtually missing: the contributors interact with the content and not with one another. Finally, regarding data, Wikipedia is extremely user-friendly and transparent: all history of the content and contributions is open and available in multiple formats. This sharply contrasts to popular user-centered unprecedented networks, where one cannot access a lot of content, most specifically, i.e *LiveJournal*, because of the privacy settings of the users and the restrictions of the network itself (i.e *Facebook*, *Twitter*) that are aimed to commercialize user data.

4.5 Identifying Political Astroturfing

Methodologically, this part of the paper describes the conservative, or, in other in other words minimizing the possibility of false positives, steps taken to identify what is further called "the initial instances" of astroturfing on the Russian Wikipedia. The major challenge of this task is the difficulty of relying solely on the content of the added or edited contributions. The content may serve as supporting evidence; however, in isolation from other criteria, it is incapable of discerning between the contribution of political trolls and genuine regime supporters. The behavioral patterns, such as the typical time when the contributions occur and the nature of their topics and content, can also be used to validate the hypothesis of astroturfing (Keller et al. 2017, working paper; Sanovich, Stukal, and J. A. Tucker 2017, working paper) but are usually not very helpful in indicating incidence of trolling at the very beginning of the astroturfing investigation process.

Acknowledging this, recent research on political astroturfing and its effects on public discourse is typically based on the exogenous (to the content) criteria, or the explicit leakage of official data. The possible documented leaks belong on the list of LiveJournal accounts of political trolls that was obtained and released by the journalists of *Novaya Gazeta* in Russia (Sobolev 2017, working paper). Likewise, a list of Twitter accounts found on confiscated computers and published during court proceedings in South Korea (Keller et al. 2017, working paper) was leaked in support of political troll presence, as was the uncovered Chinese propaganda office communication leaks from its state officials (King, Pan, and Roberts 2016).

Regarding contributions to Wikipedia, such leaks are technically impossible, since a large share of its most prolific contributors simply do not have functioning personal accounts. Indeed, data shows that although many contributors do have registered accounts (n=168,362 [45]), a significant share of the contributors do not have accounts (n = 1,563,624 [46]). The contributions of non-registered individuals are easy to spot, as they are marked with their IP addresses on an article's history page. Interestingly, some of the IP contributors are very active: 2,566 have completed more than 100 edits each, and 107 have contributed more than 1000 edits during the period from January 2013 to December 2016.

Hence, no apparent discernment of "political trolls" on Wikipedia accounts was detected. Luckily for researchers, there is a major "leakage" in traditional and social media over the last several years of the traces of the so-called "Russian political trolls". This section will start with the brief overview of these publications and, based on them, propose to investigate six possible locations of the "troll factories" employed for the purpose of skewing public perception on political and historical subjects. In some cases, the locations of these factories

[45] The total number of unique contributors with registered accounts for the period from January 2013 to December 2016.

[46] The total number of unique contributors without registered accounts for the period from January 2013 to December 2016

are provided with the exact known street address, and other times, they are known to extend into small regional settlements.

Furthermore, the section continues by describing how the IP addresses of the contributors can be connected to the geographical information of the identified troll factories. Then, the careful spatial analysis enables us to narrow down the list of suspect IP addresses to approximately ten IP ranges. The last step in the analysis of their contributions is the manual content evaluation based on specific language parameters and topic analysis elaborated on in subsequent sections (n = 88). The substantive analysis concludes that among the suspect IP addresses, two provide consistent evidence of astroturfing. Their contributions are marked as the initial astroturfing interventions. In the following section, the two IP addresses are analyzed specifically in order to find and match similar content.

4.5.1 The Troll Factories

On September 9th, 2013, the major Russian independent newspaper, *Novaya Gazeta*, famous for its field journalism, published investigation conclusions from their Saint-Petersburg correspondent (Garmajalova, 2013) . In the article, the investigative journalist describes her experience of getting a job in the so-called *Agency of Internet Research*, located in a small satellite town close to Saint-Petersburg. Later, this place would become infamous as a "troll factory", a hub of fake account holders posing as authentic individuals with the intent to promote particular sociopolitical viewpoints, or "Olgino". She finds this job through an open ad published on Russian social media networks (see the ad for the job in Figure 4.4). According to the article, her responsibilities were to write short content on various Russian online platforms discussing topics that support an assigned point of view. For example, her first assignment was to cover the G-20 Summit in Saint Petersburg, during which she had to publish around 100 comments during one business day. The journalist writes that she saw screen-shots of the published content of her predecessors who criticized America and the

Russian opposition politician Alexey Navalny (Figure 4.6 provides an example of an actual assignment performed by a political troll).

This journalism piece is crucial, since it was the first to introduce the meaning of “the troll”, “the state paid troll” or “the agents” of astroturfing in the Russian political context. This definition contrasts sharply with the previous existing notion of casual “trolling" by random individuals who share their organic perspectives. In 2012, the same newspaper published an article about the rapidly evolving phenomenon of “trolling" prevalent on the Russian Internet (Grigorieva and Chuvilyaev, 2012) . "Trolling" is ubiquitously defined as “[the] continuous generation of meaningless text that is embedded with useful information". However, there is substantial distinction between the definitions present in the articles published in 2012 and those in 2013. Above all, the 2012 content depicts “trolling" as instigated by multiple *freelance* (paid or non-paid) individuals. The article from 2012 further mentions that paid assignments may be political or commercial in nature, but it does not specify anything about “troll factories". For instance, it includes an interview with “an informed professional troll", who gave the general impression that back then, paid political trolling was more so related to "product placement" as a byproduct of mild political propaganda. Meanwhile, the article from 2013 precisely talks about institutionalized “factories of trolls”. Furthermore, in 2015 (Chen, 2015) , the New York Times Magazine dedicated a long piece to the very same “Internet Research Agency” – one of the factories of trolls – that “had industrialized the art of trolling" and that was run under management which was “obsessed with page view statistics, number of posts and the blog ranking on *LiveJournal* traffic charts" (Chen, 2015)). This specific piece talks about factories of trolls as seriously organized, well-established structures that do not have much in common with the unregulated freelance trend described in 2012 by *Novaya Gazeta*. Based on these pieces, it can be concluded that in Russia, political trolling became institutionalized somewhere between 2012 and 2013. Furthermore, in 2015 in particular, the phenomenon was already operating under an advanced infrastructure wielding its own

personnel (Chen, 2015) , which interestingly contrasts with the situation in China, where political trolling is a regular side activity of state officials (King, Pan, and Roberts, 2016) .

In the latter part of 2015, a greater number of journalist articles covering Russian political trolls emerged (BBC; Gordon, 2015; 2015) , following the empirical inquiry of a LiveJournal blogger four3. Interestingly enough, this blogger's investigation used Google Trends to show that four locations in Russia (within small indigenous populations) produced the majority of Google requests on the topic of the Ukrainian crisis, significantly outnumbering large cities like Moscow in this regard[47]. The blogger compares the geographical distributions of Google search requests for salient terms related to the controversial issues of recent Russian politics with politically neutral terms. As expected, he finds that for the non-political search requests, the major cities are on the top of the list. However, for the salient topic of Ukrainian politics, [48] the discovered results can only be explained by the existence of the Russian "11-ruble" army – the equivalent to the Chinese "50-cent" party – situated at various designated locations. Since this online tool is open to public at no cost, anyone can go to *Google Trends* and obtain the same results. Interestingly, one of the detected locations, *Olgino*, has been well-known since 2013 as a hot spot of trolls expounding one-sided political views (Garmajalova, 2013) .

Based on the analysis of the described media sources, the following section of the paper looks into six suspect locations as the possible origins of Russian political astroturfing.

[47]Perekatny, Zelenyy Gorod, Olgino, Yablonosky

[48] *Maidan* - "a city square" in Ukrainian. Recently, the term is most often used to refer to *The Square of Independence* in Kiev, Ukraine. Also, in Russian, the term became almost synonymous with the term revolution or the Ukrainian Revolution in 2014 or *Euromaidan.sanctions* - refers to the Western sanction as a response to the Russian intervention to the Crimea., *referendum* - The Crimean referendum of 2014, and *Poroshenko* - the Ukrainian president elected after the revolution in 2014.

4.5.2 IP addresses and geographical location

IP addresses carry geographic information of the user source. This research makes use of one of the most comprehensive databases that connects the IP-ranges to the exact geographic coordinates[49]. This dataset covers 11,547,800 IP ranges, each of which is accompanied by a geographic location that can be identified on the map (see Figure 4.8 for the extract of this data), [50] overall and 142,193 in Russia in particular. The analysis starts with the extraction of latitude and longitude of the suspect troll location using *Google Maps* (see Figure 4.1, Table 4.1). Next, each entry from the DB-IP is assigned the distance to each of the suspect astroturfing locations using a programming script. Additionally, another script selects all IP ranges within three miles from at least one of the suspect locations (n = 944). The position of these ranges are then imported to *Google Maps* for visual expert selection. The visual analysis is necessary because the density of the DB-IP coverage varies. It is necessary to take into account the features of the landscape while selecting the IP ranges for more detailed investigation (see Figures 4.9, 4.10, and 4.11).

Interestingly enough, three of the locations - *Olgino*, *Perekatny* and *Zelenyy Gorod* - do not have any entries from the *DB-IP* database exactly near them. Meanwhile, the other three may be linked to the database: the analysis identifies nine suspect ranges (see Table 4.1). Figures 4.9, 4.10, and 4.11 show the selection process: the stars refer to the assumed "troll factories"[51], the blue drops refer to the locations of the IP ranges in the DB-IP address dataset, and finally, the yellow speakers are the "drops" assigned to represent the "troll factories". Overall, the paper assigns *suspect* eighteen IP-ranges.

[49]The DB-IP database `https://db-ip.com/about/`

[50]*IP-range* and includes information on all IP addresses found in this dataset range. IPv4 addresses are usually represented in dot-decimal notation, consisting of four decimal numbers, each ranging from 0 to 255, separated by dots, e.g., 172.16.254.1. Each part represents a group of 8 bits (octet) of the address.

[51]In the case of Savushkina street and Varvarskaya Street, the street address is known; in the case of Yablonovsky, the analysis assumes a median position in the middle of the small settlement

The meta information on the contributions to the Russian Wikipedia is available in *xml-wiki dumps* (see Figure 4.12 for an extract). Each node (*page*) carries information about all its contribution content. Likewise, each contribution (*revision*) includes information on its author and the exact time of posting. The revisions *id* and *parentid*, along with the article's title, provide sufficient information on all the details of relevant modifications. Figure 4.13 shows the webpage and its address for one of the revisions shown on Figure 4.12.

After parsing the total of 22,674,100 contributions[52], the analysis discovers that 88 written contributions were made from one of the suspicious IPs. Interestingly, all of them are attributed to one of the two IP ranges assigned to the same location of *Yablonovsky*, a small settlement near Krasnodar with a total population of less than 2,100 people. Curiously, *Yablonovsky* was listed as a potential source of political trolling in one blogger's striking investigation using Google Trends, as described in the previous subsection. The total number of the distinct IPs among the selected is eighteen and of the distinct articles is sixty three.

4.5.3 Substantive analysis

So far, all of the identified edits only weakly confirm this hypothesis based on the available geographic information. Prior to this subsection, the research has not looked at either the time-restricted behavior patterns (the time-frame when the edits were made) nor at the topics and content of the edits. Before digging into the detailed overview of the content of the edits, let us first look at the basic summary of the Wikipedia contributions grouped by IPs provided in Table 4.2.

Meta-data

Foremost, as exemplified, thirteen of the eighteen individual contributors are responsible for editing a single page. One individual is responsible for editing two pages and another

[52] All contribution on the Russian Wikipedia during the period from January 2013 to December 2016

individual four pages. There are only three individuals in this dataset who display a relative diversity in their interests: with 18, 19, and 8 articles. Secondly, it can be concluded that out of the three active editors in the dataset, one has consistently contributed to the articles about online computer games - a topic very distant to Russian politics or history.

Meanwhile, most edits from other most active editors (with several exceptions) are made on topics that most possibly contain sensitive information related to current Russian politics and national ideology: settlements within Russia and the former Soviet Union (*Ashgabat*, *Lipetsk*, *Sochi*, *Seversky Donets*, *Donetsk (Russia)*, *Dubossary, Starominskaya, Kharkov, Kramatorsk*), Russian history (*Yuri Dolgoruky*[53], *Cossacks, The culture of ancient Russia, Russians in Kazakhstan*, *Pavlov, Sergey Vasilievich*), recent Russian wars (*Georgia; Khankarov, Khamzat Zharapovich; Semenchenko*[54], *Semyon Igorevich Semenchenko*[55]), science in Russia (*Cell theory*), and Islam and terrorism (*Charlie Hebdo*).

The interest of one of those editors in topics related to Ukraine (*Seversky Donets*, *Donetsk (Russia)*, *Dubossary*, the culture of ancient Russia, Khankarov, Semenchenko), the Chechen War (Khankarov) and Islamic terrorism (*Charlie Hebdo*) looks especially striking. Another editor mainly contributed to the articles related to the recent and ancient Russian history (Cossacks, Culture of Ancient Russia, Russians in Kazakhstan) as well as to Ukraine (Cossacks, Kramatorsk).

Based on the geographic and meta data analyses, the Wikipedia contributions from the identified editors are highly suspect to be a part of Russian political astroturfing efforts. The next step in confirming this hypothesis is to carefully evaluate the content of the contributions

[53]The founder of Moscow.

[54]Khamzat Zharapovich Khankarov (Honkarov) (April 13, 1965, Grozny, USSR - June 13, 1994, Grozny, Chechen Republic of Ichkeria) is a Chechen field commander, an active participant in the separatist movement in Chechnya.

[55]Semen Ihorovych Semenchenko is a deputy to the Verkhovna Rada of Ukraine, and the commander-founder of the volunteer territorial defence battalion "Donbas", based in Donetsk during the current military conflict in the Eastern Ukraine.

and establish whether they support the initial analysis. We can start by first comprehending the ideology and the typical pattern of behavior of Russian political trolls.

Strategies, Tactics, and Ideology of Russian political trolls

According to the leakages from the Russian "troll factories" (Chen; Garmajalova; Grigorieva and Chuvilyaev, 2015; 2013; 2012) , Russian political trolls are mostly concerned with four broad topics (*):

- *The opposition leader Alexey Navalny and the Russian political opposition in general*;

- *Ukraine and everything related to the Russo-Ukrainian conflict after the Ukrainian revolution in 2014 and the consequent annexation of Crimea*;

- *The greatness of Russia and Russians (referring primarily to Russians of Slavic ethnicity). Russia is viewed as the direct and only ethnic and cultural descendent of Kievan Rus' (the late 9th to the mid-13th century), the Tsardom of Russia (1547 – 1721), the Russian Empire (1721 – 1917), and the Soviet Union (1922 – 1991).*

- *The critique of the USA and its foreign policy.*

In particular, according to the existing journalism, social media, and academic evidence, Ukraine is the major topic of interest for Russian trolls and bots. In his New York Times article (Chen, 2015) on "The Internet Research Agency", Chen cites one of the agency employees:"*Ukraine was always a major topic, because of the civil war there between Russian-backed separatists and the Ukrainian Army; [the employee] and her co-workers would post comments that disparaged the Ukrainian president, Petro Poroshenko, and highlighted Ukrainian Army atrocities. Russian domestic affairs were also a major topic of discussion. Last year, after a financial crisis hit Russia and the ruble collapsed, the professional trolls left optimistic posts about the pace of recovery.*"

Many sources on Russian political trolls emphasize that working hours of the trolls usually range from 8am to 10pm (Garmajalova, 2013) . Hence, it is reasonable to conclude that the contributions during or close to these hours are more likely to turn out to be instances of astroturfing.

Another source is a well-known Russian blogger who provides an insider, counter perspective on troll detection: he claims that after reading more than 700,000 comments on his blog postings, he learned to detect *kremlibots* (from "Kremlin bots"), a term he uses to refer to the Russian political trolls (Chernyshov, 2014) . The blogger provides several characteristics of suspected political trolling that is related to particular language usage and which is relevant to this analysis. He claims that "*one of their characteristics is their contempt for the Russian language. This underlies that they are common people.*" By this, the blogger means that they are likely to misspell words and use vulgar slang and offensive tone (**).

Based on all the existing information about the strategies and tactics of Russian political trolls, the conclusive observation is that they are likely to employ one or a combination of two of the following strategies: direct misbehavior that is easily detectable (in the context of Wikipedia, the best way to refer to it is *vandalism* (V)), and subtle manipulation of information (in Wikipedia, this is likely to show up as an attempt to change the tone of the text through deletion (D), as well as through slight changes in wording). Likewise, suspect atroturfing contributions must also satisfy topic criteria relevant to social, economic and political affairs.

Content evaluation

Table 4.6 shows the summary of the evaluation for the contributions listed in Table 4.2. The criteria for considering whether an edit is an instance of political astroturfing is whether it falls into one of the topics presented above (*) and whether the edit parallels the ideological interests of the current Russian regime[56]. Based on this methodology, it is interesting that the same IPs the research had "suspected" based on their meta-data precisely fit the criteria. No other IP address produced anything even close to resembling astroturfing. All analyzed contributions from editors other than them were genuinely looking to improve the informational quality of the articles. Another somewhat surprising observation is the same editors are the only ones on the list who had "vandalic" contributions. Furthermore, one of the them did not produce any valuable edits at all. Also, both the suspect editors have IP addresses from the same IP range, while all others originate from other ranges.

Tables 4.4 and 4.5 present brief descriptions of the edits labeled as political astroturfing. All edits in the tables are labeled as D or V. The D stands for the "clever type" of astroturfing, when certain information pieces are simply deleted from the article yet leave it seemingly still completely legitimate. The findings of type D are especially interesting because they provide a somewhat unexpected insight into the type of information the political trolls do not wish the readers to see in the articles. They enable researchers to get a better sense of the NRCs of modern Russia (Figure 4.14 displays a typical example of D). The discussion section focuses on their content in detail. Their existence is almost surprising because political trolls are more likely to be associated with vandal type of behavior. Both of the identified editors showcase elements of type V edits as well. Furthermore, all of their type V edits involve misspelled words and/or vulgar, colloquial slang that according to Chernyshov (2014) , speaks in favor of the political trolls' authorship (Figure 4.15 displays a typical example of V).

[56]The screen-shots of all these edits are available as supplementary materials

After labeling the suspect contributions as acts of astroturfing/non-astroturfing, the last investigative component was to look at their chronological behavior. Figure 4.16 plots the times of the suspect edits. One striking observation is that except for three contributions made in the night-time there is a clear time separation between the astroturfing and non-astroturfing edits coming from these users. Also, the time-frame of the contributions looks surprising: the majority of the edits are completed within a narrow time period during weekday working days from 11am to 3pm. With one exception on October 3rd, 2016, literally all contributions are performed within the time-frame window of 7.30am to 10pm. Likewise, the working hours in the troll centers range from 8pm to 10pm. Furthermore, because of multiple changes regarding the summer/winter savings time shift in Russia since 2008, it is highly likely that the time of the edit on June 20 was misrecorded, and that it was actually made at 8.30am and not at 7.30am. Overall, the distribution of the hours of the astroturfing observations does not look random: there are 5 observations at exactly 2.30pm and 6 observations around 12pm.

4.5.4 Finding Accomplices: Extrapolation

The previous sections describe the deductive process of identifying a list of astroturfing interventions. The analysis uses the available geographical data of "the troll factories" and then examines the content of the contributions and edits coming from IPs located close to these locations. Given these evaluation parameters, the paper concludes that two IP address from the same IP range exhibit behavior consistent with the hypothesis of astroturfing. This section extrapolates the findings of the previous section to uncover more instances of similar Wikipedia contributions.

Based on the located astroturfing edits, the analysis narrows its attention to the categories that have proven to attract perpetrators of astroturfing (see Table 4.6 and 4.7).

With these categories in place to filter all Russian Wikipedia articles, the size of the number of the edits decreased from 22,674,100 to 149,894.

Most importantly, this section presents those contributions that appear very similar both in terms of the behavioral patterns they exhibit and the nature of content they showcase. Despite these analytical possibilities, it is unreasonable to claim that astroturfing content mostly derives from individuals who are state employees. Hence, this section refers to the diverse range of identified contributors as “astroturfing accomplices".

Comparison of Wikipedia Contributions

Comparing Wikipedia contributions is a tricky process. Most often, addition or deletion of content does not occur in the format of a proper sentence or sentences. It may not even be a phrase. In fact, many of such contributions consist of small fragments of text that are added or deleted in multiple locations of an article. Another challenge to detecting changes to text is how the history of Wikipedia edits is stored: the XML-wikidumps store each edited version of the article, and the discrepancies between them are not easy to capture. Luckily, the online version of the edits contains explicit indication of which parts of an article have been modified. Hence, their HTML code may be directly parsed. Most importantly, because of the fragmentary content of the edits, the best approach towards investigating their intended function is to construct a dictionary compilation with the words and their frequencies.

Finally, before the contraction of the dictionary and after the removal of stopwords from the text, there is a need for careful lemmatization, or extraction of the word stems. Russian is one of the most inflectional languages in the world. It is one of the Indo-European languages that retained a lot of its word ending flexibility from its proto-language, ancient Sanskrit. For example, all these words - *московский, московского, московским, московскому, московском, московская, московскую, московской, московское, московские, московских, московских, московскими* - are translated as *Moscow* (adjective) in English.

To perform the lemmatization or stemming, this paper utilizes the most efficient stemming tool for languages other than English, *Snowball* (StemmersNet 2012).

The algorithm for constructing a dictionary for edit comparison (***):

- Open the web-page with the description of an edit;
- By parsing the HTML of the page, collect all deletions {d} ;
- Remove the stop-words from {d};
- Do the lemmatization of {d} ;
- Make a list of pairs (further refered as *dictionary*) - a stem and the number of its occurrences in {d};

Before applying (***) to edits, the analysis generates a dictionary of the astroturfing edits to be used for comparison. To do that, it collects all deletions from the identified astroturfing edits and applies the same steps as in (***): removes the stopwords, performs lemmatization, generates a dictionary - D (see Tables 4.9 and 4.10 for this dictionary).

Then, for a similarity measure $M(.)$, the score for an edit X: $score(X) = M(D, X|\{d\})$. In other words, the score of an edit equals the similarity of the deleted part of X and D. In the following subsection, the paper will introduce the similarity measure.

4.5.5 Measure: *Conditional* Cosine Similarity

In the previous subsection, the paper presented ways that suspect trolls engage Wikipedia contributions. The proposed unit of analysis is a dictionary of used terms after stemming and removal of the stop words[57]. Tables 4.9 and 4.10 showcase the dictionary of terms for detected instances of atroturfing. In terms of data interpretation, the nature of the data calls

[57]The stop words are the most general words in the language that do not carry any substantive information. For a detailed list, see supplementary materials.

for a measuring methodology that is not sensitive to the order of the words but is rather sensitive to the term frequency.

An intuitive way to capture the distance between two dictionaries is to use *the cosine similarity* method of analysis. For simplicity's sake, let us assume that the terms are the same in the dictionaries and equal to n. Then, each dictionary becomes a vector in the n-space where each term represents a dimension. Further, the cosine similarity is the cosine of the angle θ between two vectors or dictionaries (X and Y) and can be calculated with the following formula:

$$cos(\theta) = \frac{X \cdot Y}{||X||_2||Y||_2} = \frac{\sum_{i=1}^{N} X_i Y_i}{\sqrt{\sum_{i=1}^{N} X_i^2}\sqrt{\sum_{i=1}^{N} Y_i^2}} \tag{4.1}$$

By definition, both dictionaries are not empty, hence the value is defined. Since $cos(0) = 1$, the smaller the angle, the closer is its measure to the unity. If the dictionaries have nothing in common, or in spatial terms are *orthogonal*, then $cos(\frac{\pi}{2}) = 0$. Since all frequencies must be non-negative, the cosine similarity ranges from 0 to 1. Another advantage of this measure is that it reflects only the proportions between the magnitudes of the terms in a dictionary but not the absolute magnitude, meaning that the angle between X and Y is the same as αX and βY (α and β are positive), and so is the cosine similarity.

The only problem with using this measure is that it assumes the same set of terms in both dictionaries. For example, if a term is missing in one dictionary, that increases the distance between the dictionaries. In some cases, this might not be a problem. Meanwhile, in this analysis the edits are compared with the dictionary of the detected astroturfing interventions in order to ascertain common values. This dictionary is likely to have more terminology than most contributions to compare. Substantively, it is enough if the edits are close in meaning when it comes to their common non-zero terms. For example, for the comprehensive

dictionary, *A={russia=1, moscow=1, kremlin=1}* and for the edit, *x={russia=1, moscow=1}* the desired distance is 1, since x fully corresponds to A. Meanwhile, according to (4.1), it is $\sqrt{\frac{2}{3}} < 1$.

However, for the purposes of this analysis, if a term is present in the contribution edit but is missing in the dictionary of the astroturfing contributions, which is assumed comprehensive, that should increase the association distance between them.

If Y is the comprehensive dictionary and X is the observation comparison to Y, then in spatial terms, this analysis desires to capture not the distance between X and Y but rather X and *the projection of Y to the space of X*. The paper introduces *the cosine similarity conditional on X* or *the conditional cosine similarity*, which fully satisfies the purposes of this analysis. If D is the set of the terms in Y and E is the set of terms in X

$$cos(\theta|X) = \frac{\sum\limits_{i \in E} X_i Y_i}{\sqrt{\sum\limits_{i \in E} X_i^2} \sqrt{\sum\limits_{i \in E} Y_i^2}} = \frac{\sum\limits_{i \in E} X_i Y_i}{\sqrt{\sum\limits_{i \in E} X_i^2} \sqrt{\sum\limits_{i \in D} I(i \in E) Y_i^2}} \tag{4.2}$$

where:

$$I(i \in E) = \begin{cases} 1, & \text{if } i \in E \\ 0, & \text{if } i \notin E \end{cases}$$

To simplify the findings, what the measurement implies is that only those terms originating from the Y value affect the measures that are also in X, but not the other way around. Likewise, all terms from the X value that are not found in the Y value increase the association distance between the two. Further, the paper uses (4.2) to propose that a Wikipedia edit is similar to the identified instances of atroturfing. As it will be demonstrated, this measure enables the ability to detect the edits of an examined deletion pretty accurately.

4.5.6 Results

The measure of (4.2) *conditional cosine similarity*, whicg was introduced in the previous subsection, enables comparing an edit to the collection of identified contributions of atroturfing. So far, this paper concluded that the whole telling sign of what makes these contributions suspicious is what information gets erased. Hence, the analysis of the edits focuses on the deletion aspect of an edit. The crucial point to keep in mind before performing the extrapolation of the initial findings, e.g looking for similar edits, is that the same words may be used among different word meanings and phrases. Given this dissonance between word intention and usage, the high values of the similarity score is a necessary but insufficient condition to identify greater instances of political atroturfing. Hence, the substantive step of the analysis (when an expert evaluates an edit thoroughly) is required. Another crucial point is that, as already highlighted, even the edits that substantively correspond to the criteria of political atroturfing cannot be confidently labeled as such.

Obviously, individuals making edits similar to the identified instances of astroturfing may be genuine regime supporters and have nothing to do with troll factories. If that is so, then why it is of importance to look for such edits? The core answer to this inquiry within the major argument of the paper is as follows. Even if the edits identified in this section are not direct cases of genuine political astroturfing (which some of them may easily be), they do represent the very same ideology as political trolls, and consequently impact a sizable share of the Russian (and even Russian-speaking) population that is susceptible to these sentiments. Indeed, as Cantir и Kaarbo 2012b illustrated, despite the fact that the self-perceptions of the regime are primarily shared among the political elites, a part of the population readily shares them as well. Since this paper is curious about the ideology and the self-perceptions of a regime manifested as a result of political astroturfing, then if a broader set of contributions carries the same ideological insights, they should also be included in the analysis. This is

especially the case if they lead to the same conclusions as the findings from the previous section.

Out of 149,894 contributions (of the articles of the categories shown in Tables 4.6, 4.7, and 4.8) included in the analysis, 25,586 have positive conditional cosine similarity values, 3,795 have scores greater than 0.9, and 3425 equal to 1 (see Figure 4.17). Then, the substantive manual analysis is performed regarding all edits with a score greater than 0.9. Foremost, across all categories, the share of vandalism edits is approximately 6%. Concurrently, the substantive analysis does not find any significant share of contributions that are similar to the identified instances of astroturfing from the previous section. With an in-depth examination across all categories, they only constitute approximately 2-5% of contributions examined, including vandalism, which usually involves misspelling of words that aims to taint the meaning and interpretation behind certain topics.

On the other hand, the situation changes dramatically if the analysis narrows down to the set of articles somewhat related to the 2014-present war in the Eastern Ukraine, or, as it is usually referred to, *the War in Donbass*[58]. For instance, among the edits for the articles from the category "*the cities of Donetsk region*"(n=52) with the score greater than 0.9 (n=283) only around 44% are completely ideologically neutral. The rest is basically the *editing war* with the shares approximately evenly divided between the proponents of the pro-Russian militant groups of the newly created self-proclaimed states - DPR[59] and LPR[60]- and their opponents who mainly hold a neural point of view, which is anti-DPR and LPR. The contributions from this "editing war" are very specific and because of their multiple contributors, they appear like an actual grassroots initiative and not an initiative of astroturfing. Meanwhile, what is fascinating and relevant within the main argument of this paper is that they share some interesting similarities with type D contributions from

[58]The region in the Eastern Ukraine where are military actions were happening
[59]Donetsk People's Republic - ДНР (Донецкая народная республика)
[60]Lugansk People's Republic - ЛНР (Луганская народная республика)

the previous section. Before talking about these similarities, it makes sense to describe this "editing war" in more detail.

There is a major clear repetitive pattern across most of its edits: several specific groups of words got changed back and forth. The first group is the country attribution. For the cities in this region, the country attribution was changed in both directions from the self-proclaimed "DNR" or even "DNR (Russia)" to "Ukraine". The second group of the typical substitution edits is the ways to characterize the militant groups controlling the region. Their reference got changed multiple times in both directions from the supportive "insurgent"(opolchency or ополченцы) to negative (through corresponding to the internationally accepted point of view) "separatists" (separatisty or сепаратисты) or even strictly negative "terrorists" (terroristy or террористы).

Furthermore, there is an interesting point to add concerning this "replacement editing war". Foremost, there is certain type of information that gets deleted (not substituted): the mention of Russia's direct involvement in the conflict, particularly the labeling of separatist military forces as "pro-Russian". In addition to that, in the case that the civil citizens were subjected to some danger, the reference to "insurgents" is either removed completely or replaced with the reference to the Ukrainian army.

Summing up the edits, it looks like the goal behind them is to distance Russia and its military's involvement from the conflict in the Eastern Ukraine. The only "acceptable" reference to Russia is that the insurgents are "pro-Russian". Interestingly, this corresponds to Putin's numerous interviews that he gave in 2014 and 2015, in which he insisted that there is no Russian troop presence in the Eastern Ukraine (Insider; Jazeera; News, 2015; 2014; 2014) Secondly, another apparent aim is to dissociate the insurgents from the situations that placed the civil local population in danger, and instead blame the Ukrainian military forces for those actions.

4.6 Discourse analysis: Political Astroturfing in the Historical Perspective

By combining quantitative and qualitative techniques, the research identified a set of Wikipedia contributions that satisfy geographical, *meta-data*, and content requirements to be considered as part of the political troll campaign, which were derived from existing work on Russian political trolls. This study interprets the detected instances of astroturfing (Tables 4.4 and 4.5) as the part of the Russian ideological campaign after the 2014 conflict in Crimea. Foremost, the identified instances of astroturfing correspond to the central notion of the regime's conservative ideology towards Russian history. As Aron (2017) states, all recent Russian ideology centers on the pursuit of national glory. For instance, this glorification of the past mirrors the recent move to make the celebration of the end of WWII as a main national holiday. In general, this ideology simultaneously "features a soup of czarist, Soviet, and post-Communist origin" (Shevtsova, 2017) , when in reality, these origins may often conflict with one another. For instance, as it is listed in the deletions in Table 4.4, those who fought against the Red Army to save the Russian Empire in the Civil War of 1917-1921 and joined the Nazis 20 years later to fight against the communists do not fit this idealized origin of Russia's domestic unity and regime support. At the same time, the ideology of Russia's "pursuit of national glory"is inseparable from keeping and further enlarging the country's territory. This leads to the importance of using a NRC rhetoric of a "regional integrator", and as this research argues, this agenda is most prevalent in the identified instances of astroturfing. To highlight the widespread use of the integrator dogma, C. S. Thies и M. Nieman (2017) point to the pervasive usage of this NRC on their list of identified cases of astrotufing from the 2000s.

Unfortunately, expounding ideology is quite different from the act of inciting war and enlarging the county's territory, which may involve infrastructural devastation and many

casualties. Also, the “united” territories may not always be satisfied with their occupied status. Nonetheless, the regime is innately populist and chooses to enact its patriotic vision no matter the level of destruction or backlash. As Shevtsova (2017) states: "*historically, Russian regimes prepared to resort to sweeping regression and the human loss it entails have faced a population that is accustomed to viewing self-sacrifice as a natural duty, even a way of life. Today, however, Russia's people are not ready to make sacrifices for the sake of the system.*"This means that the current ideology must work hard to “sell” the exhausted need for territorial expansion as an act of greater good for the common Russian. The general impression must be that the joining of annexed territories is a welcome act of those territories who desire to be a part of Russia's new destiny and that their compliance would be carried out without military involvement. Further, the vision suggests that the only military that might be fighting to join Russia are local volunteers, who will ensure that there is no blood on their hands and that the Russian public will never hear of any victims. Also, the ideology insinuates that all territories occupied by Russia have always lived happily in peace and desire to be "Russified"under the label of "friendship of the peoples" (Fish, 2017) . And for those who were against the influence (if it rarely ever happened), were traitors and ungrateful after everything beneficial that Russia had done for them.

In regards to the Wikipedia article on the country of Georgia (see Table 4.4), the following edits were found to fit this precise pattern of the glorification ideology: the removal of the information about the anti-Russian appraising in 1819, the pro-Georgian language demonstrations in 1978, the bloodshed during the suppression of the pro-independence protests in 1989, and the Russian occupation of the Georgian territories of Abkhazia and South Ossetia in 2008. The same can be noted about the deletion of the mention of the disastrous Circassian genocide that occurred at the end of the Russo-Circassian War in 1864-67 in the article on Sochi. Since all the territories readily “joined” Russia and were never “conquered the replacement of the word “conquest” with “inclusion” in the article

about Ashgabat, the capital of Turkmenistan, matches the same pattern of strategic word replacement as well. The modification in the article about Russians in Kazakhstan provides a more subtle example. Whatever was the actual reason why a large number of ethnic Russians left Kazakhstan after the fall of the Soviet Union, the initial phrase "the loss of the privileged status within the USSR"(that would refer to the violation of the perfect picture of "the friendship of the peoples" in USSR) was replaced with the phrase "the suppression", which told a very different story of the events that took place. This filtered version of history does complement another case of censorship of the word "bandit", a recently deleted term and a name previously associated with the Cossacks, a group of East Slavic people located in Russia and Ukraine, who, despite frequent conflicts with the Russian Empire, played a crucial role in the Empire's expansion.

Even the most recent Russian military conflicts are used as ploy to degrade the enemy and bolster Russia as a heroic nation. For instance, the way some articles related to the wars in Chechnya (the first of 1994-1996 and the second during 1999-2000, as well as the whole Chechen independence movement) and the hybrid war in the Eastern Ukraine starting in 2014 were also manipulated through "vandalism signals the (strategically senseless) desire of the content editors to ideologically shape the attitude towards "traitors". In an article about Khankarov, one of the guerrilla fighters during the first Chechen War, the neutral word for "died" (*umer*) is replaced with its rude colloquial synonym, (*sdokh*), which is usually used to refer to the death of a person not worthy of respect. In an article about Lipetsk, the adjective "bandit" is added to describe Poroshenko, the current Ukrainian President and oligarch, who owns a candy factory in this Russian city. Similarly, one of the Ukrainian fighters in the Ukrainian conflict, Semenchenko, was given "a war criminal" association in a biographic page and his ethnic group was changed from "Russian" to "unknown". Interestingly, this pattern of content manipulation resembles another "vandal" edit to an article about a WWI soldier from Ukraine, Pavlov Sergey Vasilievich, who later fought against the Bolsheviks in the Civil War

and then joined the Nazis in the war against the USSR. In the edit, "the prominent figure of the Don history" is replaced with "a traitor and criminal". Apparently, the contributors to these edits are concerned with the favorable presence of the Ukrainian influence as well as Ukrainian nationalism. The mention of the famous writers and Ukrainian nationalists, the Kapranowowie brothers, was removed from the article on Dubossary, a Moldovian city where they resided.

Considering the implication of all these edits, it is clear that the political trolls sought not only to promote the self-perceived role of Russia as a *regional integrator* but to also act as a peaceful *regional integrator* that feeds into the desired view of homeland for the average Russian citizen. According to C. S. Thies и M. Nieman (2017) , holding this NRC or its modification as a *regional liberator* has been a typical approach of Russia and the Soviet Union for most of the time-frame they cover in their book.

Among the numerous detected astroturfing edits, two specific "deletion" edits relate to the military conflict in the Eastern Ukraine. For instance, this is most evident in an article about the river Seversky Donets, which was mentioned in an edited note to serve as the border between the territories controlled by the Ukrainian military forces and the insurgent army, and in an article about Kramatorsk, where there was deliberate removal of the information about the attacks on the city by the insurgent army that placed the civil population in danger. Curiously, as described in detail in the previous section, the extrapolation of the detected instances of political astroturfing had uncovered only edits related to the Ukrainian topic. Hence, according to the findings of this paper, only this topic became viral among regular contributors. Most likely, this behavior was not calculated among groups of political trolls but most likely conducted independently. But why? As already noted before, the annexation of Crimea and the consequent rise of Russian nationalism was the main (and probably the only) successful ideological outcome of Putin's regime. The findings in this paper confirm this assertion.

Another self-perceived role that left a trace of influence in the detected astroturfing edits is "defender of the faith". As Fish (2017) notices about Putin, "he ostentatiously blesses Russia's traditional religions, especially, the majority faith of Orthodox Christianity, but also Islam, Judaism, and Buddhism", while being nonreligious. For the current Russian ideology, religion is a part of a conservative and traditionalist paradigm of authority. Religion also serves the purpose of distinguishing Russia from the West, which is not Orthodox Christian. While the idea of religion is mainly centered on Orthodox Christianity, Islam, as the second largest practiced religion in Russia, is handled very strategically as well. For example, according to Russian law, when mentioning the Islamic State publicly, one must also distinguish that it is an " organization forbidden in Russia" in order to separate the organization from the common Islamic faith (Fish, 2017) . The traces of this self-perceived role are significantly less apparent than the "regional integrator", and it was not confirmed that its influence attracted any genuine contributors. The research detected five edits corresponding to it in a page about the French magazine *Charlie Hebdo*. Four out of five of these edits were "vandalistic", in that they used offensive words to characterize the main journalist of the magazine.

4.7 Conclusion

This work has confirmed that political astroturfing attempts are detectable on Wikipedia and their contents clearly confirm the most up-to-date insights of political science research about the nature of the ideology and self-perception of the current Russian political regime. Foremost, the findings are consistent with the historical-political construct of "pursuing national glory" where a peaceful, mighty and expanding Empire serves as a focal source of power and protection for the communities historically and terrestrially linked to its fate. Additionally, as a result of the extrapolation of the astroturfing interventions, the study verifies that the recent rise of nationalism relates directly to the notion of the volunteer

secessionist movement in the Eastern Ukraine. A somewhat surprising finding is that despite the overall current authoritative and *macho* style of governance of the Putin regime, the mainstream version of the regime is likely to be constantly supplied with filtered content free from unpleasant and controversial details. Meanwhile, this last observation is easily explained with another feature of this regime's embodiment- its populist nature- that gives it the persuasive power to convince average citizens to accept a personal economic loss for the belief that unlike the recent past, their sacrifices will enable a reality where greatness will no longer be contingent on violence and human sacrifice.

Tables and Figures

Figure 4.1: Soviet Union/Russia's Role Sets (from Thies and Nieman (2017)

1940s	*anti-imperialist agent*, *defender of the faith*, *regional liberator*, *guardian*, *example*, *U.S. rival*
1950s	*anti-imperialist agent*, *defender of the faith*, *regional liberator*, *guardian*, *example*, *U.S. rival*
1960s	*anti-imperialist agent*, *regional-subsystem collaborator*, *regional protector*, *regional liberator*, *example*, *liberation supporter*, *developer*, *defender of the peace*, *defender of the faith*, *guardian*, *example*, *U.S. rival*, *revolutionary power*
1970s	*anti-imperialist agent*, *regional-subsystem collaborator*, *regional protector*, *regional liberator*, *example*, *liberation supporter*, *developer*, *defender of the peace*, *defender of the faith*, *guardian*, *example*, *U.S. rival*, *revolutionary power*
1980s	*status quo power*, *integrator of Russia with the West*, *member of a democratic community of states*, *promoter of arms control and disarmament*, *international collaborator*, *reformer of the international order*, *reformer of the regional European order*, *defender of universal human values*, *supporter of international organizations*, *new member of the international community*, *collaborator/partner*, *independent and influential player*, *great power*, *excluded from the international system*
1990s	*promoter of a regional geopolitical order*, *integrator of Russia into the Western world*, *promoter of international cooperation*, *protector of the Russian state and independent player*, *instrument for changing Russia*, *defender of democratic principles and human rights*, *supporter of international organizations*, *supporter of arms control and disarmament*, *(normal nonaggressive) great power*, *stabilizer of the Eurasian geopolitical environment*, *democratic state*, *U.S. ally*, *new state*
2000s	*Eurasian power*, *excluded from restrictive Western institutions*, *collaborator/partner*, *leading power*, *(normal) great power*, *supporter of international organizations*, *regional peacekeeper/defender of the peace*, *reformer*, *defender of democratic principles and human rights*, *supporter of international organizations*, *supporter of arms control and disarmament*, *defender of international law*, *cultural and civilizational bridge*, *intermediary*, *regional integrator*, *balancer and guardian of global stability*, *U.S. rival*, *major power*

Figure 4.2: Russian Wikipedia: Pageviews for January 2013 - May 2015.

Figure 4.3: The articles from categories Ukraine, Kiev, and Crimea on Russian Wikipedia: Pageviews for September 2013 - July 2014

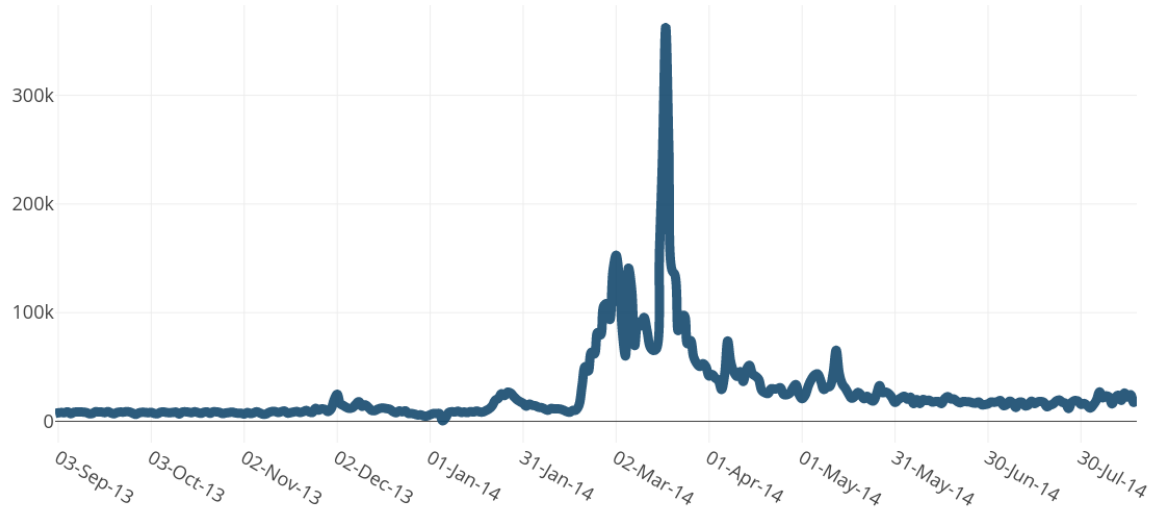


Figure 4.4: Actual ad from Russian social networks. August 2013 (Garmajalova, 2013)

Translation: "Internet operators are required! Work in a chic office in OLGINO!!!! (m. Staraya derevnya), payment of 25 960 per month. Task: posting comments on thematic Internet sites, writing thematic posts, blogs, posts on social networks. Work reports in the form of screen-shots. Work schedule selected individually ... ... Weekly payments, 1180 [rubbles] per shift (from 8.00 to 16.00, from 10.30 to 18.30, from 14.00 to 22.00). WEEKLY PAYMENTS AND FREE MEALS!!! Full-time employment or with a contract (by choice). Training is possible!".

Figure 4.5: Typical assignment in a "troll factory" (from Volchek and Sindelar (2015))

A TYPICAL ASSIGNMENT

Topic: NATO troops are embedded with Ukrainian armed forces

Keywords: *ukraine news, russia and ukraine, ukraine policy, ukraine, NATO, PMC (private military company)*

Task: Raise this topic on 35 municipal forums

Figure 4.6: A joke on the Russian Internet dedicated to political trolls: A banknote of 11 rubles.

Explanation of the details: Eleven (rubles) was the amount of money the Russian political troll received for one comment. The building on the banknote is the "The Internet Information Center" on Savushkina in Saint Petersburg. The settlement on the map says "Olgino", the original location of “the troll factory”.

Figure 4.7: The troll factories on the map.

Figure 4.8: The example of the data from the DB-IP database.

```
"62.152.89.0","62.152.89.15","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
"62.152.89.16","62.152.89.19","RU","St.-Petersburg","Saint Petersburg","59.9596","30.3234","3","Europe/Moscow"
"62.152.89.20","62.152.89.23","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
"62.152.89.24","62.152.89.27","RU","Samara","Usinskoye","55.0654","56.3653","5","Asia/Yekaterinburg"
"62.152.89.28","62.152.89.71","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
"62.152.89.72","62.152.89.79","RU","Samara","Usinskoye","55.0654","56.3653","5","Asia/Yekaterinburg"
"62.152.89.80","62.152.89.95","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
"62.152.89.96","62.152.89.127","RU","Samara","Usinskoye","55.0654","56.3653","5","Asia/Yekaterinburg"
"62.152.89.128","62.152.92.255","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
"62.152.93.0","62.152.93.255","RU","Russia","г. Санкт-Петербург","59.9684","30.3432","3","Europe/Moscow"
"62.152.94.0","62.152.94.255","RU","St.-Petersburg","Saint Petersburg","59.9596","30.3234","3","Europe/Moscow"
"62.152.95.0","62.152.95.255","RU","St.-Petersburg","Saint Petersburg","59.9343","30.3351","3","Europe/Moscow"
```

Figure 4.9: The DB-IP spatial analysis: The area around Varvarskaya.

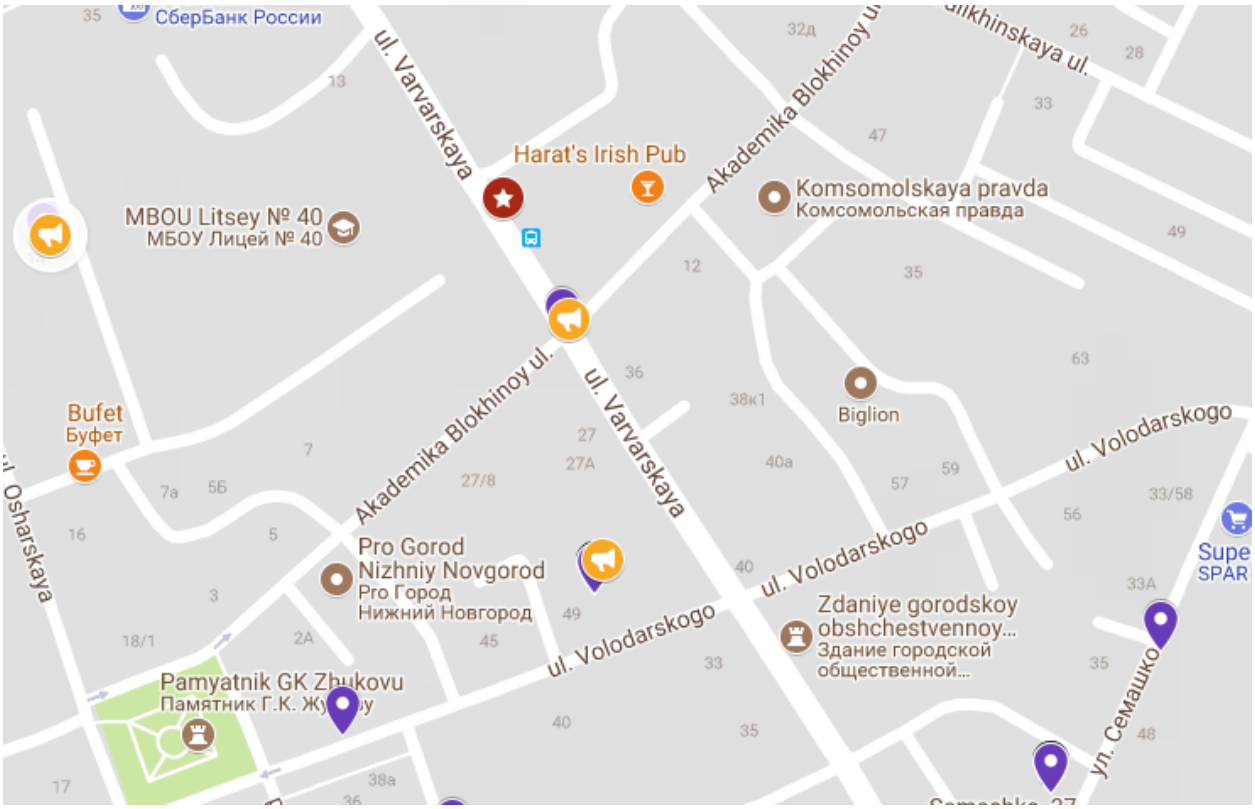


Figure 4.10: The DB-IP spatial analysis: The area around Savushkina.

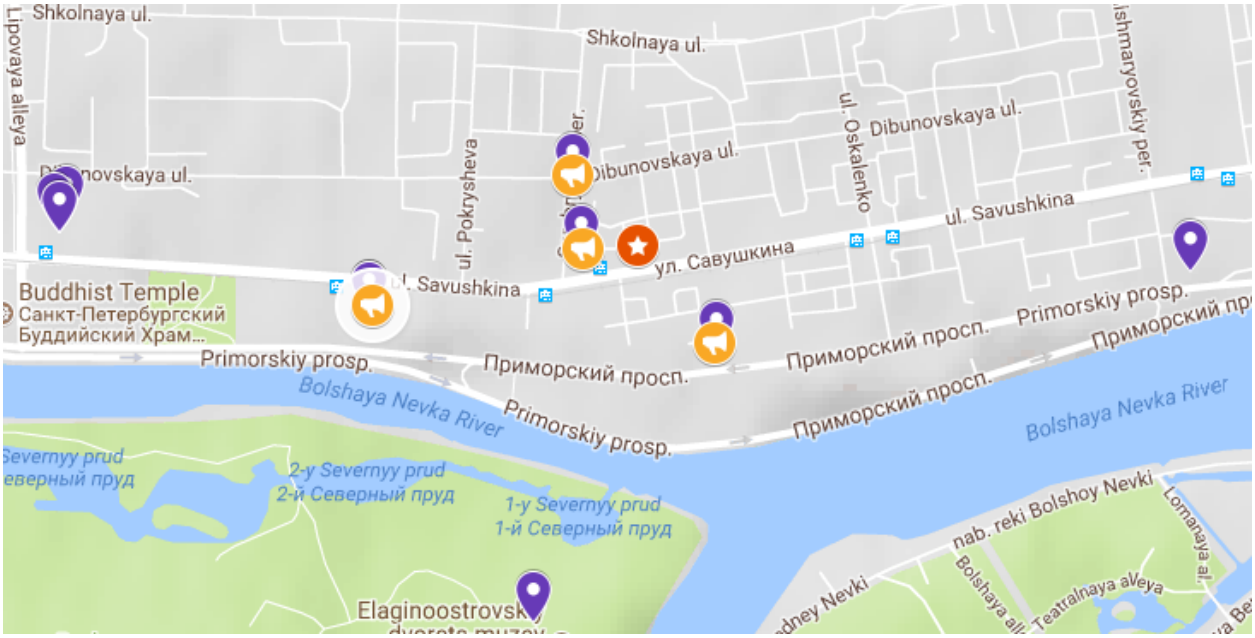


Figure 4.11: The DB-IP spatial analysis: The area around Yablonovsky.

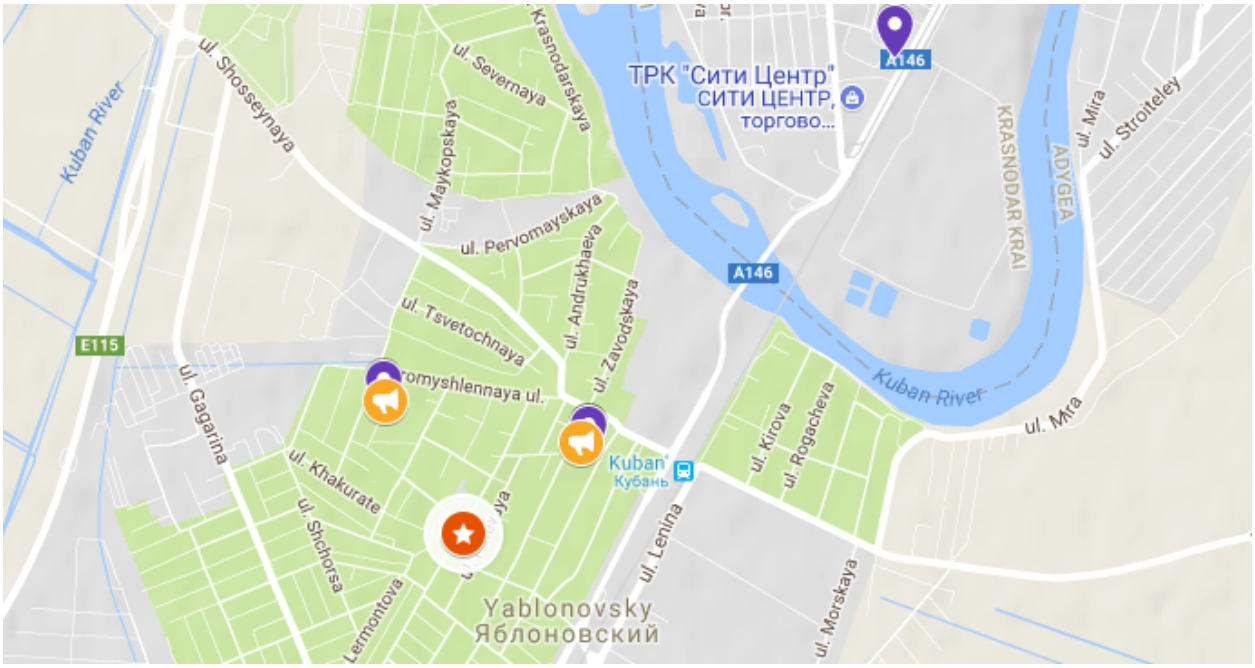


Figure 4.12: Wikipedia revisions' dump: an extract.

```
<page>
  <title>Литва</title>
  <ns>0</ns>
  <id>7</id>
  <revision>
    <id>25</id>
    <timestamp>2002-11-13T18:48:05Z</timestamp>
    <contributor>
      <username>Victor~ruwiki</username>
      <id>5</id>
    </contributor>
    <comment>test</comment>
    <model>wikitext</model>
    <format>text/x-wiki</format>
    <text id="25" bytes="129" />
    <sha1>ghq6oyr2gzls2tjqg1gkiovnsm1twmd</sha1>
  </revision>
  <revision>
    <id>82</id>
    <parentid>25</parentid>
    <timestamp>2002-11-30T16:11:57Z</timestamp>
    <contributor>
      <ip>80.73.68.239</ip>
    </contributor>
    <model>wikitext</model>
    <format>text/x-wiki</format>
    <text id="82" bytes="167" />
    <sha1>o1r20z20n03wi72sycgwkikxqltwydd</sha1>
  </revision>
  <revision>
    <id>239</id>
    <parentid>82</parentid>
    <timestamp>2003-01-07T19:18:00Z</timestamp>
    <contributor>
      <ip>212.59.62.90</ip>
    </contributor>
    <model>wikitext</model>
    <format>text/x-wiki</format>
    <text id="239" bytes="199" />
    <sha1>n3e6gl9twowetvdrrd3ulqxqhd0qcce</sha1>
  </revision>
```

Figure 4.13: Revision's online page.

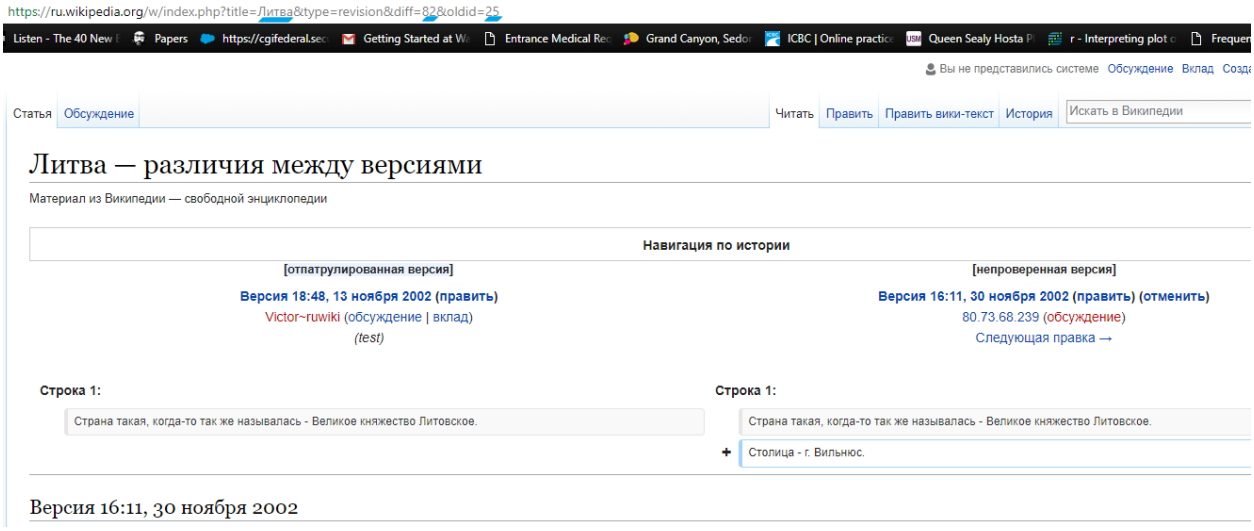


Table 4.1: Geographical coordinates of the troll factories

	name	latitude	longitude	IP ranges	detected IPs	contributions
1	Olgino	59.99671	30.13154	0	0	0
2	Savushkina, 55	59.98433	30.27192	4	0	0
3	Zelenyy Gorod	56.16729	44.06608	0	0	0
4	Varvarskaya, 32	56.32139	44.01241	3	0	0
5	Yablonovsky	44.98674	38.94344	2	19	88
6	Perekatny	45.03046	38.9449	0	0	0

Table 4.2: Suspect IP contributors: Basic Stats

IP	ip range	n	articles of the edits
c1	yablonovsky2	26	14 June; 14 June; Georgia; Georgia; Ashgabat; Lipetsk; Sochi; Seversky Donets; Donetsk (Russia); Cell theory; Dubossary; Starominskaya; Mr. Credo; Khankarov, Khamzat Zharapovich; Khankarov, Khamzat Zharapovich; Multiverse (DC Comics)(l); Pavlov, Sergey Vasilievich; Charlie Hebdo (l); Charlie Hebdo (l); Charlie Hebdo (l); Charlie Hebdo (l); Charlie Hebdo (l); Flash (television series, 2014) (l); Flash (television series, 2014) (l); Flash (television series, 2014) (l); Semenchenko, Semyon Igorevich
c2	yablonovsky2	24	January 9; Ozon.ru (l); Social Psychology; Yablonovsky (Adygea); Yablonovsky (Adygea); M Video; Connected (company); Dosage aerosol dispenser; Vettel, Sebastian; Social network; Robinson, Ross; Chia (plant); Stronghold Kingdoms (l); Prime World (l); War Thunder (l); Blood and Soul (l); Black Star Inc . (l); Black Star Inc . (l); Star Conflict (l); Star Conflict (l); Star Conflict (l); Panzar (l); L'One (l); L'One (l)
c3	yablonovsky2	10	Cossacks; Kramatorsk; Economic laws; Yury Dolgoruky; Yury Dolgoruky; Russians in Kazakhstan; Culture of Ancient Russia; Culture of Ancient Russia; Anti-collector agency; Legal consultation
c4	yablonovsky2	6	Planet of the Apes (novel); Planet of the Apes (novel); Planet of the Apes (novel); Planet of the Apes (novel); Planet of the Apes (novel); Planet of the Apes (novel)
c5	yablonovsky2	4	Manticore (monster); Arianism; Glade; Pozharsky, Dmitry Mikhailovich
c6	yablonovsky2	2	Malcolm X; Krasnodar musical theater
c7	yablonovsky2	2	Krasnodar; Krasnodar
c8	yablonovsky2	2	International Society for Krishna Consciousness; The International Society for Krishna Consciousness
c8	yablonovsky2	2	Margaritovo (Rostov Region); Margaritovo (Rostov region)
c9	yablonovsky1	1	Minecraft (l)
c10	yablonovsky1	1	Cherkessia
c11	yablonovsky2	1	Steel (stadium, Alchevsk)
c12	yablonovsky2	1	Univer (TV series)
c13	yablonovsky2	1	Clickjacking
c14	yablonovsky2	1	Assolkolai
c15	yablonovsky2	1	Yablonovsky (Adygea)
c16	yablonovsky2	1	Adverb
c17	yablonovsky2	1	Sofia (name)
c18	yablonovsky2	1	Silent Hill 4: The Room (l)

Notes: 1) The IPs are masked in the main text because of the privacy reasons. They are available by request or can be reconstructed through the replication code. 2) The articles names are originally in Russian; the translated version is provided unless it is explicitly indicated - *latin* or l means that the article name was originally in the latin script.

Table 4.3: Suspect IP contributors: Evaluation of the contributions

IP	total	detected	*van|d*	*ca|d*	*malicious|nd*	good edits	good edits (%)
c1	26	18	8	10	2	5	15
c2	24	0	0	0	0	24	100
c3	10	7	4	3	3	0	0
c4	6	0	0	0	0	6	100
c5	4	0	0	0	0	4	100
c6	2	0	0	0	0	2	100
c7	2	0	0	0	0	2	100
c8	2	0	0	0	0	2	100
c9	1	0	0	0	0	1	100
c10	1	0	0	0	0	1	100
c11	1	0	0	0	0	1	100
c12	1	0	0	0	0	1	100
c13	1	0	0	0	0	1	100
c14	1	0	0	0	0	1	100
c15	1	0	0	0	0	1	100
c16	1	0	0	0	0	1	100
c17	1	0	0	0	0	1	100
c18	1	0	0	0	0	1	100

Notes: detected refers to the number of contributions substantively satisfying the criteria of astroturfing; $van|d = vandalism|detected$ - the vandal contributions that satisfy the criteria of astroturfing; $ca|d$ = clever astroturfing | detected; malicious contributions that do not satisfy the conditions of the astroturfing; good edits - the contributions that improve the quality of an article

Table 4.4: Detected instances of astroturfing on the Russian Wikipedia (c1)

article	type	brief description
Georgia	D	Removal of information related to Russian occupation of the Georgian territories of Abkhazia and South Ossetia in 2008
	D	Removal of information related to the anti-Russian uprising (1819), the demonstration in Georgia (1978), the bloodshed during the suppression of the demonstration (1989), and the occupation of the Georgian territories of Abkhazia and South Ossetia.
Ashgabat	D	In the description of Turkmenistan's association with the Russian Empire, the word "conquest" was replaced with "inclusion".
Lipetsk	V	The adjective "bandit" was added to the description of the candy factory Roshen, which belongs to Poroshenko, the recent president of Ukraine.
Sochi	D	Removal of information of the Circassian genocide at the end of the Russo-Circassian War in Russia in 1864 – 1867.
Seversky Donets	D	Removal of the information that this river served as a border between the Ukrainian military and the insurgent army in the Eastern Ukraine in 2014.
Donetsk (Russia)	D	Removal of information from the section on famous citizens about an officer who fought in the White Army during the Civil War and in the Nazi Army in WW2.
Cell theory	D	Removal of some information about the scientific achievements of foreign scientists that could detract from the successes of their Russian counterparts.
Dubossary	D	Removal from the famous citizens section the information about the Kapranowowie brothers, who are famous writers and Ukrainian nationalists.
Khankarov	V	In the description of his death, the neutral word "died" was replaced with a more colloquial term, "sdokh", which delegates a negative attitude to the deceased person.
Pavlov	V	In the description, the phrase "prominent figure of the newest Don history" is replaced with "traitor and criminal".
Charlie Hebdo	V	In the description, the words "well-known, sketch-cartoonists" were replaced with "rascals".
	V	The words "French satirical weekly" were replaced with "French porn weekly". Also, in the description where the magazine mentions "ridicules politicians, ultra-rightists, Islam and Christianity" it is also added "for that they paid when the Islamists killed 12 scoundrels on the editorial office of this newspaper" (in colloquial language).
	V	The addition of the colloquial expression for "get afraid", pointing to the threatening disrespect of the people indicated.
	D	The removal of the word "famous" from their description.
	V	The addition of the colloquial expression" cowardly creatures" to describe them.
Semenchenko	V	The addition to the description the terms "a war criminal". The replacement of his ethnic group from "Russian" to "unknown".

Table 4.5: Detected instances of the atroturfing on Russian Wikipedia (c3)

article	type	brief description
Cossacks	D	The removal of one of the meanings of the word "bandit".
Kramatorsk	D	The removal of part of the description of the attacks of the insurgent army in the city when it was under control of the Ukrainian military.
Dolgoruky	V	The misinformation act of replacing "Volyn" to "Volynsk".
	V	The replacement of the actual name of "The Tale of Igor's Campaign" with "The Tale of Tortoise's Campaign".
Russians in Kazakhstan	D	The removal of certain details about the reasons why Russians left Kazakhstan in the beginning of the 1990s. The reason given as "the loss of privileged status within the USSR" is replaced with "suppression".
Culture of Ancient Russia	V	Rewriting a sentence while misspelling several words.
	V	Change of a date to a wrong one

Figure 4.14: The article on *Georgia*: A typical example of subtle manipulation by deletion.

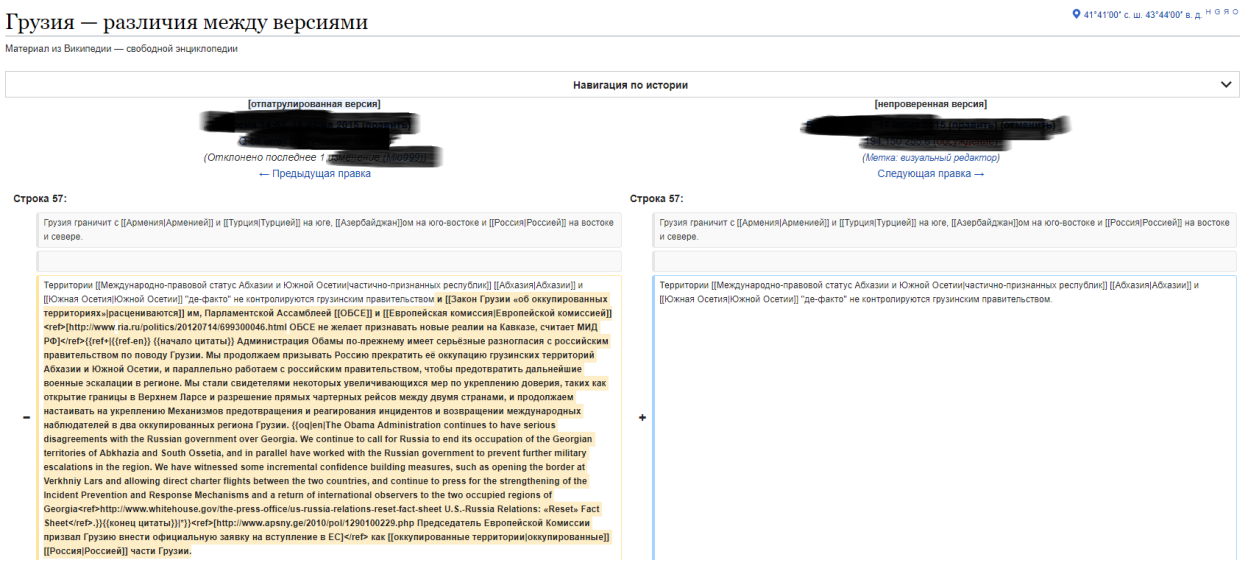
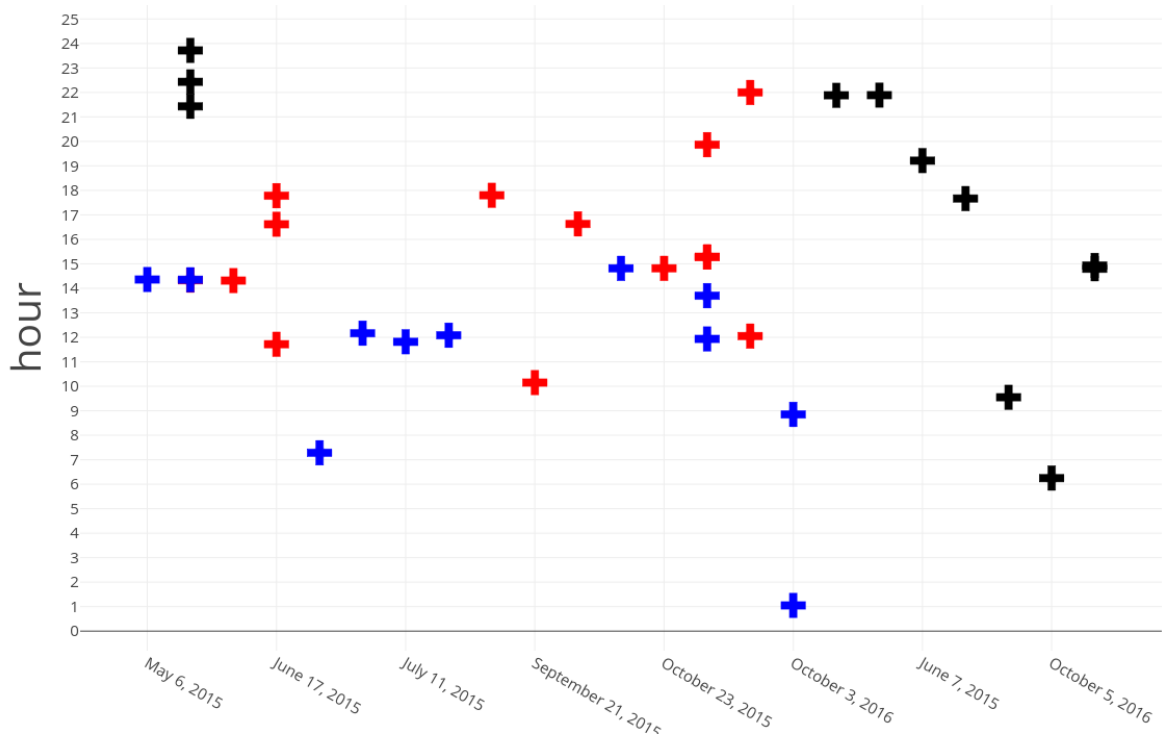


Figure 4.15: The article on *Charlie Hebdo*: A typical example of abusive (*hate speech*) manipulation .

Figure 4.16: Time of the day and the date: c1 and c3's contributions to Wikipedia



Note: The black crosses mark non-astroturfing edits; the blue are the astroturfing edits of type D; the red crosses - of type V.

Table 4.6: Detected instances of the atroturfing on Russian Wikipedia: Categories

title (source)	title (translation)	n
участники первой мировой войны*	participants of the First World War*	5405
персоналии: белое движение*	personalities: the white movement (during the Civil War in 1917-1922 in Russia)*	803
рюриковичи по алфавиту	rurikovichi in alphabetical order	506
переименованные населённые пункты россии	renamed settlements of Russia	384
генералы вермахта	Wehrmacht generals	305
политики украины	policy of Ukraine	282
национальные меньшинства	national minorities	248
выпускники николаевского кавалерийского училища	graduates of the Nikolayev Cavalry School	233
пограничные реки	border rivers	233
общественные деятели украины	public figures of ukraine	230
народные депутаты верховной рады украины viii созыва	people's deputies of the Verkhovna Rada of Ukraine viii convocation	218
народы россии	peoples of russia	179
реки ростовской области	the river of the Rostov region	176
участники евромайдана	euromaidan participants	172
унитарные государства	unitary states	162
социалистическая пресса	socialist press	125
цитология	cytology	116
основатели городов	founders of cities	114
русские коллаборационисты во второй мировой войне	Russian collaborators in the Second World War	112
вооружённые силы российской империи	armed forces of the Russian Empire	106
кавалерия	cavalry	100
палеолит европы	Paleolithic of Europe	93
административные центры субъектов российской федерации	administrative centers of the subjects of the Russian Federation	84
притоки дона	tributaries	83
отравленные	poisoned	82
города, основанные в xix веке	cities founded in the xix century	80
объекты книги рекордов гиннесса	objects of the Guinness Book of Records	75
города, основанные в xviii веке	cities founded in the xvi century	74
псевдонимы	aliases	70

Note: * are excluded from the further analysis due to their generality.

Table 4.7: Detected instances of the atroturfing on Russian Wikipedia: Categories

title (source)	title (translation)	n
чеченские полевые командиры	Chechen field commanders	68
казачество	Cossacks	63
участники грузино-абхазских войн	participants of the Georgian-Abkhaz wars	63
реки белгородской области	the rivers of the Belgorod region	62
курорты ссср	resorts ussr	59
города молдавии	cities of moldavia	57
реки луганской области	the Lugansk region	54
сословия	estates	53
города донецкой области	the cities of Donetsk region	52
поп-музыканты	pop musicians	48
реки харьковской области	the river of the Kharkov region	46
великие князья киевские	the Grand Dukes of Kiev	45
сатирические журналы	satirical magazines	44
социальная структура русского государства	social structure of the Russian state	43
чеченские поэты	Chechen poets	42
российские военные укрепления	Russian military fortifications	40
русские в казахстане	Russians in Kazakhstan	38
приграничные с украиной районы российской федерации	bordering with Ukraine regions of the Russian Federation	36
районные центры краснодарского края	district centers of the Krasnodar Territory	34
скифы	Scythians	34
журналы франции	magazines of france	31
курганы	mounds	31
народы казахстана	peoples of Kazakhstan	31
военачальники северо-восточной руси	commanders of north-eastern Russia	30
князья ростовские	princes of Rostov	30
реки донецкой области	rivers of the Donetsk region	29
россия при романовых (1613—1917)	russia with the novels (1613-1917)	29
военные украины	military Ukraine	28
выпускники донского кадетского корпуса	graduates of the Don Cadet Corps	28
история республики казахстан (с 1991 года)	history of the Republic of Kazakhstan (since 1991)	27
экономические законы	economic laws	25
краматорск	kramatorsk	24
мономаховичи	monomachovichi	24
города краснодарского края	city of Krasnodar region	23

Table 4.8: Detected instances of the atroturfing on Russian Wikipedia: Categories

title (source)	title (translation)	n
населённые пункты староминского района	populated areas of the Starominsky district	21
города на днестре	cities on the Dniester	19
вооружённые силы русского государства	armed forces of the Russian state	18
грузия	Georgia	18
диаспорные группы средней азии	diaspora groups of Central Asia	17
сатирики франции	French satirists	17
персоналии:суздаль	people: suzdal	16
районные центры липецкой области	district centers of the Lipetsk region	16
дубоссары	Dubossary	15
армия русского государства	army of the Russian state	14
города воинской славы	city of military glory	14
населённые пункты краматорского городского совета	populated places of the Kramatorsk city council	14
курорты федерального значения	resorts of federal importance	13
липецк	Lipetsk	13
города на северском донце	city in the northern part of the Donets	12
командиры украинских добровольческих батальонов	commanders of Ukrainian volunteer battalions	12
государства постсоветского пространства	states of the post-Soviet region	10
биологические теории	biological theories	8
культура древней руси	the culture of ancient Russia	7
сочи	Sochi	7
страны у чёрного моря	countries near the Black Sea	7
ашхабад	Ashgabat	6
курорты приднестровской молдавской республики	resorts of the Transnistrian Moldovan Republic	6
почётные граждане парижа	honorary citizens of Paris	6
члены партии «самопомощь»	members of the "self-help"party	6
герои абхазии	heroes of Abkhazia	5
города приднестровской молдавской республики	the city of the Transnistrian Moldovan Republic	5
донецк (ростовская область)	Donetsk (Rostov region)	5
переименованные населённые пункты туркмении	renamed settlements of Turkmenia	5
северский донец	Seversky Donets	5
стрельба в редакции газеты charlie hebdo	Charlie Hebdo	5
авторы-исполнители чечни	authors/executors of Chechnya	3
похороненные в церкви спаса на берестове	buried in the church of Spas on Berestov	3
похороненные в новогрудке	buried in Novogrudka	2

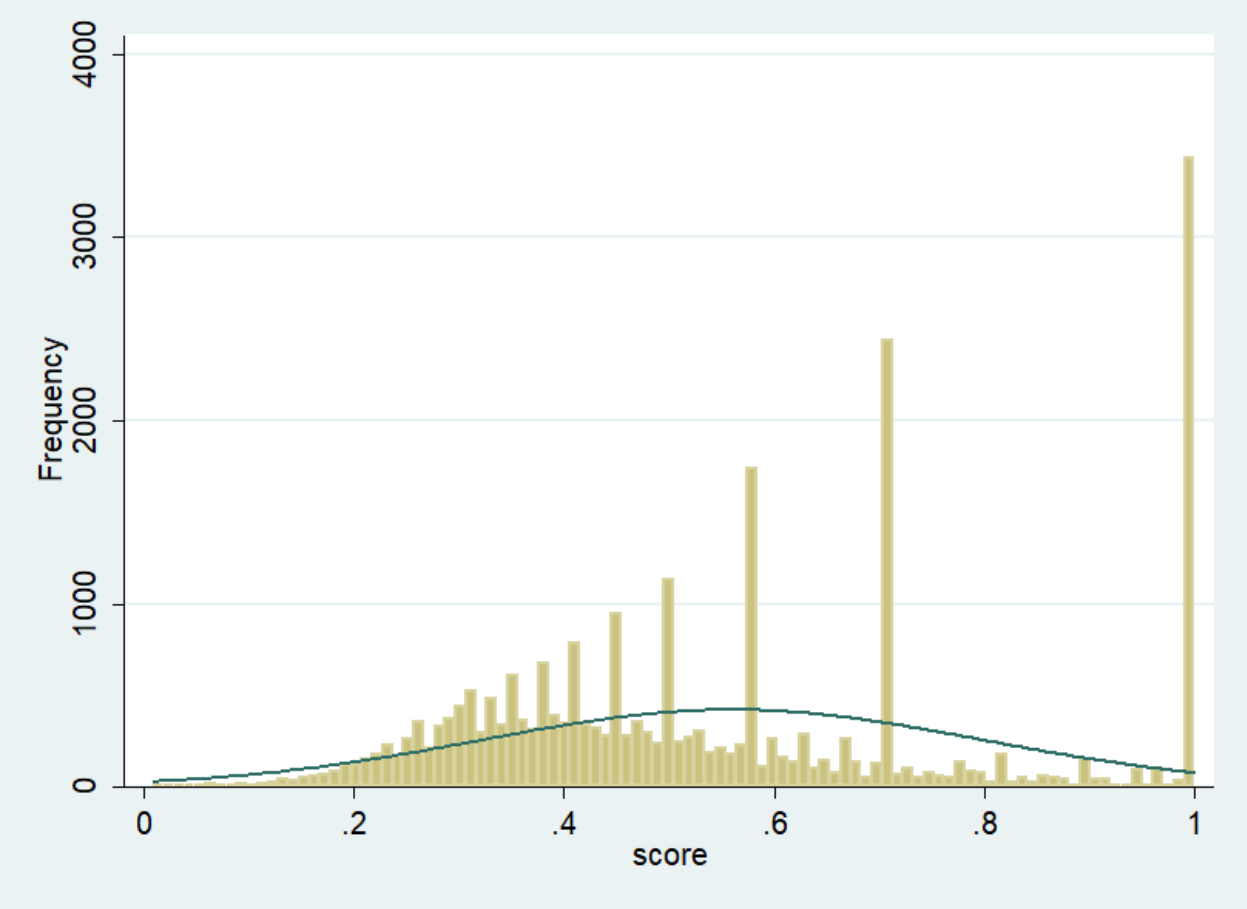
Table 4.9: Initial astroturfing contributions: Dictionary arranged by the frequency of the used terms (I)

stem	n	stem	n	stem	n	stem	n
теор	9	пророк	4	числ	2	чтко	2
росс	8	июл	3	безопасн	2	полн	2
шван	8	февр	3	казахст	2	бож	2
клетк	8	сист	3	игоревич	2	оконч	2
украин	7	дет	3	укр	2	враждебн	2
семенченк	7	русл	3	им	2	горц	2
груз	7	некотор	3	политическ	2	последн	2
клеточн	7	сем	3	майд	2	насел	2
структур	7	род	3	верховн	2	матер	2
животн	7	командир	3	рад	2	турц	2
краматорск	6	парт	3	соз	2	засел	2
территор	6	прож	3	самопомощ	2	западн	2
раст	6	европейск	3	стб	2	казак	2
обстрел	5	комисс	3	сентябр	2	переселенц	2
республик	5	кавказ	3	родственник	2	велик	2
батальон	5	российск	3	брат	2	княз	2
донбасс	5	правительств	3	переж	2	мих	2
работ	5	южн	3	откр	2	николаевич	2
оккупир	5	границ	3	публик	2	склон	2
част	5	стран	3	ус	2	отряд	2
войн	5	мир	3	выполн	2	труп	2
элементарн	5	липецк	3	успешн	2	женщин	2
обс	4	сокол	3	факт	2	собак	2
обществ	4	комп	3	нов	2	болезн	2
экономическ	4	вод	3	цитат	2	едв	2
русск	4	жив	3	продолж	2	берег	2
июн	4	кубанск	3	грузинск	2	мал	2
украинск	4	положк	3	абхаз	2	выбрас	2
во	4	сравн	3	осет	2	луганск	2
известн	4	доказ	3	регион	2	народн	2
крым	4	гор	2	укрепл	2	классическ	2
половин	4	карачун	2	явл	2	оказ	2
александр	4	сил	2	гагрск	2	генл	2
черкес	4	донецк	2	бзыбск	2	принцип	2
кавказск	4	закон	2	завоев	2	стро	2
исследов	4	рамк	2	свободн	2	рост	2
клеток	4	определ	2	завод	2	ядр	2
микроскопическ	4	сост	2	лебедянск	2	основн	2
соответств	4	главн	2	прогресс	2	бесструктурн	2
шлейд	4	представител	2	центр	2	веществ	2
отказ	4	сообщ	2	ядр	2	сочин	2
карикатур	4	орг	2	рескрипт	2	книг	2

Table 4.10: Initial astroturfing contributions: Dictionary arranged by the frequency of used terms (II)

stem	n	stem	n	stem	n
хорд	2	объективн	1	президент	1
хрящ	2	обусловл	1	вдзнач	1
показ	2	чм	1	державн	1
век	2	функц	1	нагород	1
впоследствиикогд	2	волын	1	горович	1
банд	1	игор	1	прошл	1
власт	1	кор	1	константин	1
днр	1	медл	1	гришин	1
вооружн	1	темп	1	переход	1
всу	1	потер	1	вопрос	1
подверг	1	привилегир	1	наст	1
внешн	1	статус	1	севастопол	1
годовщин	1	ссср	1	деятел	1
освобожд	1	террористическ	1	получ	1
видеорепортаж	1	группировк	1	вес	1
реактивн	1	террорист	1	качеств	1
залп	1	глав	1	добровольческ	1
огн	1	общин	1	собр	1
рсзо	1	юр	1	автобус	1
оперативн	1	бунак	1	добровольц	1
отчт	1	утвержд	1	зерк	1
специальн	1	виктор	1	недел	1
мониторинг	1	пугач	1	депутат	1
мисс	1	соратник	1	номер	1
смм	1	явн	1	списк	1
отчет	1	психическ	1	самопомич	1
сми	1	ненормальн	1	передач	1
боцюрк	1	организац	1	окна	1
майкл	1	устькаменогорск	1	телека	1
гуд	1	неоднократн	1	рассказ	1
ирин	1	правоохранительн	1	близк	1
госадминстрац	1	комитет	1	мат	1
погибл	1	национальн	1	отец	1
пострад	1	кнб	1	поддерж	1
жертв	1	серг	1	присоедин	1
трагед	1	козл	1	захотел	1
доног	1	восточн	1	уезж	1
наряд	1	письм	1	вывезт	1
разорв	1	многочисл	1	бойц	1
ввод	1	семн	1	слов	1
кажд	1	указ	1	жизн	1
				оттуд	1

Figure 4.17: Histogram: The non-zero conditional cosine similarities for the articles from the selected categories



References

Abbott, Peter (2011). *Modern African Wars (2): Angola and Mozambique 1961-74*. Bloomsbury Publishing.

Abir, Mordechai and Aryeh Yodfat (2013). *In the Direction of the Gulf: The Soviet Union and the Persian Gulf*. Taylor & Francis.

Allison, Roy (2008). "Russia Resurgent? Moscow's Campaign to 'Coerce Georgia to Peace'." In: *International Affairs* 84.6, pp. 1145–1171.

Ambrosio, Thomas (2006). "The Political Success of Russia-Belarus Relations: Insulating Minsk from a Color Revolution." In: *Demokratizatsiya* 14.3, p. 407.

Arceneaux, Kevin, Martin Johnson, and Chad Murphy (2012). "Polarized Political Communication, Oppositional Media Hostility, and Selective Exposure." In: *Journal of Politics* 74.1, pp. 174–186.

Aron, Leon (2017). "Putinism After Crimea." In: *Journal of Democracy* 28.4, pp. 76–79.

Aydin, Aysegul (2012). *Foreign Powers and Intervention in Armed Conflicts*. Stanford University Press.

Aydin, Aysegul and Patrick M Regan (2012). "Networks of Third-party Interveners and Civil War Duration." In: *European Journal of International Relations* 18.3, pp. 573–597.

Babiracki, Patryk (2015). *Soviet Soft Power in Poland: Culture and the Making of Stalin's New Empire, 1943-1957*. University of North Carolina Press.

Bapat, Navin A (2006). "State Bargaining with Transnational Terrorist Groups." In: *International Studies Quarterly* 50.1, pp. 213–230.

BBC, Russian (2015). *Troll factories: self-disclosure thanks to Google*. Ed. by BBC Russian.

Beevor, Antony (2006). *The Battle for Spain: The Spanish Civil War, 1936-1939*. Penguin Books.

Beissinger, Mark R (2013). "The Semblance of Democratic Revolution: Coalitions in Ukraine's Orange Revolution." In: *American Political Science Review* 107.03, pp. 574–592.

Bideleux, Robert and Ian Jeffries (2007). *The Balkans: A Post-communist History*. Routledge.

Bidwell, Robin (1994). *Obituary: Abdullah Sallal*. URL: `http://www.independent.co.uk/news/people/obituary-abdullah-sallal-1429912.html`, Accessed 6/5/2017. March 18.

Blake, Aaron (2016). *The first Trump-Clinton presidential debate transcript, annotated*.

Bohlen, Celestine (1994). "Russia and Georgia Sign Military Cooperation Treaty." In: *The New York Times*. URL: http://www.nytimes.com/1994/02/04/world/russia-and-georgia-sign-military-cooperation-treaty.html.

Bradberry, Benton L (2012). *The Myth of German Villainy*. Author House.

Breuning, Marijke (2011). "Role Theory Research in International Relations." In: *Role Theory in International Relations: Approaches and Analysis*. Ed. by Sebastian Harnisch, Cornelia Frank, and Hanns W. Maull. London: Routledge.

Brummer, Klaus and Cameron G. Thies (2015). "The Contested Selection of National Role Conceptions." In: *Foreign Policy Analysis* 11.3, pp. 273–293.

Brzezinski, Zbigniew (2012). *Strategic vision: America and the crisis of global power*. Basic Books.

Bueno de Mesquita, Bruce and Alastair Smith (2007). "Foreign Aid and Policy Concessions." In: *Journal of Conflict Resolution* 51.2, pp. 251–284.

Byrne, Hugh (1996). *El Salvador's Civil War: a Study of Revolution*. Lynne Rienner.

Cantir, Cristian and Juliet Kaarbo (2012a). "Contested Roles and Domestic Politics: Reflections on Role Theory in Foreign Policy Analysis and IR Theory." In: *Foreign Policy Analysis* 8.1, pp. 5–24.

– (2012b). "Contested roles and domestic politics: reflections on role theory in foreign policy analysis and IR theory." In: *Foreign Policy Analysis* 8.1, pp. 5–24.

Carment, David and Dane Rowlands (1998). "Three's Company Evaluating Third-Party Intervention in Intrastate Conflict." In: *Journal of Conflict Resolution* 42.5, pp. 572–599.

Carothers, Tom H (1991). *In the Name of Democracy: US Policy toward Latin America in the Reagan Years*. University of California Press.

Chen, Adrian (2015). *The Agency*. Ed. by The New York Times Magazine.

Chenoweth, Erica and Maria J Stephan (2011). *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. Columbia University Press.

Chernyshov, Dmitry (2014). *Guideline (to discover political trolls)*.

Chib, Siddhartha (1996). "Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models." In: *Journal of Econometrics* 75, pp. 79–97.

– (1998). "Estimations and Comparison of Multiple Change-point Models." In: *Journal of Econometrics* 86, pp. 221–241.

Cho, In-Koo and David M Kreps (1987). "Signaling games and stable equilibria." In: *The Quarterly Journal of Economics* 102.2, pp. 179–221.

Chyzh, Olga V. and Elena Labzina (n.d.). "Bankrolling Repression? Modeling Third-Party Influence on Protests and Repression." In: *American Journal of Political Science* 0.0. DOI: 10.1111/ajps.12341. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12341.

Ciment, James (2015). *Encyclopedia of Conflicts Since World War II*. Taylor & Francis.

Ciment, James, Kenneth L. Hill, Carl Skutsch, and David. MacMichael (1998). *Encyclopedia of Conflicts Since World War II*. Vol. 2. Taylor & Francis.

CNN (2014). *Ukraine cries 'robbery' as Russia annexes Crimea.*

Committee to Protect Journalists (2017). `https://cpj.org`.

Conrad, Courtenay R (2014). "Divergent Incentives for Dictators Domestic Institutions and (International Promises Not to) Torture." In: *Journal of Conflict Resolution* 58.1, pp. 34–67.

Conrad, Courtenay Ryals and Will H Moore (2010). "What Stops the Torture?" In: *American Journal of Political Science* 54.2, pp. 459–476.

Cook, Bernard A (2001). *Soviet Soft Power in Poland: Culture and the Making of Stalin's New Empire, 1943-1957*. Taylor Francis.

Couloumbis, Theodore A. (1966). *Greek Political Reaction to American and NATO Influences*. Yale University Press.

Coyne, Christopher J. and Peter T. Leeson (2009). *Media, Development, and Institutional Change*. Cheltenham, Gloucestershire, UK: Edward Elgar Publishing.

CSIS (2016). *The Ukrainian Crisis Timeline*.

Cunningham, David E (2010). "Blocking Resolution: How External States Can Prolong Civil Wars." In: *Journal of Peace Research* 47.2, pp. 115–127.

Dagiev, Dagikhudo (2013). *Regime Transition in Central Asia: Stateness, Nationalism and Political Change in Tajikistan and Uzbekistan*. Taylor & Francis.

Daugherty, Leo J. (2015). *Counterinsurgency and the United States Marine Corps*. McFarland, Incorporated Publishers.

Davenport, Christian (2007). *State Repression and the Domestic Democratic Peace*. Cambridge University Press.

Davenport, Christian and David A Armstrong (2004). "Democracy and the Violation of Human Rights: a Statistical Analysis from 1976 to 1996." In: *American Journal of Political Science* 48.3, pp. 538–554.

Davenport, Christian, David A Armstrong, and Will H. Moore (2008). *Waterboarding in a Democracy? Torture, Political Threats and the Palliative Impact of Democratic Institutions*. Available at `https://whmoore.net/research/research-papers/`.

Davis, David and Will Moore (1997). "Ethnicity Matters." In: *International Studies Quarterly* 41.1, pp. 171–184.

de Lama, George (1985). "Reagan Frees $67 Million In Aid To Troubled Sudan." In: *Chicago Tribune* April 2. URL: `http://articles.chicagotribune.com/1985-04-02/news/8501180686_1_aid-package-sudanese-senior-administration-official` Accessed 6/6/2017.

Debs, Alexandre and Hein E Goemans (2010). "Regime type, the Fate of Leaders, and War." In: *American Political Science Review* 104.03, pp. 430–445.

Dimitrov, Vesselin (2013). *Bulgaria: The Uneven Transition*. Taylor & Francis.

Dowell, Nia M., Leah C. Windsor, and Arthur C. Graesser (2015). "Computational linguistics analysis of leaders during crises in authoritarian regimes." In: *Dynamics of Asymmetric Conflict* doi:10.1080/17467586.2015.1038286.

Egorov, Georgy, Sergei Guriev, and Konstantin Sonin (2009). "Why Resource-poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data." In: *American Political Science Review* 103.4, pp. 645–668.

Ellison, Herbert J. (1988). *A Significant Gesture From Moscow : Pullback in Mongolia Signals Sino-Soviet Cooperation*. URL: `http://articles.latimes.com/1988-12-27/local/me-962_1_soviet-relations`, Accessed 6/8/2017. December 27.

Findley, Michael G and Tze Kwang Teo (2006). "Rethinking Third-Party Interventions into Civil Wars: An Actor-Centric Approach." In: *Journal of Politics* 68.4, pp. 828–837.

Fish, M Steven (2017). "What Is Putinism?" In: *Journal of Democracy* 28.4, pp. 61–75.

Forsythe, David P (1992). "Democracy, War, and Covert Action." In: *Journal of Peace Research* 29.4, pp. 385–395.

Frazier, Derrick and Robert Stewart-Ingersoll (2010). "Regional Powers and Security: A Framework for Understanding Order within Regional Security Complexes." In: *European Journal of International Relations* 16.4, pp. 731–753.

Fruhwirth-Schnatter, Sylivia and Helga Wagner (2006). "Auxiliary Mixture Sampling for Parameter-Driven Models of Time Series of Small Counts with Applications to State Space Modelling." In: *Biometrika* 93.4, pp. 827–841.

Gallagher, Jim (1982). "Jaruzelski to ask for Soviet Help." In: *Chicago Tribune*. URL: http://archives.chicagotribune.com/1982/03/01/page/2/article/jaruzelski-to-ask-for-soviet-help.

Galván, Javier A (2012). *Latin American dictators of the 20th century: The Lives and Regimes of 15 Rulers*. McFarland.

Garmajalova, Alexandra (2013). *Where trolls live. And who feeds them*. Ed. by Novaya Gazeta.

Gartner, Scott Sigmund and Patrick M Regan (1996). "Threat and Repression: The Non-linear Relationship between Government and Opposition Violence." In: *Journal of Peace Research* 33.3, pp. 273–287.

Gehlbach, Scott and Konstantin Sonin (2014). "Government Control and the Media." In: *Journal of Public Economics* 118, pp. 163–171.

Gent, Stephen E (2008). "Going in When it Counts: Military Intervention and the Outcome of Civil Conflicts." In: *International Studies Quarterly* 52.4, pp. 713–735.

Gentleman, Amelia (2014). *Putin asserts right to use force in east Ukraine*. Ed. by The$_G$*uardian*.

Gibler, Douglas M. (1999). "East or Further East?" In: *Journal of Peace Research* 36.6, pp. 627–637.

Gibler, Douglas M. and Jamil A. Sewell (2006). "External Threat and Democracy: The Role of NATO Revisited." In: *Journal of Peace Research* 43.4, pp. 413–431.

Gill, Jeff (2009). *Bayesian Methods: A Social and Behavior Sciences Approach*. Second. Boca Raton, FL: Chapman and Hall/CRC Press.

Goemans, Hein E (2008). "Which Way Out? The Manner and Consequences of Losing Office." In: *Journal of Conflict Resolution* 52.6, pp. 771–794.

Goemans, Henk E, Kristian Skrede Gleditsch, and Giacomo Chiozza (2009). "Introducing Archigos: A Dataset of Political Leaders." In: *Journal of Peace Research* 46.2, pp. 269–283.

Gonzales, Michael J (2002). *The Mexican Revolution, 1910-1940*. University of New Mexico Press.

Gordon (2015). *Bloggers discover the troll factories thanks to the new Google service*. Ed. by Gordon.

Granville, Johanna (2006). "East Germany in 1956: Walter Ulbricht's Tenacity in the Face of Opposition." In: *Australian Journal of Politics and History* 52.5, 417438.

Grigorieva, Inna and Pavel Chuvilyaev (2012). *Valley of the Trolls. What Russian Internet gradually turns into*. Ed. by Novaya Gazeta.

Grimmer, Justin and Brandon M. Stewart (2013). "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." In: *Political Analysis* 21.3, pp. 267–297.

Gunitsky, Seva (2015). "Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability." In: *Perspectives on Politics* 13.1, pp. 43–54.

Hassid, Jonathan (2008). "Controlling the Chinese Media: An Uncertain Business." In: *Asian Survey* 48.3, pp. 414–430.

Hastedt, Glenn P. (2014). *Cases in International Relations: Pathways to Conflict and Cooperation*. SAGE Publications.

Hegelich, Simon and Dietmar Janetzko (2016). "Are Social Bots on Twitter Political Actors? Empirical Evidence from a Ukrainian Social Botnet." In: *Tenth International AAAI Conference on Web and Social Media*.

Hevesi, Dennis (1988). *Mohammad Zia ul-Haq: Unbending Commander for Era of Atom and Islam*. URL:`http://www.nytimes.com/1988/08/18/world/mohammad-zia-ul-haq-`

unbending-commander-for-era-of-atom-and-islam.html?pagewanted=all, Accessed 6/8/2017. August 18.

Higgins, Andrew and Peter Baker (2014). "Russia Claims U.S. is Meddling Over Ukraine." In: *New York Times* 6 Feb. https://www.nytimes.com/2014/02/07/world/europe/ukraine.html.

Hill, Daniel W and Zachary M Jones (2014). "An Empirical Evaluation of Explanations for State Repression." In: *American Political Science Review* 108.03, pp. 661–687.

Hill, Fiona (2013). "The Real Reason Putin Supports Assad." In: *Foreign Affairs* 25.

Holsti, K. J. (1970). "National Role Conceptions in the Study of Foreign Policy." In: *International Studies Quarterly* 14.3, pp. 260–70.

Hopkins, Daniel and Gary King (2010). "Extracting Systematic Social Science Meaning from Text." In: *American Journal of Political Science* 54.1, pp. 229–247.

Huang, Haifeng (2015). "Propaganda as Signaling." In: *Comparative Politics* 47.4, pp. 419–437.

IBP, Inc (2012). *Philippines Country Study Guide Volume 1 Strategic Information and Developments*. International Business Publications, USA.

Inozemtsev, Vladislav (2017). "Why Putinism Arose." In: *Journal of Democracy* 28.4, pp. 80–85.

Insider, Business (2015). *Putin: 'I will say this clearly: There are no Russian troops in Ukraine'*.

Interfax-Ukraine (2013). *Ukraine to issue Eurobonds; Russia will purchase* $15 *bln, says Russian finance minister*. http://en.interfax.com.ua/news/economic/182491.html. Retrieved on September 30, 2015.

International Business Publications, USA. and U. Ibp (2004). *Morocco Foreign Policy And Government Guide*. 1. International Business Publications, USA.

Its time for the U.S. to rethink its approach to Uganda (2016). URL: https://www.washingtonpost.com/opinions/its-time-for-the-us-to-rethink-its-approach-to-uganda/2016/04/08/1b1d80ae-fce1-11e5-80e4-c381214de1a3_story.html?utm_term=.91134dbab467, Accessed 6/6/2017. April 8.

Iyengar, Shanto and Kyu S. Hahn (2009). "Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use." In: *Journal of Communication* 59.1, pp. 19–39.

Iyengar, Shanto, Kyu S. Hahn, Jon A. Krosnick, and John Walker (2008). "Selective Exposure to Campaign Communication: The Role of Anticipated Agreement and Issue Public Membership." In: *Journal of Politics* 70.1, pp. 186–200.

Jazeera, Al (2014). *Putin says no Russian troops in east Ukraine*.

Kara-Murza, Vladimir V (2017). "Putin Is not Russia." In: *Journal of Democracy* 28.4, pp. 110–116.

Kass, Robert E. and Adrian E. Raftery (1995). "Bayes Factor." In: *Journal of the American Statistical Association* 90.430, pp. 773–795.

Keller, Franziska, David Schoch, Sebastian Stier, and JungHwan Yang (2017). "Identifying and Characterizing Electoral Astroturfing: Evidence from South Korea." Working Paper.

King, Gary, Jennifer Pan, and Margaret E Roberts (2013a). "How Censorship in China Allows Government Criticism but Silences Collective Expression." In: *American Political Science Review* 107.2, pp. 326–343.

– (2013b). "How censorship in China allows government criticism but silences collective expression." In: *American Political Science Review* 107.02, pp. 326–343.

– (2016). "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument." In: *Copy at http://j. mp/1Txxiz1 Download Citation BibTex Tagged XML Download Paper* 2.

Kivelson, Valerie A. and Ronald Grigor Suny (2016). *Russia's Empires*. Oxford University Press.

Knarr, James C (2012). "Uruguay and the United States, 1903–1929: Diplomacy in the Progressive Era." In:

Kohlberg, Elon (1990). "Refinement of Nash equilibrium: The main ideas." In: *Game Theory and Applications*. Elsevier, pp. 3–45.

Kohlberg, Elon and Jean-Francois Mertens (1986). "On the strategic stability of equilibria." In: *Econometrica: Journal of the Econometric Society*, pp. 1003–1037.

Kornbluh, Peter (2013). *The Pinochet File: A Declassified Dossier on Atrocity and Accountability*. The New Press.

Kydd, Andrew H (2006). "When Can Mediators Build Trust?" In: *American Political Science Review* 100.03, pp. 449–462.

Kydd, Andrew H. and Roseanne W. McManus (2015). "Threats and Assurances in Crisis Bargaining." In: *Journal of Conflict Resolution* doi:10.1177/002200271557657.

Labzina, Elena and Mark Nieman (2017). "State-controlled Media and Foreign Policy: Analyzing Russian-language News." Working paper.

Labzina, Elena and Mark David Nieman (2017). *State-controlled Media and Foreign Policy: Analyzing Russian-language News*.

Lake, David A. (2009). *Hierarchy in International Relations*. Cornell University Press.

Lally, Kathy and Will Englund (2014). *Putin says he reserves right to protect Russians in Ukraine*. Ed. by The Washington Post.

Lavrov, Sergei (2006). "The Rise of Asia, and the Eastern Vector of Russia's Foreign Policy." In: *Russia in Global Affairs* 4.3, pp. 68–80.

Leapman, Michael (1998). "British interests, Nigerian tragedy." In: *Independent* January 4. URL: `http://www.independent.co.uk/voices/british-interests-nigerian-tragedy-1136684.html`, Accessed 6/6/2017.

Lebanon's Camille Chamoun Dies; Leader in Fight for Independence (1987). URL: http://articles.latimes.com/1987-08-08/news/mn-300_1_lebanon-s-christians, Accessed 6/6/2017. August 8.

Lee, Chin-Chuan (1998). "Press Self-Censorship and Political Transition in Hong Kong." In: *Harvard International Journal of Press/Politics* 3.2, pp. 55–73.

Lemke, Douglas (2002). *Regions of War and Peace*. Cambridge University Press.

Leonard, Thomas M (2004). *Encyclopedia of Cuban-United States Relations*. McFarland.

Levy, Jack S. and William R. Thompson (2010). *Causes of War*. Chichester, UK: Wiley-Blackwell.

Licht, Amanda A (2010). "Coming into Money: The Impact of Foreign Aid on Leader Survival." In: *Journal of Conflict Resolution* 54.1, pp. 58–87.

Lo, Bobo (2002). *Russian Foreign Policy in the Post-Soviet Era: Reality, Illusion and Mythmaking*. London: Palgrave Macmillan.

– (2004). "The Long Sunset of Strategic Partnership: Russia's Evolving China Policy." In: *International Affairs* 80.2, pp. 295–309.

Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley (2016). "Computer-Assistaed Text Analysis for Comparative Politics." In: *Political Analysis* 23.2, pp. 254–277.

MacFarlane, Neil (2006). "The 'R' in BRICs: Is Russia an Emerging Power?" In: *International Affairs* 82.1, pp. 41–57.

MacFarquhar, Neil (2018). "Inside the Russian Troll Factory: Zombies and a Breakneck Pace." In: *The New York Times*. URL: https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html.

Marks, Thomas A. and Tej Pratap Singh Brar (2016). *Sri Lanka: State Response to the Liberation Tigers of Tamil Eelam as an Illicit Power Structure*. URL: http://cco.

ndu.edu/News/Article/780214/chapter-9-sri-lanka-state-response-to-the-liberation-tigers-of-tamil-eelam-as-a/, Accessed 6/8/2017. May 24.

Martin, Andrew D., Kevin M. Quinn, and Jong Hee Park (2011). "MCMCpack: Markov chain Monte Carlo in R." In: *Journal of Statistical Software* 42.9, pp. 1–21.

Martinez Machain, Carla and T Clifton Morgan (2013). "The Effect of US Troop Deployment on Host States Foreign Policy." In: *Armed Forces & Society* 39.1, pp. 102–123.

Mas-Colell, Andreu, Michael Dennis Whinston, Jerry R Green, et al. (1995). *Microeconomic theory*. Vol. 1. Oxford university press New York.

McManus, Roseanne W. (2014). "Fighting words The effectiveness of statements of resolve in international conflict." In: *Journal of Peace Research* 51.6, pp. 726–740.

– (2016). "The Impact of Context on the Ability of Leaders to Signal Resolve." In: *International Interactions* doi:10.1080/03050629.2016.1171764.

Meduza (2017). *Russian journalists publish massive investigation into St. Petersburg troll factory's U.S. operations*. Ed. by Meduza.

Messing, Solomon and Sean J. Westwood (2014). "Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation when Selecting News Online." In: *Communication Research* 41.8, pp. 1042–1063.

Moehler, Devra C. and Naunihal Singh (2011). "Whose News Do You Trust? Explaining Trust in Private versus Public Media in Africa." In: *Political Research Quarterly* 64.2, pp. 276–292.

Moore, Will H (1995). "Action-Reaction or Rational Expectations? Reciprocity and the Domestic-International Conflict Nexus during the Rhodesia Problem." In: *Journal of Conflict Resolution* 39.1, pp. 129–167.

Moser, Michael (2014). *Language Policy and Discourse on Languages in Ukraine Under President Viktor Yanukovych*. Columbia University Press.

New York Times (2003). *In $87.5 Billion Bill, $2 Million Bounty for Exiled Liberian*. URL: http://www.nytimes.com/2003/11/10/world/in-87.5-billion-bill-2-million-bounty-for-exiled-liberian.html, Accessed 6/6/2017. November 11.

News, CBS (2014). *Putin insists no Russian troops in eastern Ukraine*.

Nieman, Mark David (2016a). "Moments in Time: Temporal Patterns in the Effect of Democracy and Trade on Conflict." In: *Conflict Management and Peace Science* 33.3, pp. 273–293.

– (2016b). "The Return on Social Bonds: Social Contracts and International Conflict." In: *Journal of Peace Research* 53.3, pp. 665–679.

– (2016c). "The Return on Social Bonds: Social Hierarchy and International Conflict." In: *Journal of Peace Research* 53.5, pp. 665–679.

Nygren, Bertil (2008). "Putin's Use of Natural Gas to Reintegrate the CIS Region." In: *Problems of Post-Communism* 55.4, pp. 3–15.

Oates, Sarah (2007). "The Neo-Soviet Model of the Media." In: *Europe-Asia Studies* 59.8, pp. 1279–1297.

Oberdorfer, Don (1980). *The Making of a Soviet Coup*. URL: https://www.washingtonpost.com/archive/politics/1980/01/02/the-making-of-a-soviet-coup/175c9707-dddb-4052-87be-9bf4f048bf56/?utm_term=.39bc6c51bcf7, Accessed 6/8/2017. January 2.

OECD (2016). "Main Economic Indicators." In: 2016.10.

Olson, J.Stuart, L. Brigance Pappas, and N. Charles Pappas (1994). *An Ethnohistorical Dictionary of the Russian and Soviet Empires*. Greenwood Press.

Park, Jong Hee (2010). "Structural Change in U.S. Presidents' Use of Force." In: *American Journal of Political Science* 54.3, pp. 766–782.

Pascoe, William H. (1986). "Mozambique Is A Test For Reagan." In: *Chicago Tribune* November 20. URL: http://articles.chicagotribune.com/1986-11-20/news/8603270269_

1_mozambique-national-resistance-renamo-president-samora-machel, Accessed 6/6/2017.

Patrick, Peebles (2006). *The History of Sri Lanka*. Greenwood Press.

Paul, Christopher (2008). *Marines on the Beach: The Politics of US Military Intervention Decision Making*. Greenwood Publishing Group.

Peterson, Nolan (2015). "Why a Russian is Fighting for Ukraine." In: *Newsweek* April 8. URL: http://www.newsweek.com/why-russian-fighting-ukraine-320723, Accessed 6/7/2017.

Pierskalla, Jan Henryk (2010). "Protest, Deterrence, and Escalation: The Strategic Calculus of Government Repression." In: *Journal of Conflict Resolution* 54.1, pp. 117–145.

Poggo, Scopas S (2008). *The First Sudanese Civil War: Africans, Arabs, and Israelis in the Southern Sudan, 1955-1972*. Palgrave Macmillan US.

Prior, Markus (2013). "Media and Political Polarization." In: *Annual Review of Political Science* 16, pp. 101–127.

Rahr, Alexander (2007). "Germany and Russia: A Special Relationship." In: *Washington Quarterly* 30.2, pp. 137–145.

Redlawsk, David P. (2002). "Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." In: *Journal of Politics* 64.4, pp. 1021–1044.

Regan, Patrick M (1998). "Choosing to Intervene: Outside Interventions in Internal Conflicts." In: *The Journal of Politics* 60.3, pp. 754–779.

Regan, Patrick M. (2002). "Third-Party Interventions and the Duration of Intrastate Conflicts." In: *Journal of Conflict Resolution* 46.1, pp. 55–73.

Regan, Patrick M. and Sam R. Bell (2010). "Changing Lanes or Stuck in the Middle: Why Are Anocracies More Prone to Civil Wars?" In: *Political Research Quarterly* 63.4, pp. 747–759.

Regan, Patrick M and M Scott Meachum (2014). "Data on Interventions During Periods of Political Instability." In: *Journal of Peace Research* 51.1, pp. 127–135.

Renouard, J. (2015). *Human Rights in American Foreign Policy: From the 1960s to the Soviet Collapse*. University of Pennsylvania Press.

Richards, David L., Alyssa Webb, and K. Chad Clay (2015). "Respect for Physical-Integrity Rights in the Twenty-First Century." In: *Journal of Human Rights* 14.3, pp. 291–311.

Ritter, Emily Hencken (2014). "Policy Disputes, Political Survival, and the Onset and Severity of State Repression." In: *Journal of Conflict Resolution* 58.1, pp. 143–168.

Ritter, Emily Hencken and Courtenay R Conrad (2016). "Preventing and Responding to Dissent: The Observational Challenges of Explaining Strategic Repression." In: *American Political Science Review* 110.1, pp. 85–99.

Robertson, Graeme and Samuel Greene (2017). "How Putin Wins Support." In: *Journal of Democracy* 28.4, pp. 86–100.

Roman, Eric (2003). *Austria-Hungary & the Successor States: A Reference Guide from the Renaissance to the Present*. Infobase Publishing.

Rule, Sheila (1988). "Soviet Arms Aid Is Seen as Pivotal to Ethiopia." In: *The New York Times* April 30. URL: `http://www.nytimes.com/1988/04/30/world/soviet-arms-aid-is-seen-as-pivotal-to-ethiopia.html` Accessed 6/6/2017.

Salamey, Imad (2013). *The Government and Politics of Lebanon*. Taylor & Francis.

Salganik, Matthew J (2017). *Bit by bit: social research in the digital age*. Princeton University Press.

San-Akca, Belgin (2016). *States in Disguise: Causes of State Support for Rebel Groups*. Oxford University Press.

Sanovich, Sergey, Denis Stukal, and Joshua A. Tucker (2017). "Detecting Political Bots on Russian Twitter." Working Paper.

Schimpfossl, Elisabeth and Ilya Yablokov (2014). "Coercion or Conformism? Censorship and Self-censorship Among Russian Media Personalities and Reporters in the 2010s." In: *Demokratizatsiya* 22.2, pp. 295–311.

Scranton, Margaret E (1991). *The Noriega Years: US-Panamanian Relations, 1981-1990*. Lynne Rienner.

Selten, Reinhard (1975). "Reexamination of the perfectness concept for equilibrium points in extensive games." In: *International journal of game theory* 4.1, pp. 25–55.

Semple, Robert B. (1970). *Nixon Sends Combat Forces to Cambodia to Drive Communist from Staging Zone*. URL:`http://www.nytimes.com/1970/05/01/archives/nixon-sends-combat-forces-to-cambodia-not-an-invasion-president.html`, Accessed 6/8/2017. May 1.

Sheinin, David MK (2010). *Argentina and the United States: An Alliance Contained*. University of Georgia Press.

Shevtsova, Lilia (2017). "Paradoxes of Decline." In: *Journal of Democracy* 28.4, pp. 101–109.

Shi, Tianjian, Jie Lu, and John Aldrich (2011). "Bifurcated Images of the U.S. in Urban China and the Impact of Media Environment." In: *Political Communication* 28.3, pp. 357–376.

Sobolev, Anton (2017). "Fantastic Beasts and Whether They Matter: How Internet 'Trolls' Influence Political Conversations in Russia." Working Paper.

StemmersNet (2012). *Snowball stemmers on C#*. `https://stemmersnet.codeplex.com/`. Accessed: 2017-10-15.

Stockmann, Daniela (2011). "Race to the Bottom: Media Marketization and Increasing Negativity Toward the United States in China." In: *Political Communication* 28.3, pp. 268–290.

Stockmann, Daniela and Mary E. Gallagher (2011). "Remote Control: How the Media Sustain Authoritarian Rule in China." In: *Comparative Political Studies* 44.4, pp. 436–467.

Stroud, Natalie Jomini (2011). *Niche News: The Politics of News Choice*. Oxford University Press.

Taber, Charles S. and Milton Lodge (2006). "Motivated Skepticism in the Evaluation of Political Beliefs." In: *American Journal of Political Science* 50.3, pp. 755–769.

Tarling, Nicholas (1999). *The Cambridge History of Southeast Asia*. Cambridge University Press.

Thies, Cameron G. (2013). *The United States, Israel, and the Search for International Order: Socializing States*. New York: Routledge.

Thies, Cameron G. and Marijke Breuning (2012). "Integrating Foreign Policy Analysis and International Relations through Role Theory." In: *Foreign Policy Analysis* 8.1, pp. 1–4.

Thies, Cameron G. and Mark David Nieman (2017). *Rising Powers and Foreign Policy Revisionism: Understanding BRICS Identity and Behavior through Time*. Ann Arbor, MI: University of Michigan Press.

Thies, Cameron S. and Mark Nieman (2017). *Rising Powers and Foreign Policy Revisionism: Understanding BRICS Identity and Behavior Through Time*.

Thomas, Jakana L, William Reed, and Scott Wolford (2016). "The Rebels Credibility Dilemma." In: *International Organization* 70.03, pp. 477–511.

Thompson, Wayne C. (2012). *Nordic, Central and Southeastern Europe*. Rowman & Littlefield Publishers.

Trenin, Dmitri (2009a). "Reimagining Moscow's Foreign Policy." In: *Foreign Affairs* 88.6, pp. 64–78.

– (2009b). "Russia's Spheres of Interest, Not Influence." In: *The Washington Quarterly* 32.4, pp. 3–22.

Tsygankov, Andrei P. and Matthew Tarver-Wahlquist (2009). "Duelling Honors: Power, Identity and the Russia–Georgia Divide." In: *Foreign Policy Analysis* 5.4, pp. 307–326.

Tucker, Spencer C. (2010). *The Encyclopedia of Middle East Wars*. ABC-CLIO.

Tucker, Spencer C. (2014). *World War I: The Definitive Encyclopedia and Document Collection*. ABC-CLIO, LLC.

Valentino, Nicholas A., Antoine J. Banks, Vincent L. Hutchings, and Anne K. Davis (2009). "Selective Exposure in the Internet Age: The Interaction between Anxiety and Information Utility." In: *Political Psychology* 30.4, pp. 591–613.

Vallone, Robert P., Lee Ross, and Mark R. Lepper (1985). "The Hostile Media Phenomenon: Biased Perception and Perceptions of Media Bias in Coverage of the Beirut Massacre." In: *Journal of Personality and Social Psychology* 49.3, pp. 577–585.

Vanderhill, R. et al. (2013). *International Dimensions of Authoritarian Persistence: Lessons from Post-Soviet States*. Lexington Books.

Vartanova, Elena (2012). "The Russian Media Model in the Context of Post-Soviet Dynamics." In: *Comparing Media Systems Beyond the Western World*. Ed. by Daniel C. Hallin and Paolo Mancini. Cambridge University Press New York, NY, pp. 119–142.

Walker, Shaun (2015). "Putin Admits Russian Military Presence in Ukraine for First Iime." In: *The Guardian* 17 December. `https://www.theguardian.com/world/2015/dec/17/vladimir-putin-admits-russian-military-presence-ukraine`.

Walker, Stephen G. (1992). "Symbolic Interactionism and International Politics: Role Theory's Contribution to International Organization." In: *Contending Dramas: A Cognitive Approach to International Organizations*. Ed. by Martha Cottam and Chih-Yu Shih. New York: Praeger.

Warren, Camber (2014). "Not by the Sword Alone: Soft Power, Mass Media, and the Production of State Sovereignty." In: *International Organization* 68.1, pp. 111–141.

– (2015). "Explosive Connections? Mass Media, Social Media, and the Geography of Collective Violence in African States." In: *Journal of Peace Research* 52.3, pp. 297–311.

Wendt, Alexander (1994). "Collective Identity Formation and the International State." In: *American Political Science Review* 88.2, pp. 384–396.

Whitaker, Brian (2013). *Hizbullah Backs Syria to Force Political Divide in Lebanon*. URL: https://www.theguardian.com/world/2005/mar/07/syria.lebanon, Accessed 6/8/2017. March 7.

Whitmore, Brian (2017). "Belarus And The Ghosts Of 2014." In: *Radio Free Europe* March 17. URL: https://www.rferl.org/a/belarus-and-the-ghosts-of-2014/28375750.html, Accessed 6/7/2017.

Windsor, Leah C., Nia Dowell, and Art Graesser (2014). "The Language of Autocrats: Leaders' Language in Natural Disaster Crises." In: *Risk, Hazards & Crisis in Public Policy* 5.4, pp. 446–467.

Windsor, Leah, Nia Dowell, Alistair Windsor, and John Kaltner (2017). "Leader Language and Political Survival Strategies." In: *International Interactions*, pp. 1–16.

Wolfers, Arnold (1984). *Discord and Collaboration: Essays on International Politics*. Hopkins.

Zaller, John (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.

– (1996). "The Myth of Massive Media Impact Revived: New Support for a Discredited Idea." In: *Political Persuasion and Attitude Change*. Ed. by Diane Carole Mutz, Paul M. Sniderman, and Richard A. Brody. Ann Arbor, MI: University of Michigan Press.