Appendix

Author names redacted

2020-12-02

This appendix accompanies the paper “After Deterrence: Explaining Conflict Short of War”. It provides supplemental information concerning proofs for the formal model, the dataset of Russian gray zone campaigns introduced in the paper, and robustness checks and alternate specifications for the statistical model.

# Formal Model

## Formal statement of assumptions

First, we express the assumption that the kinks in the P function are never activated in equilibrium. Letting and denote the optimal levels selected by C and D conditional on the actors selecting into gray zone conflict (these are defined below), when Assumption 1 holds, the ``min-max’’ statements in the function will never be relevant to analysis.

Second, we express the assumption that if C’s resolve increases, C becomes more willing to go to war over using gray zone conflict. As some intuition, conditional on gray zone conflict occurring, C selects one of two values for . For the first value, the selected will be the largest possible that is tailored to keep D from going to war. This is . For the second value, the selected will be based on C’s own resolve and represents the solution to C’s internal optimization problem or C’s internal efficiency. This is . For C’s utility from war to be increasing in at a faster rate than the utility from gray zone conflict, we must consider both values of .

## Proving Proposition 1

### Equilibrium Intuition

Outside of gray zone conflict, C will prefer the status quo to initially going to war when or

Here we discuss the intuition of the equilibrium in the paper. Assume for now that C is optimally selecting a such that the game ends in gray zone conflict (in other words assume that and ). Also assume that D selects an optimal such that (this will be borne out by Assumption 1). D selects characterized by I take first-order conditions with respect to and solve the expression above to identify the optimal level of D’s gray zone response . This unique value is

Using the expression for , D’s utility in terms of the selected is .

I can then begin considering C’s utility. There are two matters to consider. First, it could be that C will select an optimal that is constrained by D’s willingness to go to war. Essentially, if , then D’s utility from war is greater than D’s utility from gray zone conflict; thus, if C wants to remain in gray zone conflict and will be constrained by D’s deterrent threat, C will select , where is the greatest that would make D indifferent between gray zone conflict and war, or Second C may select an optimal that is constrained by their own internal costs. When this is the case, C will select , defined by the optimization which yields Before discussing the true behavior, we highlight two things that do not happen. First, note that C will never select an that provokes D to go to war in the final stage, because this is strictly worse than initially going to war. Second, note that C will never select into gray zone conflict (i.e. set and ) if as defined above is greater than because C could do strictly better not paying the costs of war and selecting into the status quo ().

With this is place, if C optimally selects into gray zone conflict, C will select , where We have now characterized what happens withing gray zone conflict. We now need to describe how the game optimally plays out across the possibility of selecting into the status quo, war (at the onset; ), or gray zone conflict. Because C moves first, this is ultimately C’s choice. We can calculate C’s decision within the two cases of gray zone conflict:

First, we consider the case when . This condition implies that the selected gray zone conflict will be constrained by D’s deterrent threat and not C’s internal costs. So, if C selects into gray zone conflict, C will select . We can then express C’s behavior in terms of . C prefers the status quo to gray zone conflict when or Note that the above derivation relies on , lest the inequality sign would flip. This is assumed by Assumption 1.

Next, C prefers war to gray zone conflict when or Note that the above derivation relies on , lest the inequality sign would flip. this is assumed by Assumption 2.

Next, we assume . This condition implies that the selected gray zone conflict will be constrained by C’s internal costs and not D’s deterrent threat. So, if C selects into gray zone conflict, C will select . We can then express C’s behavior in terms of . C prefers the status quo to gray zone conflict when or

Next, C prefers war to gray zone conflict when or Note that the above derivation relies on , lest the inequality sign would flip. This is implied by Assumption 2.

With all of this defined, we can characterize C’s strategy in terms of ; as increases, C prefers more degrees of conflict (i.e. larger ’s or war) to get what they want.

### Equilibrium Behavior

Proposition 1A and the text below contains a more complete discussion on the equilibrium behavior characterized in Proposition 1.

Working backwards, D will declare war for all . If , D will select . When , D is selecting their optimal level of gray zone response based on their internal optimization. When , it implies that D would be willing to select a greater gray zone response, but does not need to, essentially driving the political impact of C’s limited challenges back to zero (at cost).

## Observation 1 Discussion

Assume for now the parameters are such that the Case 1.C. conditions hold, and consider what happens when decreases. Because here C selects the greatest level of limited challenges that will not provoke D to war, C’s selected is a decreasing function of ; therefore, because is fixed, the final extent of gray zone conflict will be less. Of course, the analysis does not stop there. Improvements in D’s willingness to go to war constrain how useful gray zone conflict is to R, and, within case 1.C., C’s utility is decreasing in . Thus, if becomes small enough, C will leave gray zone conflict and instead select into either accepting the status quo (entering into case 1A) or going to war (entering into Case 1B). Additionally, it is worthwhile noting that as decreases, the condition that selects into Case 1 (over Case 2) has more slack, implying that improvements in D’s willingness to go to war will keep D in within Case 1.

Now assume the parameters are such that the Case 2.C. conditions hold, and consider what happens when decreases. Note that this will not change the selected here, but it could break the inequality that determines whether the equilibrium is defined in Case 1 or Case 2. thus, for a small enough , the conditions for Case 2 will break and the conditions for Case 1 will hold. When this happens, either the selected is increasing in (Case 1.C.) or gray zone conflict is not selected (Case 1.A. or 1.B.).

## Extension 1: Endogenous

In the model in the paper, we treated D’s gray zone efficiency as exogenous. In some special cases or under some conditions, this may be too strong an assumption. In this section, we characterize an equilibrium for the game when D can have complete flexibility in selecting some , where cannot equal zero lest D’s costs from their gray zone response will be undefined. The key take away from this extension is that if is endogenous (and its selection costless), then D’s selection of will be arbitrated by two properties. As the first property, it matters whether C prefers war to the status quo (formally, if C is type ), or C prefers the status quo to war (). When C prefers the status quo to war, then D is in a position where D can, by selecting a low enough , influence C to stop undertaking limited challenges and select into the status quo. Intuitively, when D is very good at gray zone conflict, D would select a high , which makes gray zone conflict less productive for C. But, when C prefers war to the status quo, then D could pressure C to stop undertaking limited challenges, but this will result in C going to war with D.

As the second property, D’s decision will also be arbitrated by whether D can select a gray zone efficiency that pushes C into a level of gray zone conflict where the deterrent threat does not bind. Recall that if C optimally conducts gray zone conflict, C selects , implying that C will either select an optimal based on their own internal cost-benefit analysis, or select an optimal tailored to make D indifferent between war and gray zone conflict (where the deterrent threat binds), with C ultimately choosing the smaller of the two. This means that if D can select a small enough so that , then C will selecting a level of limited challenge that is below the point that would make D indifferent between war and gray zone conflict, thus granting D some surplus.

The above two properties interact. Based on Assumptions 1 and 2, D will always prefer the status quo to gray zone conflict where the deterrent threat doesn’t bind, and gray zone conflict where the deterrent threat doesn’t bind to gray zone conflict where the deterrent threat does bind or war. Proposition A identifies how D selects in one possible equilibrium. Note that this is not the only possible equilibrium.

As one example of how this one equilibrium plays out, we adapt Figure 4 in the text. Now the solid black lines denote the selected levels of (with plotted so that greater y-axis values represent greater gray zone efficiencies for D), and the dotted lines separate equilibrium spaces.

Moving left to right, for between and , D’s optimal is described in Proposition A Case 1.A. As the outcome, C will optimally select into the status quo. For this selected , C knows that C would face enough of a challenge in gray zone conflict to make competing there too costly. Thus within this region, D could select a low enough to compel C to forgo limited challenges and conflict, and stick to the status quo.

Moving right, for between and , D’s optimal is described in Proposition A Case 2.A. As the outcome, C will optimally select into gray zone conflict, but will be constrained by C’s internal costs. For this selected , D wants to challenge C in gray zone conflict (which a lower accomplishes), but does not want to push C into forgoing gray zone conflict, because within this region C prefers war to accepting the status quo. Thus here, D selects the where C selects into gray zone conflict and is not bound by the deterrent threat, because this gives D some surplus beyond what war or C selecting gray zone conflict and being bound by the deterrent threat produces.

Finally, for between and , D’s optimal is described in Case 2.B. As the outcome, C will optimally select into gray zone conflict, and will be constrained by D’s deterrent threat. Essentially here, D is in a bad situation. If D modifies either C will adapt by selecting the new that makes D indifferent between war and gray zone conflict, or will go to war over the issue. Within this region, it does not matter what is selected, because C will always select an action that gives D their wartime utility.

## Extension 2: Probabilistic Escalation to War

A useful feature of the model above is that everything that occurs is deterministic. It is only if a state wants to go to war or wants to enter gray zone conflict does it actually happen. However, this may not perfectly represent reality. Perhaps in some cases, one state behaving aggressively in lower-levels of conflict can create an incident that necessitates an escalation to higher levels of conflict. To speak to this issue, we introduce the possibility of probabilistic escalation out of gray zone conflict. Our results are substantively similar, but this change shifts some equilibrium properties. Intuitively, now gray zone conflict can probabilistically lead to C’s worst outcome: where C invests in limited challenges, war happens, and C must pay the costs of limited challenges with the costs of war. Strategically, because here gray zone conflict is overall worse for R, C will be more willing to accept the status quo or go to war.

There are many possible ways to model this. For ease, we choose (in our opinion) the simplest way, which is that selecting introduces a likelihood of an escalation to war. Thus, when C selects , C’s new expected utility is To offer some intuition, , , , and remain the same as it was in the model in the text (as defined in Proposition 1). However, the cut-points that distinguish C’s decision to enter into the status quo, gray zone conflict, or war change slightly; overall, the key take-away is that considering probabilistic escalation makes gray zone conflict less appealing relative to the status quo and war.

I express equilibrium behavior in Proposition B. Then below, we derive the new cut-points, Additionally in the derivations, we discuss how the new cut-points imply that gray zone conflict is less appealing and fewer types will select into it relative to the game without a probabilistic likelihood of escalation to war from gray zone conflict.

### Equilibrium Intuition

First, we consider the case when . This implies that C will select . We can then express C’s behavior in terms of . C prefers the status quo to gray zone conflict when or Note that the inequality sign does not flip because, by Assumption 1, . We am able to say that because . Furthermore, this constraint (on when the status quo is preferred to gray zone conflict) matters only when C prefers the status quo to war, or when ; this condition implies , which means , which in turn implies that there are more C’s with some resolve that will select into the status quo in the game here relative to the game in the text without probabilistic escalation.

Next, C prefers war to gray zone conflict when or

Note that based on Assumption 2 (as is written: that ), the above sign does not flip. We can say that . This implies that

In other words, there are more C’s with some resolve that will select into war in the game here relative to the game without probabilistic escalation.

Next, we assume . This condition implies that the selected gray zone conflict will be constrained by C’s internal costs and not D’s deterrent threat. So, if C selects into gray zone conflict, C will select . We can then express C’s behavior in terms of . C prefers the status quo to gray zone conflict when or

To speak to this inequality, we will need to consider a few different cases here.

First, it could be possible that . When this is the case, then C would never want to select into gray zone conflict as doing so would always be strictly worse for R.

Next, consider when and . In this case, C’s wartime payoff is greater than C’s status quo payoff, meaning that C would never select into the status quo over selecting into war, meaning this constraint would never be activated.

Finally, consider when and . We can re-write the above as Where note that , where the inequality holds by Assumption 1. Altogether, this means that $\text{\ensuremath{\theta\left(\frac{\theta}{4\beta\_{C}}-\frac{1}{2\beta\_{D}}\right)}}>\theta\left(\zeta\left(\frac{\theta}{2\beta\_{C}}-\frac{1}{2\beta\_{D}}\right)-\frac{\theta}{4\beta\_{C}}\right)+(1-\zeta)(\theta\rho\_{W}-\theta\rho\_{0}-\kappa\_{C})$. This implies that there are more C’s with some resolve that will select into the status quo in the game here relative to the game without probabilistic escalation.

Finally, assuming , C prefers war to gray zone conflict when or Note the inequality sign does not slip because . Furthermore, by that condition, . Therefore . This implies that there are more C’s with some resolve that will select into war in the game here relative to the game without a random chance of escalation.

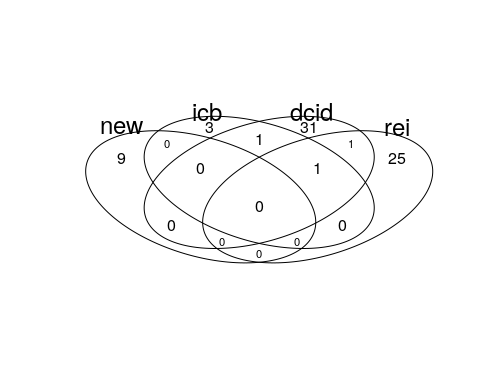
Finally, note that D’s strategies in this game are unchanged from the game without probabilistic escalation.

# New data

The universe of cases was created by first identifying cases of Russian foreign interventions from 3 prior datasets; ICB, DCID, and REI. Code replicating those findings is provided in the appropriate RMarkdown files. These cases were then supplemented with additional cases of Russian interference the authors were able to identify.

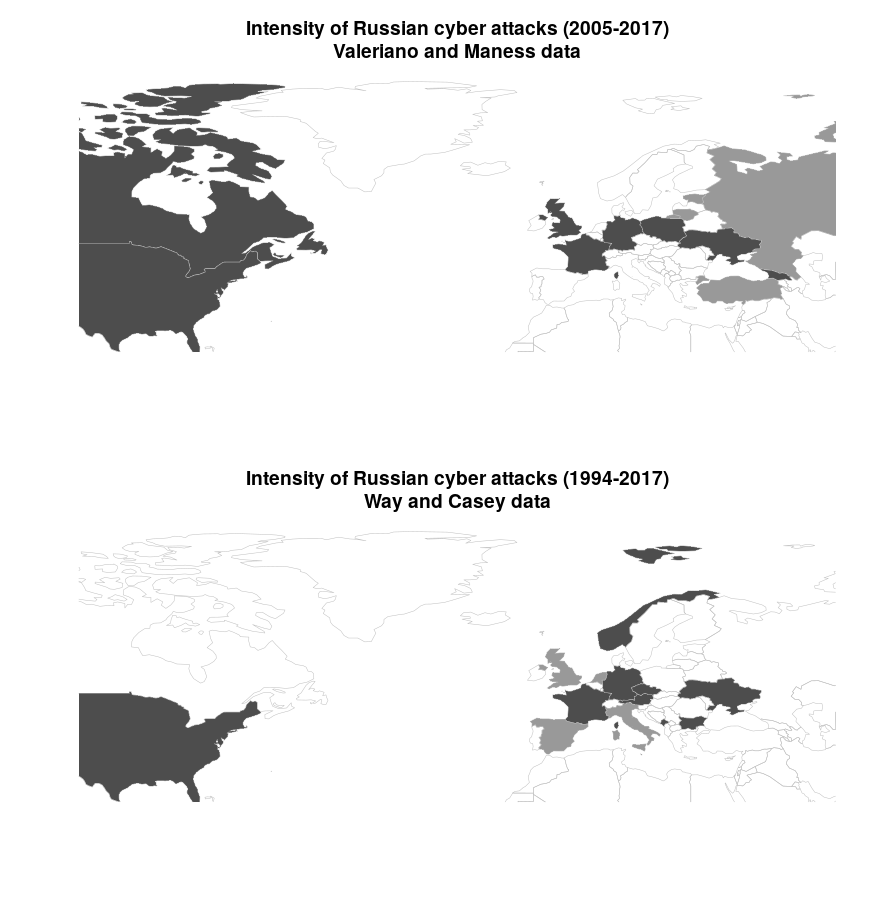
## Coverage of current datasets

A comparison of what cases were covered in each individual dataset is provided here:



## Consistency of current datasets

Aside from the cases covered, the intensity codings for current datasets are difficult to compare given their different scales. A more thorough analysis is provided in the appropriate R Markdown files, but a comparison of intensity codings in DCID (Valeriano and Maness) and REI (Way and Casey) is visualized here:



The DCID data identifies the United States, United Kingdom, Poland and Ukraine as targets of the most severe Russian cyber operations. In the cases documented by REI, the most severe Russian attacks occurred against France, Austria, and Ukraine. Part of this discrepancy is due to the respective foci of each dataset; DCID seeks out cases of cyber incidents and disputes while REI focuses on Russian electoral interference. While a majority of the REI cases include some form of Russian cyber activity, there are a few cases where only material support was provided (eg. Moldova 2014 and Belarus 1994). This discrepancy exemplifies not only the challenges of relying on open source reporting for identifying cyber influence or disruption campaigns, but also differences in defining what counts as an attack. The only country-year that appears in both datasets is Ukraine 2014. We standardized codings across the two datasets using variable definitions from respective codebooks. A severity less than or equal to 2 in DCID’s coding is synonymous in our recoding with REI’s coding for disinformation, a severity between 3 and 7 equals REI’s coding for cyberattack, and no cases in DCID have a severity greater than 7. We adopted Valeriano and Maness (2014)’s approach of sampling on intensity when there are multiple observations in a given time unit.

## Variable codings

For each incident, we code whether Russia used conventional ground forces, conventional air or sea forces, paramilitary or covert forces, cyber disruption, and information operations. By distinguishing between these five types of aggression, we obtain a clearer picture of the intensity of each case of Russian intervention. The vast majority of cases include at least some type of cyber operations. In a few cases, data limitations preclude coding of non-kinetic activity by Russia or other actors. In Moldova 2005, for example, Russia provided material support for the Communist Party but there is no credible evidence of cyber activities.

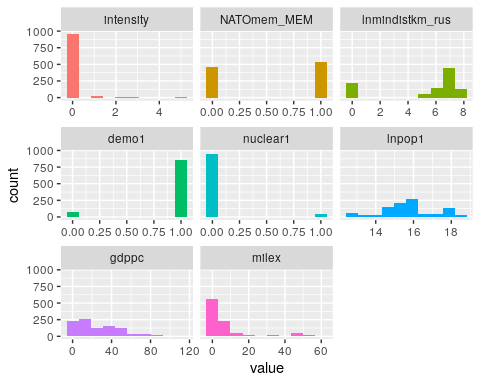
The following binary coding criteria were used for each case:

* **resp\_infoops** - Did Russia use information operations during this event? That includes propaganda, misinformation campaigns, theft of information, and other simple intrusions
* **resp\_cyberdisrup** - Did Russia use cyber attacks during this operation? That includes hacking, phishing, cyber espionage, DDoS attacks, etc. that constitute a system shut down rather than simple intrusions
* **resp\_paramil** - Did Russia use paramilitary troops during this event? Special forces, covert troops, speznatz, etc all count
* **resp\_convmil\_airsea** - Did Russia use conventional naval or air forces during this event?
* **resp\_convmil\_gro** - Did Russia use conventional ground troops like their army, artillery, tanks, etc during this event?

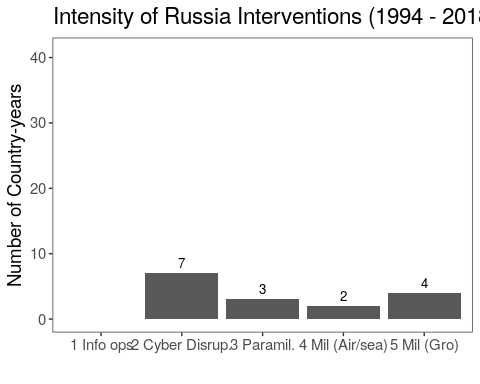
The complete dataset is provided in the appropriate .csv file. It includes sources used for the codings as well as justifications and explanations where needed.

## Summary statistics

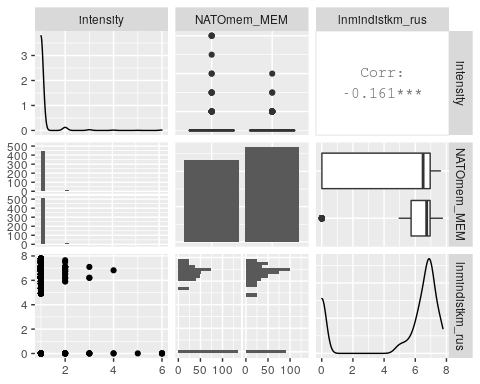
Although data was compiled on Russian intervention against all states from 1994-2018, the statistical analysis is limited to a sample from European states. In alignment with that, we present descriptive statistics of the sample used in the models provided in the main text



The distribution of our dependent variable, intensity, is shown below for the full sample (including non-European states):



The bivariate correlations between the DV and the two EVs are shown below



# Alternate model specifications

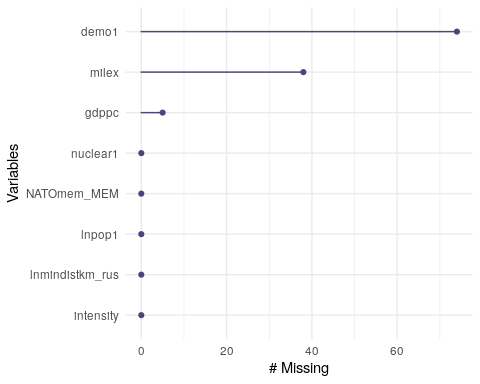
We run a set of alternate model specifications as robustness checks. Those results are shown below.

## Main regression results

The table of results in the main text shows the odds ratios. The regression results used to produce the odds ratios is shown here:

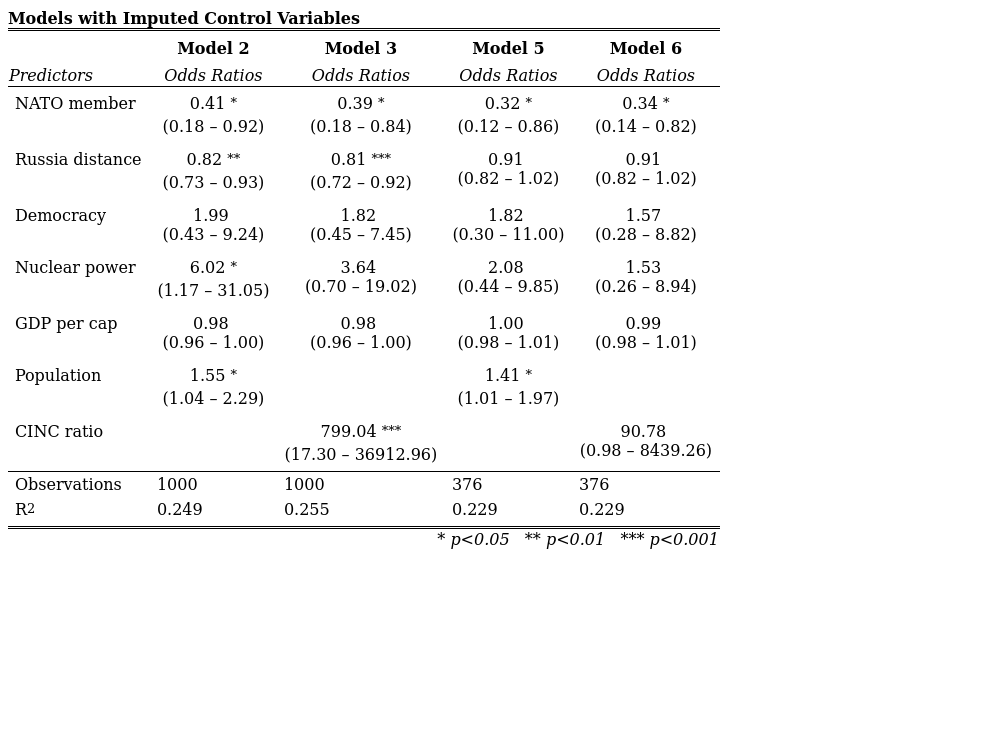
## Imputed models

Models 2 and 3 lose some observations due to missing values for control variables; primarily those not available after 2012. Variables with missing data are shown here:



We impute values into those control variable columns to avoid losing those observations and display the results below. We do not show results for models 1 and 4 in the original text since those had no missing values and are thus identical.

The full results replicating models 2, 3, 5, and 6 in the original text with imputed missing valus are shown here. Note that for models 3 and 5, CINC ratios replace the military expenditure and population variables since they are components of the composite CINC index:



# Case Study: US 2016

The main text presents case studies of Russian interventions in Estonia, Ukraine, and Georgia, which are all contiguous to Russia and former Soviet republics, and thus more comparable. Yet we expect the logic of the argument to apply more generally. Thus Russian intervention should be even more restrained against targets that are further away and more capable. Intervention in the 2016 U.S. election is consistent with this expectation.

A U.S. intelligence assessment released soon after the 2016 election concluded with “high confidence” that “Russian President Vladimir Putin ordered an influence campaign in 2016 aimed at the US presidential election. Russia’s goals were to undermine public faith in the US democratic process, denigrate Secretary Clinton, and harm her electability and potential presidency. We further assess Putin and the Russian Government developed a clear preference for President-elect Trump” (Office of the Director of National Intelligence 2017). Moscow’s influence operations might thus be described as unrestrained, even brazen, and thus motivated entirely by efficiency calculations. Yet the choice to pursue this course of action in the first place was very much constrained by the implicit deterrence posture of the United States. Russia could safely assume that the most powerful military in the world would retaliate for armed attacks against U.S. vital interests. While the United States had not designated its electoral process as “critical infrastructure” to explicitly signal that cyber interference was proscribed, Russia still had to consider the potential for American retaliation. Russia thus sought opportunities to impose costs and seek benefits while minimizing the risk of escalation. It found them through covert manipulation of democratic discourse. Indeed, Russia’s electoral interference has gone essentially unpunished by the United States to date, aside from the expulsion of some Russian intelligence officers and the application of some additional sanctions to an already heavy regime put in place after Ukraine. Of course, if Trump’s victory in 2016 or any of his administration’s subsequent policies can ever be credited to active measures by the Russian Federation, even in part, it would amount to one of the most consequential intelligence coups in history. It is just as likely, however, that the Russian campaign simply added noise to one of the most chaotic campaigns in U.S. presidential history (Gelman and Azari 2017). Russian information operations appear to be a low-cost gamble to influence an over-determined outcome.

# References

Gelman, Andrew, and Julia Azari. 2017. “19 Things We Learned from the 2016 Election.” *Statistics and Public Policy* 4 (1): 1–10. <https://doi.org/10.1080/2330443X.2017.1356775>.

Office of the Director of National Intelligence. 2017. “Assessing Russian Activities and Intentions in Recent US Elections.” Intelligence Community Assessment ICA 2017-01D. Washington, DC: National Intelligence Council. <https://www.dni.gov/files/documents/ICA_2017_01.pdf>.